A Non-Parametric Approach to Evaluate the Performance of Social Service Organizations

by

N. Alexandra Medina-Borja

Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Industrial and Systems Engineering

APPROVED:

Konstantinos Triantis, Chair

Eileen M. Van Aken

Brian M. Kleiner

James Bohland

William Seaver

April 17, 2002

Blacksburg, Virginia

Key Words: Performance Measurement System, Data Envelopment Analysis, Service Quality, Efficiency, Effectiveness, Nonprofit Organizations

Copyright 2002, N. Alexandra Medina-Borja
A Non-Parametric Approach to Evaluate the Performance of Social Service Organizations

N. Alexandra Medina-Borja

(Abstract)

Determining the best way for evaluating organizational performance is a complex problem as it involves assessment of indicators in multiple dimensions. In the case of nonprofit social service provision this evaluation needs to consider also the outcomes of the service. This research develops a performance measurement system that collects performance indicators, evaluates them and provides concrete performance improvement recommendations to decision-makers in the nonprofit sector. Three dimensions of performance are identified for social services: effectiveness or outcome achievement, service quality and efficiency.

A framework for measuring performance in four stages or nodes is advanced. The nodes represent the most important production functions for nonprofit organizations dedicated to social services. These are: (a) financial (fundraising or income generation activities); (b) capacity creation; (c) service delivery; and, (d) effectiveness. Survey instruments were developed to collect service quality and effectiveness indicators for the last two nodes. Effectiveness measures were identified following a well-structured 7-step approach to develop outcome-based objectives.

To effectively deal with this problem, the Data Envelopment Analysis (DEA) formulation was adapted to evaluate performance at each node. DEA computes performance scores, optimal target performance levels, and the performance frontier for different branches, units, or other comparable decision-making units (DMUs). Two basic formulations were developed for this framework as follows: Model I as a four stage formulation that carries the actual values of output variables of one node to the successive node, and Model II as a formulation that carries the projections —i.e. the recommended targets— from one node to the other. This last formulation assumes that the DMUs have undergone a reengineering effort and that their indicators are set at their maximum potential. Several environmental factors affecting social service provision were included in the analysis. Additionally, variable selection recommendations were developed for DEA analysis and DEA graphical reports produced.

It was concluded that decision makers could use Model I to identify performance improvement targets in each production node. The results from Model II can be used for resource planning after the targets are achieved. Finally, this performance measurement framework is being implemented at one of largest national social service agencies in the United States.
To Felipe and Camile
Acknowledgements

This is the end of a long journey and the beginning of a new one. I am amazed to realize how many things have taken place during this period of over five years and how time flew. I once read that a dissertation is always the product of the work of many, never of the doctor-to-be alone. Mine is not an exception. I have a long list of persons that contributed to my work, to my growing as a person and as a professional, to whom I will be forever indebted.

First, I want to extend my gratitude to the members of my dissertation committee: Dr. William Seaver for his statistical advice; Dr. James Bohland for his guidance during my incursion in a new field; Dr. Brian Kleiner for teaching me what rigorous experimental research is and what is not, also for his support during my years in Blacksburg and for his constant and kind concern throughout my last pregnancy. I want to give special thanks to Dr. Eileen Van Aken for her friendship, guidance and support. Eileen has been a role model and an inspiration throughout my doctoral studies. I admire her work ethics and teaching skills. She taught me many important lessons about research, work and life. When I was her teaching assistant, we were both pregnant and were trying to balance our professional goals and family. Just as Eileen never complained and was never tired, I was up and in class 3 days after my first child was born. Her example convinced me that everything is doable — if you have a loving and understanding husband and mom. I also remember several Sundays in the Management Systems Lab, away from our families, sharing few minutes of coffee break and stories about Maddie and Felipe. Further, I am able to find any file, any e-mail, any article in my computer; thanks to the “Eileen’s method” — one cannot imagine how much stress “organization” can release from anyone’s life.

I want to make a special mention of Dr. Kostas Triantis, my “master”, for being the best advisor any person could ever have had: tough, critical, skeptical, opinionated, but more importantly, caring, fair, understanding and fun. Kostas has been to me a mentor, a guardian and a friend. He has been able to get the best out of me by never being “satisfied”. He has been patient and understanding throughout the early years of my marriage, my two pregnancies, my decision to stay away in Blacksburg to raise my kids, my sudden adventure in Central America that lasted almost 2 years, and my “triumphant return to life” a year ago. I am convinced that not being for his boundless support, and for the DEA research community he has sponsored at the Northern Virginia Center, this day would not have come. We will always be friends, and I am very proud of this fact.

I also want to thank the Grado Department of Industrial and Systems Engineering at Virginia Tech for their constant and steady financial support while I was in Blacksburg and here in Northern Virginia. Also, thanks for the teaching and research opportunities that the department provided me. The ISE staff needs also to be acknowledged. I want to thank Mrs. Lovedia Cole and Marty Simpson in particular. Mrs. Cole always received me with a smile, always prompt to help me solve my administrative problems. After two years away from school she recognized my voice immediately and greeted me by my name over the phone. Marty Simpson was always kind and caring and made sure that I received my financial supports on time. Many thanks to both!

Professor Ken Harmon deserves a special mention as well. He and his wife Mary Ann were always concerned about me and my family. I remember the snowy day that I walked from the Dunn Loring metro station to his office in the old NVGC with my application package to the Master of
Engineering Administration program. He was the first kind and warm contact I had with Virginia Tech an institution that has given me so much. Thanks.

I want to thank two very special persons that have been the force enabling this accomplishment: my bosses, Dr. Jacqueline D’Alessio and Bernie Benson. They are by far the most visionary people an organization can have. With no operations research training they grasped the potential of DEA in few minutes and made the strategic and gigantic decision that my research proposal made sense. Furthermore, Jacqui and Bernie are two of the most open minded, caring and easygoing persons I have ever met in the US. Going to work everyday is for me fun, regardless of the tons of problems encompassed by actual implementation. Basically, it is fun because of Jacqui and Bernie. Working with you both is for me a honor and a blessing, and I thank you for that. I know your trust will not be betrayed.

During my stay at Tech, I have been blessed with a number of honors. In particular, I want to thank the American Society for Quality for awarding me the 1999 International Freund Scholarship for graduate studies in Quality Management. Also the International Test and Evaluation Association, for awarding their 1997 scholarship.

Life in Blacksburg was beautiful and productive. Particularly due to the people I encountered there. In particular, I want to remember my close class group, Ami, Wolfgang, Sam, Medio, Sunil, Moncef and Yonathan, and all others. Adar Kalir was also an inspiration for me.

My research skills have grown steadily over the years, but my quantitative abilities have grown exponentially. This has all been possible through the rich interaction with a group of brilliant colleagues at Tech. From my early DEA days I have been honored with the friendship and professional interaction of many Indian students after many hours at the GA’s office and through endless DEA research meetings. Thanks to John Phillips, who patiently guided my first DEA steps, Narayan Partangel, who explained how to use SAS, and how to eat chutney and predicted my son’s temperament based on Indian astrology. Chintan Sheth, for his friendship, inspirational calm in times of despair, and OR insights; Rina Sadia for her friendly and supportive attitude, and so many others that if included, my list would go on for pages and pages. Also, Dr. Barbara Hoopes and Dr. Dusan Teodorovic for their active participation in our research meetings.

I need to make a special mention to a very special person in my life. I firmly believe that God made it possible for my path to cross with that of my great friend Kalyan Pasupathy in order to finish this endeavor. As it turns out, we moved into an apartment in the same building without knowing each other. Kalyan has been the genius, the teacher and the thinker that made me understand DEA at last. I do not know any other student that knows linear programming as well as Kalyan does. Not only that, he has been my friend, my supporter and my bodyguard during the hundreds of nights I have spent in the GA’s office alone, working until the next day. He has endured my frustration when something did not run smoothly and shared my excitement when we finally got the first results. I used the algorithm he created in Excel solver for his thesis and without Kalyan’s contributions there would be no model, nor would I have kept my cool. We have shared culture, experiences, DEA insights, subway, paneer, sweets from Coimbatore, aplanachados from Quito and so many other things. I know how to say Vanakkam and he knows much more than just buzz words in Spanish. My kids call him the “best friend ever”. Kalyan, we will always have you in our hearts. Thanks.

Last but not least, I want to thank my loving and supportive family. They are my foundation, the reason I was ever able to accomplish anything. First, to my parents. My Mom Norma is responsible
of me being here because she never doubted I would be able to accomplish big things, and because she taught me that a woman has to grow professionally and be independent. Moreover, I owe her so much because of all the small and big sacrifices that she has done throughout her life for me. Particularly, for being a second Mom for Felipe and Camile, and being there for my prelim, for my proposal defense, for my pregnancies and every time I ever needed her. Gracias Baba. Te quiero mucho.

To my Dad Edgar, who has been an example of hard work, of never-ending vitality and strength and who transmitted to me my Granddad Jose’s belief that the only thing that a person will always own is the education earned. He has been the pressure needed to come back and finish. My Dad supported me to go to Brazil to study engineering, to the US to get a Master’s degree, and has always been there whenever I needed him. His daily e-mails have always kept me in touch with my own and with reality. He is a brilliant man, and yet the most down to earth person and has been my role model since childhood. I owe him everything.

I also want to thank the other members of my great family. My siblings Viviana, Edgar Jose, Amaury and Auxiliadora, and my beautiful nephew JE. I do not know a person more fortunate for having brothers and sisters as supportive, loving and caring as mine who are also so much fun to be with. The best part of finishing this endeavor is the expectation of being around you again soon. I also need to thank my “other” siblings. As loving and dedicated though no blood link exist among us, Jose Francisco and Luis, thanks for being there supporting every new crazy Borja adventure.

Finally, I thank Riqui (Armando), my husband, for keeping the house afloat during the last year, for always having high expectations, for never complaining for the mess, for always having a kiss and a hug even if it is 4 am in the morning and for sharing wild dreams for the future. His “lessons of life” made me what I am now. Many years ago, when we were dating he taught me “how to write”, and all the nuances of American culture that I would have never figured out by myself. I specially thank you for your mix of total trust and annoying innocence that obliged me to get things done even when I thought them to be unachievable. I love you.

Finally, one day in the near future, Felipe and Camile will be able to read this acknowledgements section and perhaps will be the only section that they will ever read. Hence, I want them to know that I spent so many days and nights away from you because Mommy had a dream, and when the dream is plausible and good, one should always pursue it. I did this for me, but also for us, and from now on I will compensate for all those hours of absence. Thanks principes for all your unconditional, cheering and loud love.
Table of Contents

List of Figures .................................................................................................................. xii
List of Tables ................................................................................................................... xiii

CHAPTER 1 Introduction and Scope of the Research ....................................................... 1
1.1 Background .................................................................................................................. 2
  1.1.1 What is Performance for a Social Service Agency? .............................................. 4
  1.1.1.1 Performance Measurement Dimensions ......................................................... 6
        Effectiveness .............................................................................................................. 6
        Technical Efficiency .................................................................................................. 7
        Service Quality ......................................................................................................... 8
  1.1.2 What is DEA .......................................................................................................... 9
1.2 The Problem and Its Setting ...................................................................................... 10
  1.2.1 Problem Statement ............................................................................................. 10
  1.2.2 Relevance of the Problem .................................................................................. 11
    1.2.2.1 Relevance of the Problem for Industrial and Systems Engineering ............ 12
1.3 Research Purpose ..................................................................................................... 13
1.4 Global Research Model ........................................................................................... 13
1.5 Research Objectives ................................................................................................ 17
1.6 Conceptual Research Model .................................................................................... 17
1.7 Research Questions .................................................................................................. 21
1.8 Hypotheses .............................................................................................................. 21
1.9 Operational Research Model .................................................................................... 22
1.10 Contribution of this Research .............................................................................. 24
    1.10.1 Outputs and Desired Outcomes of this Research ............................................. 24
1.11 Assumptions and Delimitations ............................................................................ 25
1.12 Organization of this Document ............................................................................. 26

CHAPTER 2 Review of Literature .................................................................................. 29
2.1 Background .............................................................................................................. 29
2.2 Performance Measurement Systems for Nonprofit Organizations ....................... 31
    2.2.1.1 The effects of the Environment in organizational performance .................... 36
    2.2.2 Performance Measurement Dimensions .......................................................... 37
    2.2.2.1 Effectiveness .................................................................................................... 37
    2.2.2.2 Productivity and Technical Efficiency .............................................................. 39
    2.2.2.3 Service Quality and Customer Satisfaction ..................................................... 40
        Models to Measure Service Quality ......................................................................... 41
    2.2.3 Relationships among performance dimensions ................................................. 45
5.1.3 Results for Node 3: Service Delivery ................................................................. 218
  5.1.3.1. Results with Volume Data for Node 3 .......................................................... 219
  5.1.3.2. Results with Standardized Data: Node 3 ....................................................... 222
5.1.4 Node 4: Customer Effectiveness ................................................................. 224
  5.1.4.1. Results for Node 4 with Service Quality Indexes ........................................... 225
  5.1.4.2. Results for Node 4 Using Transformed Inputs to Bad Quality Index .............. 229

5.2 Application of Model II: Considering the Optimal Level of the Outputs from the
Previous Node in Disaster Services ............................................................................. 233
  5.2.1 Node 2: Capacity Creation Considering the Optimal Level of the Outputs of the Financial
Node .............................................................................................................................. 233
  5.2.2 Node 3: Service Delivery Considering the Optimal Level of the Outputs of the Capacity
Creation Node .................................................................................................................. 235
  5.2.3 Node 4: Effectiveness Considering the Optimal Level of the Service Quality Indicators of
the Service Delivery Node ............................................................................................ 238

5.3 Application of Model I in the Total Field Office Operation .................................. 242
  5.3.1 Node 1 (Financial Node) Results for Total Field Office Operations ..................... 247
  5.3.2 Node 2 (Capacity Creation) Results for Total Field Office Operations ................. 250
  5.3.3 Node 3: (Service Delivery) Results for Total Field Office Operations ..................... 254
  5.3.4 Node 4 (Customer Effectiveness Node) Results for Total Field Office Operations .... 258

5.4 A Note about the Inclusion of Environmental Variables in the Model .................... 262

5.5 Conclusions about Models .................................................................................. 264

5.6 Development of an Algorithm for the New DEA Formulation ............................... 267

5.7 Deployment to Field Offices: Performance Improvement Reports, Service Quality
and Effectiveness Reports ............................................................................................... 267

CHAPTER 6 Conclusions and Recommendations for Future Research .................. 273

6.1 Summary .............................................................................................................. 273
  6.1.1 Validity and Generalizability of this Study ............................................................. 277
    6.1.1.1. Content and Face Validity of the Study .......................................................... 278
    6.1.1.2. Mathematical Model Validation ................................................................. 279
    6.1.1.3. Survey Instruments: Construct Validity and Reliability ................................ 282

6.2 Contribution of this Research ............................................................................. 284

6.3 Limitations of this Research ............................................................................... 287
  6.3.1 Limitations of this Application in the Nonprofit Sector ....................................... 287
  6.3.2 Mathematical Model Limitations ........................................................................ 288
  6.3.3 Data limitations .................................................................................................. 288
  6.3.4 Limitations of the Survey Instruments ............................................................... 289

6.4 Recommendations for Future research ............................................................... 290

Bibliographic References .......................................................................................... 293

Appendix A: The Outcome-Based Objectives Development Process ....................... 316

Appendix B: Service Quality and Effectiveness Disaster Services Final Survey and
Guidelines for Sampling and Survey Administration .................................................. 321
Health and Safety Course Takers: ................................................................. 325
Community Services .................................................................................... 325
International Services .................................................................................. 326
Mail Survey Checklist .................................................................................. 334

Appendix C: Data Sets .................................................................................. 336
APPENDIX D: Excel Solver Codes ................................................................. 343
VITA .............................................................................................................. 360
List of Figures

Figure 1.1 Overlapping Areas of the Two Approaches to Measure Performance in Social Service Organizations..............................................................................................................................................3
Figure 1-2. Performance Measurement Model for Social Services Framework adopted by the Agency’s Outcome Measurement Project (adapted from Sink, 1985) .........................................................15
Figure 1.3. Elements of the Performance Control System for Social Service Organizations .................................................................16
Figure 1.4. Hypothesized Model of the Relationships between Dimensions of Performance in Social Service Organizations .........................................................................................................................20
Figure 1-5. Overall Operational Research Model ..............................................................................................................................................22
Figure 1-6. Operational Research Model: Relationship among Variables ..............................................................................................................23
Figure 2-1. Performance Relationships in Hospitals (from Kunst and Lemmink, 2000) .................................................................46
Figure 2-2. Interrelationship between Organizational Performance Criteria (adapted from Sink & Tuttle, 1989) ..............................................................................................................................................47
Figure 2-3. The integration of the Balanced Scorecard with the Performance Pyramid (adapted from Rouse et al., 1997). ..............................................................................................................................................49
Figure 2-4. Main Components of the Process of Value Generation (Sola & Prior, 2001) ..............................................................................................................................................................................................................88
Figure 2-5. Production System for District Schools ..............................................................................................................................................90
Figure 2-6. Output-Maximizing Orientation .................................................................................................................................................................................................................................................................................................................90
Figure 2-7. The Network Technology .............................................................................................................................................................................................................................................................................................................119
Figure 2-8. Capabilities, Service Quality and Performance in the Banking Industry (Athanassopoulos 2000) ...........................................................................................................................................................................................................................................................................................................123
Figure 2-9. Two -Stage Process of Assessing the Effectiveness of Bank Branch Operations ................. 124
Figure 2-10. Decomposition of Malmquimist productivity index .................................................................................................................................................................................................128
Figure 2-11. The Production of Social Services (target population A). (adapted from Byrnes et al. 1997) .................................................................................................................................................................................................................................................................................................................131
Figure 3-1. Deming Diagram for the Financial Node .................................................................................................................................................................................................................................................................................................................147
Figure 3-2. Deming Diagram for the Capacity Creation Node .................................................................................................................................................................................................................................................................................................................148
Figure 3-3. Deming Diagram for the Service Delivery Node .................................................................................................................................................................................................................................................................................................................149
Figure 3-4. Deming Diagram for the Customer Node .................................................................................................................................................................................................................................................................................................................150
Figure 3-5. Conceptual Model of the Production Technology of Nonprofit Organizations ................. 150
Figure 5.1. Sample of Trend Analysis Report .................................................................................................................................................................................................................................................................................................................270
Figure 5.2 Samples of Service Quality and Effectiveness Report.................................................................................................................................................................................................................................................................................................................272

xii
List of Tables

Table 1-1. Outputs and Outcomes for Social Services ................................................................. 7
Table 1-2. Hypothesized Relationships between Effectiveness, Technical Efficiency and Customer Satisfaction ................................................................................................................. 18
Table 2-1. Mathematical Notation Used in this document for DEA ............................................. 86
Table 2-2. Inputs and Outputs used by Tankersley (2000) to Determine Levels of Service Quality Dimensions ................................................................................................................................. 129
Table 4.1 Performance Indicators Selected to Measure Each Performance Dimension .............. 181
Table 4.2. Inputs and Outputs Used for the First Data Set regarding Disaster Services collected by FOCUS ...................................................................................................................................................... 182
Table 4-3. Sample of Labels Used in Organizational Research (Bruner et al, 2001; NCS Pearson, 2002) .................................................................................................................................................................. 191
Table 4.4 Constructs and Items for the Customer & Outcome Survey ........................................ 196
Table 4.5 Results of Internal Consistency and Reliability Analysis Of Service Quality And Effectiveness Surveys ......................................................................................................................................................... 197
Table 4.6 Example of the Calculation of Service Quality Index .................................................. 203
Table 4.7 Hypothesized Relationships between Effectiveness, Technical Efficiency and Customer Satisfaction ................................................................................................................................. 206
Table 4.8 Results with Volume Data for Node 1 ........................................................................ 210
Table 4.9 Results with Standardized Data for Node 1 ................................................................ 211
Table 4.10 Performance Score and Peers for Node 2 with Volume Data .................................... 213
Table 4.11 Performance Scores for Node 2 .................................................................................. 214
Table 5.1. Demographic Characteristics of DMUs in the Sample ............................................... 217
Table 5.2. Results with Volume Data for Node 1 ........................................................................ 220
Table 5.3. Results with Standardized Data for Node 1 ............................................................... 221
Table 5.4. Performance Score and Peers for Node 2 with Volume Data .................................... 222
Table 5.5. Performance Scores for Node 2 with Standardized Data .......................................... 223
Table 5.6. Performance Scores for Node 2 with Standardized Data ........................................... 224
Table 5.7. Performance Scores for Node 2 with Standardized Data ........................................... 225
Table 5.8. Performance Scores for Node 2 with Standardized Data ........................................... 226
Table 5.9. Service Quality Inputs for Node 4 ............................................................................... 227
Table 5.10. Actual and Target Values of Outcomes in Node 4 with Service Quality as Input ...... 228
Table 5.11. Performance Scores for Node 4 Using Transformations in the Input Side ................ 230
Table 5.12. Performance Scores for Node 4 Using Transformations in the Output Side ............. 230
Table 5.13. Performance Scores for Node 4 Using Transformations in the Input Side ............. 231
Table 5.14. Performance Scores for Node 4 Using Transformations in the Output Side ............. 231
Table 5.15. Performance Scores and Peers for Standardized Data: Node 2 ............................... 232
Table 5.16. Performance Scores and Peers for Standardized Data: Node 2 ............................... 232
Table 5.17. Performance Scores and Peers for Standardized Data: Node 2 ............................... 233
Table 5.18. Performance Scores and Peers for Standardized Data: Node 2 ............................... 234
Table 5.19. Performance Scores and Peers for Standardized Data: Node 2 ............................... 235
Table 5.20. Performance Scores and Peers for Standardized Data: Node 2 ............................... 236
Table 5.21. Performance Scores and Peers for Standardized Data: Node 2 ............................... 237
Table 5.22. Performance Scores and Peers for Standardized Data: Node 2 ............................... 237
Table 5.23. Performance Scores and Peers for Standardized Data: Node 2 ............................... 238
Table 5.24. Performance Scores and Peers for Standardized Data: Node 2 ............................... 239
Table 5.25. Performance Scores and Peers for Standardized Data: Node 2 ............................... 240
Table 5.26. Performance Scores and Peers for Standardized Data: Node 2 ............................... 241
Table 5.27. Performance Scores and Peers for Standardized Data: Node 2 ............................... 242
Table 5.28. Performance Scores and Peers for Standardized Data: Node 2 ............................... 243
Table 5.29. Performance Scores and Peers for Standardized Data: Node 2 ............................... 244
Table 5.30. Performance Scores and Peers for Standardized Data: Node 2 ............................... 245
Table 5.31. Performance Scores and Peers for Standardized Data: Node 2 ............................... 246
Table 5.32. Performance Scores and Peers for Standardized Data: Node 2 ............................... 247
Table 5.33. Performance Scores and Peers for Standardized Data: Node 2 ............................... 248
Table 5.34. Performance Scores and Peers for Standardized Data: Node 2 ............................... 249
Table 5.35. Performance Scores and Peers for Standardized Data: Node 2 ............................... 250
Table 5.36. Performance Scores and Peers for Standardized Data: Node 2 ............................... 251
Table 5.37. Performance Scores and Peers for Standardized Data: Node 2 ............................... 252
Table 5.38. Performance Scores and Peers for Standardized Data: Node 2 ............................... 253
Table 5.39. Performance Scores and Peers for Standardized Data: Node 2 ............................... 254
Table 5.40. Performance Scores and Peers for Standardized Data: Node 2 ............................... 255
Table 5.41. Performance Scores and Peers for Standardized Data: Node 2 ............................... 256
Table 5.42. Performance Scores and Peers for Standardized Data: Node 2 ............................... 257
Table 5.43. Performance Scores and Peers for Standardized Data: Node 2 ............................... 258
Table 5.44. Performance Scores and Peers for Standardized Data: Node 2 ............................... 259
Table 5.45. Performance Scores and Peers for Standardized Data: Node 2 ............................... 260
Table 5.46. Performance Scores and Peers for Standardized Data: Node 2 ............................... 261
Table 5.47. Performance Scores and Peers for Standardized Data: Node 2 ............................... 262
Table 5.48. Performance Scores and Peers for Standardized Data: Node 2 ............................... 263
Table 5.49. Performance Scores and Peers for Standardized Data: Node 2 ............................... 264
Table 5.50. Performance Scores and Peers for Standardized Data: Node 2 ............................... 265
Table 5.28 List of Inputs and Outputs used for the Evaluation of Field Office’s Effectiveness ..... 246
Table 5.29. List of Non-controllable Environmental Variables Used in the Evaluation of Total Field Office’s Operations............................................................................................................................... 246
Table 5.30 Results in Node 1 for Total Field Office Operations.............................................. 248
Table 5.31 Peers of Under Performing Units in Node 1 for Total Field Office Operations ....... 249
Table 5.32 Environmental Variables for the Peers of Field Office 7038.................................................... 249
Table 5.33 Scores from Node 2 .................................................................................................................... 251
Table 5.34 Results from Node 2 for Total Field Office Operations ...................................................... 252
Table 5.35 Peers in Node 2 for Total Field Office Operations ............................................................... 253
Table 5.36 Results from Node 3 for Total Field Office Operations ...................................................... 256
Table 5.37 Peers in Node 3 for Total Field office Operations ............................................................... 258
Table 5.38 Results from Node 4 for Total Field Office Operations ...................................................... 260
Table 5.39 Peers in Node 4 for Total Field office operations ................................................................. 261
Table 5.40 Performance Scores for Node 1 without Environmental Harshness ................................. 262
Table 5.41. Peers for Inefficient Units in Node 1 without Environmental Harshness ....................... 264
Table 5.42. Performance Report for FIELD OFFICE A....................................................................... 269
Table 6.1 Conclusions about Models ........................................................................................................... 286
CHAPTER 1  Introduction and Scope of the Research

Evaluating performance has increasingly gained strategic importance for the nonprofit sector of the economy. As a result, during the last decade, several performance measurement initiatives have been launched with this purpose (e.g. ASPA, 1992; Brown & Pyers, 1988; GPRA, 1993; United Way of America, 1996; Wholey & Hatry, 1992). These initiatives are a natural reaction to new evaluation and social pressures that demand greater accountability from these organizations. Thus, measuring and demonstrating results has become a question of survivability for many nonprofit organizations (e.g. Hatry et al., 1992; Kaplan, 2001; Light, 2000; Maurrasse, 2002; Medina-Borja and Triantis, 2001; Schallock, 1995; Triantis & Medina-Borja, 1996). The reasons are several. Nonprofit organizations providing social services are being impacted by the growing number of new agencies, reduced number of donors, increased budget cuts, and increased challenges emanating from emergent social and health problems (Jennings & Saunders, 1993; Lovelock & Weinberg, 1984; Rees, 1998; Schallock 1995). Moreover, most stakeholders associated with nonprofit organizations want to know if the programs or services they support are having a positive, visible, and consequential impact on their communities (Magura and Moses, 1986; Schallock, 1995).

Complying with these new challenges is not an easy task for social service organizations as performance is still a concept not clearly understood (Magura & Moses 1987; Schallock 1995). These organizations need increasingly not only to demonstrate the effectiveness of the programs they provide —i.e. the accomplishment of the results sought originally by their charter— but also are required to demonstrate stewardship in the use of resources endowed to them (e.g. Light, 2000; Ospina et al., 2002). Hence, an ideal performance measurement system for social service organizations should account for the outcomes or results obtained by their programs in addition to reporting on other important process performance measures that provide a general overview of their
program operations and accountability (Hargreaves and Attkisson, 1978; Kaplan, 2001; Kravchuck & Schack, 1996; Nyhan & Martin, 1999).

According to Ospina et al (2002), dominant taxonomies\(^1\) classify nonprofit organizations based either on fiscal considerations or on the activities such organizations perform (Boris and Mosher-Williams, 1998; Salamon and Anheimer, 1996). I will focus on developing a model for nonprofit charitable organizations dedicated to the delivery of social programs of diverse nature (Major groups V and VII of human services, public and societal benefit). Although the models developed here can be extrapolated to other types of nonprofits as well. I am thus referring to this type of charitable organizations whenever alluding to nonprofits in the context of the conceptual and mathematical models developed in this study, unless otherwise specified.

This research advances an integrated, multi-dimensional performance measurement system for evaluating performance in nonprofit organizations. A non-parametric linear programming framework has been devised and implemented for evaluating the metrics collected through this system. The mathematical formulation for this framework is an adaptation of a linear programming technique called Data Envelopment Analysis (DEA) (Charnes et al, 1978). DEA has been used to calculate performance scores and the best practice frontier for a variety of different organizational units (decision-making units / DMUs) operating in a variety of organizational environments. Furthermore, an application of this performance measurement system and the DEA technique is illustrated at a large social service organization (referred in this document as the “agency”.)

1.1 Background

Evaluating nonprofit performance is not new. Government and private social service agencies have traditionally measured their performance using two disconnected approaches and measurement systems (Cameron, 1981; Connolly, Conlon

\(^1\) The Internal Revenue Service classification, the National Taxonomy of Exempt Entities, and the International Classification of Nonprofit Organizations.
and Deutsch, 1980). One of these measurement systems is based on a managerial perspective of performance where data related to budgetability, productivity, service quality and customer satisfaction are regularly collected (e.g. Forbes, 1998; Gruber, 1981; Hatry, 1980; Hayes & Millar, 1990; Kleiner, 1994). The second approach is the one used in the evaluation research field, and is concerned with the experimental measurement of program impact and other process variables influencing the extent of outcome achievement (e.g. Grinnell, 1988; Lipsey, 1993; Littell, 1986; Rossi & Freeman, 1993). Dimensions such as the quality of the service provider are investigated to account for their influence, for example, in the achievement of clients’ outcomes.

One of the problems with this performance evaluation practice is that scientific program evaluation approaches are not integrated into the day-to-day operations of the organizations being evaluated. In fact, the evaluation of impact or results is usually treated as a completely independent approach than the one used to measure productivity, service quality and financial health indicators. An analysis of both measurement systems—managerial and program evaluation—will reveal that there is an overlapping area between them since there are common variables or dimensions of performance being measured by both systems such as quality of service, service provider characteristics, etc. This overlapping area is depicted in Figure 1.1.

![Figure 1.1 Overlapping Areas of the Two Approaches to Measure Performance in Social Service Organizations](image-url)
Because of the shortcomings of these two disconnected approaches new evaluation trends have emerged (Lohman, 1999; Vogt, 2000; Volkman, 1999). The theoretical foundation behind new measurement frameworks stress that the ultimate objective of social service agencies is not only to produce a unit of output —e.g. number of services provided, number of students that participated in a training program, number of clients served, etc. — but to produce an outcome or a positive change on their clients’ lives. Of course, stewardship in the use of funds is also an important performance aspect.

The managerial approach needs to make use of program evaluation theory to develop outcome indicators for social programs and then incorporate them into the measurement and evaluation framework of other operational dimensions. The latest performance measurement movements at the government and nonprofit levels have reiterated the need for agencies to develop a performance evaluation system that is continuous, internal to the organization and that provides decision makers with up-to-date information on program performance based on reliable and significant decision-making dimensions like efficiency and effectiveness.

The GPRA and the United Way's outcome-based evaluation initiatives, for example, integrate outcome measures in the organizational strategic plan by the way of outcome-based objectives. In fact, the measurement of effectiveness indicators (outcome achievement) is now being required by most initiatives (e.g. Hargreaves and Attkisson, 1978; Magura & Moses, 1986; Martin & Kettner, 1996, 1997). This phenomenon has become especially important since the early 90s (GPRA, 1993; Medina-Borja and Triantis, 2001; Triantis & Medina-Borja, 1996; United Way of America, 1996).

1.1.1 What is Performance for a Social Service Agency?

There is agreement in the management and evaluation literature that performance is a multi-dimensional construct (e.g. Kaplan, 2001; Kaplan & Norton, 1992; Sink, 1985; Brown, 1996). Recent performance measurement efforts combine different performance dimensions such as the balance scorecard (Kaplan, 2001; Kaplan & Norton, 1992, 1996), and the family of measures approach (Brown, 1996; Provost and
Leddick, 1993). Other authors (Sink, 1985; Sink & Smith, 1992; Sink & Morris, 1995) have presented seven interrelated and interdependent performance criteria for an organizational system, namely effectiveness, efficiency, productivity, quality, quality of work life, innovation, and profitability/budgetability. All of these approaches are being translated from the private sector to government and nonprofit organizations. The Reinventing Government movement and the National Performance Review initiatives have become the catalysts for many of these initiatives (Gore, 1993; NPR, 1997).

Kaplan (2001) in particular, adapted his balanced scorecard approach for the nonprofit sector by focusing on shareholders objectives driving the organization's strategy — the organization's mission drives it — rather than the organization’s financial growth. As such, measures in the financial, customer and internal perspectives answer to the following questions:

- "To achieve our vision, how must we look to our customers/ recipients?"
- "To satisfy our customers, financial donors, and mission, at which business processes must we excel?"
- "If we succeed, how will we look to our financial donors?"

All three questions provide the framework to establish which dimensions define the construct "performance" for a nonprofit organization. Organizations need to address these questions to achieve external growth (funds raised); to provide internal stability (balance income and expenses); to achieve community building (increase the amount of funds that go into services), customer satisfaction, market growth, service quality and effectiveness or community building, to name just a few.

Performance in the context of the present research work considers three dimensions of performance. These have been suggested by several authors as fundamental for social programs: service quality, efficiency and effectiveness (e.g. Hargreaves and Attkisson, 1978; Nyhan & Martin, 1999; Urban Institute, 1980).
1.1.1.1. Performance Measurement Dimensions

The exact definitions of effectiveness, service quality, and efficiency have been a matter of extensive debate. I will briefly introduce all these three dimensions for the purpose of this research.

Effectiveness

Effectiveness measures the extent to which the program achieves the organization’s goals and objectives while at the same time fulfills its mission statement (Schalock, 1995). A program is effective to the extent that it accomplishes what it was designed to accomplish. I will use effectiveness to explain the extent of outcome achievement of a social program or service.

There is a lot of confusion regarding the difference between outputs and outcomes. The next section clarifies these differences.

Outcomes

While outcomes are different from outputs, there is a relationship that cannot be ignored. In most cases, outcomes are the results obtained after the provision of outputs (Triantis & Medina-Borja, 1996). In general, outcomes for the social and health services are related to a perceived benefit or improvement of the quality-of-life of the client as a direct result from the program’s services and projects.

I will use the approach adopted by the United Way (1996) to define “outcomes.” The United Way approach differentiates outcomes according to the nature of the service being provided. The first definition applies to those programs aimed at producing a change in the behavior, knowledge, attitude, or condition of the target population. “Outcomes” for these services are conditions reached by the population because of the presence of a program. They are the after-effects of a program’s intervention.

The second definition applies to those programs whose purpose is to sustain or process services for clients, rather than to change them. In such cases where
the desired effect is the appropriate delivery of the service itself, agencies can use the quality of program operations as an outcome. In this case, client satisfaction is a valid alternative descriptor of outcomes. Table 1-1 provides examples that illustrate the difference between outcomes and outputs for social services.

<table>
<thead>
<tr>
<th>OUTPUTS</th>
<th>OUTCOMES</th>
</tr>
</thead>
</table>
| # of participants enrolled in an AIDS prevention program | ♦ Change in intentions and risky sexual behavior of the participants (immediate outcome)  
♦ Reduced mortality due to AIDS (long-term outcome) |
| # of shelters provided during a disaster | ♦ Reduced homelessness after a disaster (intermediate outcome) |
| # of emergency communications provided | ♦ Reduced anxiety due to long-term separation from a loved one (immediate outcome) |

**Table 1-1. Outputs and Outcomes for Social Services**

**Technical Efficiency**

Any system can be represented with inputs (resources), outputs (units of service), and processes that transform inputs to outputs. This system is a production system even if the goods it produces are services. The efficiency of such a system can be expressed in the simplest case as Output(s)/Input(s). However, most processes require multiple inputs to produce multiple outputs. This multiple input/output representation requires that these input/output variables be combined into an aggregated scalar form (Girod, 1996).

Debreu (1951) and later Farrell (1957) developed indices of technical efficiency for comparing efficient and inefficient production units that takes into account
multiple input-output configurations. This definition will be used throughout this research:

\[
TE = \frac{\text{aggregated measure of outputs}}{\text{aggregated measure of inputs}} \quad (1.1)
\]

**Service Quality**

Nowadays, quality performance measures routinely take their place alongside more traditional efficiency and effectiveness performance measures (Brinkerhoff & Dressler, 1990; Martin & Kettner, 1996, 1997). When posed with the question of what does quality mean for services, one has to go back to one of the most accepted "definitions of quality." It is well recognized that the *quality* of any product is a concept perceived by the consumer of the product. Juran’s concept suggests that quality as “fitness for use” underlies this precept. Equivalently, a service is perceived of as having high quality or not to the extent that the service organization meets its clients’ needs in the most appropriate way. The quality of service organizations is directly related to customer satisfaction (Triantis and Medina-Borja, 1996) and one can claim that client satisfaction for service organizations is a good proxy measure of the *quality of a service* but not the only one.

It is also possible to investigate and define a series of desirable attributes or characteristics for each service or product that the customer values and in this way define the construct "service quality". The presence of all these attributes is perceived as "quality of service" which in turn generates client satisfaction. The actual conceptual model of service quality and customer satisfaction has been the object of a great deal of research and debate. An extensive review of this field is presented in Chapter 2 of this document. The service quality construct investigated and used during the course of this
research is further explained and developed in Chapters 2 and 4 of this dissertation. In the next section a brief overview of Data Envelopment Analysis is provided.

1.1.2 What is DEA

Data envelopment analysis (DEA) was developed by Charnes, Cooper and Rhodes (1978) to compute the technical efficiency performance of organizational units or DMUs. DEA is based on a linear programming approach to assess how efficiently a firm, agency, program site, or other such decision-making unit utilizes its resources (inputs) and produces its outputs. DEA uses a set of DMUs to construct a production efficiency frontier consisting of all efficient DMUs in the sample. Each DMU is compared to all other DMUs (Silkman, 1986).

The original DEA formulation (Charnes et al, 1978) has been modified and adapted afterwards in hundreds of papers and publications (e.g., Banker et al, 1984, Banker and Morey, 1986). Efficiency scores are obtained by computing the distance between each DMU and the production frontier. This means that the benchmark frontier envelops the data and is an extreme, that is, maximum or minimum values, as opposed to an average representation of performance.

DEA circumvents many of the standard economic problems associated with nonprofit input-output analyses. First, it does not require that the production function have a specific functional form. Second, resulting efficiency scores are based on extreme methods (Banker et al, 1989). Third, input prices are not required. In most nonprofit settings it is not possible to obtain the price of the social services provided. Finally, inputs and outputs can be of any form and units.

DEA has been successfully used to develop performance standards and evaluation in such diverse nonprofit settings as education (e.g. Callen, 1991; Chalos and Cherian, 1995; Charnes et al., 1981; Kao, 1994; Lovell, Walters and Wood, 1994); healthcare (e.g. Banker et al., 1986; Juras, Kaspin and Martin, 1994; Nunamaker, 1983; Sherman, 1984, 1986) and municipal police departments (Parks, 1991). A complete
DEA is useful computing relative traditional technical efficiency scores and new performance measures of effectiveness and service quality related to social services. In most cases, nonprofit organizations regularly collect a variety of data of several forms: categorical, continuous, and also social and service delivery ratios as indicators of performance. A modification of DEA to account for these data is appropriate for evaluating the performance of social service units. In general, DEA models compute relative performance scores, provide information on performance targets for each DMU, and provide a set of peers to which the DMU under study can most meaningfully compare itself to. A more exhaustive review of DEA has been included in Chapter 2 of this document.

The following section introduces the research problem for this research and its relevance.

1.2 The Problem and Its Setting

1.2.1 Problem Statement

Most researchers agree that a large number of social service programs have been designed with inadequate evaluation capabilities (Cronchbach, 1980). In many cases, even when experimental or quasi-experimental evaluation designs are in place, these are motivated by the need to show proof of causality. Furthermore, it is almost impossible to rigorously demonstrate that these programs have had any effect (e.g. Campbell, 1969; Coursey, 1977; Freeman & Solomon, 1979; Hargreaves and Attkisson, 1978; Rossi and Freeman, 1993). In addition, most of these designs try to account for impact but rarely take into consideration other important dimensions such as technical efficiency, customer satisfaction and service quality, among others. Finally, the implementation of these designs is generally expensive, sporadic and the analysis of the data collected requires specialized expertise.
For nonprofit organizations it has become inherently difficult to measure program effectiveness and to relate it to other more traditional performance dimensions (Moore, 1995; Schallock, 1995). But, how does one measure social service outcomes? And how can those outcomes be related to the costs incurred to produce them? To address these questions several performance evaluation initiatives have been launched in the last decade for measuring performance in social service agencies (National Performance Review, 1997; United Way of America, 1996). Nonetheless, there still is a lack of a framework that integrates and evaluates different performance dimensions relevant for the nonprofit sector, particularly when dealing with outcomes. A review of the current literature suggests that there are too many approaches that use inconsistent terminologies and lack adequate analytical tools for evaluating performance. Moreover, there is also a need to integrate all relevant dimensions into a common score for providing managers with the ability to benchmark their services against other comparable service providers and more importantly to assist these managers in their decision making process.

1.2.2 Relevance of the Problem

If the current trend for evaluating performance persists, the need to know how to successfully design and implement performance measurement systems will continue to increase. The difficulty in developing measurement tools that account for relevant variables that affect important dimensions of performance, including service outcomes in addition to the lack of analytical tools that evaluate the data collected through these systems makes this research effort extremely relevant.

There have been few attempts to combine linear programming approaches to performance measurement in the social services arena. Although dozens of applications in diverse nonprofit settings have clearly demonstrated the potential applicability of DEA to nonprofit/government sector systems, there are still few studies that examine how outcome indicators can be used to develop DEA effectiveness measures (Byrnes et al, 1997; Chalos and Cherian, 1995). If by applying a linear
programming approach one can point out to the possible causes for reduced program effectiveness and efficiency, identify other operating units that can be considered peers, and prescribe quantifiable performance targets, then the contribution of this research effort is even more significant.

The DEA framework developed in this research will also help identify “best practices” across sub-units (field agency offices). Also, the increased understanding of these “best practices” will assist decision-makers, both in government and in private sector, to make wiser choices about the future of social services. The current challenges posed by nonprofit downsizing and restructuring could also be better managed by having more effective performance information systems in place. Furthermore, the decision making process would be better informed so that decisions are made based on what could legitimately be phase out without adversely jeopardizing community and social needs.

1.2.2.1. Relevance of the Problem for Industrial and Systems Engineering

This research effort relates to Industrial Engineering (IE) in different ways. First, it is clearly in the realm of performance measurement of organizations, a field of concern for IEs. Furthermore, DEA is a linear programming tool which is one of the core IE’s techniques. Second, the methodology for definition of outcome-based objectives developed by Triantis and Medina-Borja (1996) is connected to the strategic planning function of organizations (see Appendix A), another action area for IE (Sink & Tuttle, 1989; Sink & Morris, 1995). In general terms, measurement needs to be viewed as a key step in a strategic management process, not the reverse (Sink & Smith, 1992). The process of identifying expected outcomes for social programs and incorporating them in the program philosophy and rationale of an organization is intrinsically related to its strategic planning process.

Third, this research relates to total quality management (TQM). One can relate the information gathered through this research to a TQM system. IE scholars study methodologies and best practices of TQM, which include customer satisfaction and service quality theories for all types of organizations, including the nonprofit sector. The
data transformed to performance improvement recommendations feed the PDSA cycle of TQM interventions.

This research represents a unique attempt to define an interface between the performance measurement and evaluation fields thus creating a unique opportunity for social workers and industrial engineers to join in a multidisciplinary research effort.

The following sections 1.3 through 1.6 present the research purpose, global, operational and conceptual research models devised for this research, in addition to the description of the research objectives.

### 1.3 Research Purpose

The purpose of this research effort is to develop an analytical framework to enable nonprofit organizations (particularly social services agencies) to measure, evaluate and benchmark the performance of their services.

### 1.4 Global Research Model

Figure 1-2 is an adaptation of Sink’s model for performance measurement (1985) and provides a graphical representation of how performance might be measured in social services organizations.

It depicts two process measures: technical efficiency and service quality, including customer satisfaction plus an effectiveness measure: outcome achievement. Based on this Input/Output analysis, a tri-dimensional performance system that argues for the importance of measuring outcomes in addition to other process related measures (e.g. Magura & Moses, 1986; Rossi & Freeman, 1993; Schalock, 1995) has been developed. This multi-dimensional approach satisfies three important precepts: that an agency needs to not only achieve its goals or aimed outcomes (being therefore effective), but for doing so in an efficient manner, making a wise use of resources (technical efficiency) while delighting the customer or client.
A truly high performing agency or unit, will, therefore, achieve optimum performance balancing these three dimensions. The performance measurement system developed in this research will make use of what Hargreaves and Attkisson call “routinely monitored client change information” to collect outcomes and customer satisfaction data through large reporting systems.
Figure 1-2. Performance Measurement Model for Social Services Framework adopted by the Agency’s Outcome Measurement Project (adapted from Sink, 1985)
Figure 1-3 displays three elements of a performance control system for social service organizations. The vertices of the triangle represent the metrics obtained from the process, the sides of the triangle represent the methodology or procedure employed to obtain the metrics. This model hypothesizes that all three performance dimensions are interrelated, but as I will discuss in the next section, empirical studies have not shown conclusive evidence about the direction of these relationships (Chalos and Cherian, 1995; Hayes and Millar, 1990).

![Figure 1.3. Elements of the Performance Control System for Social Service Organizations](image-url)
1.5 Research Objectives

The research objectives for this research study follow:

O1: To define performance for social service organizations and to clarify the definitions of effectiveness, efficiency and quality that are best suited for the DEA analysis.

O2: To design and test a performance measurement system for social service organizations. This performance measurement system will include:

a. the design of a data collection tool to collect data with respect to service quality and program outcomes

b. sampling guidelines to be used in the DEA approach.

c. DEA model definition, formulation, and implementation — this model will measure service quality, effectiveness and technical efficiency.

d. sensitivity analysis for the DEA model.

e. interpretation of the results produced by the DEA model — particularly an interpretation of the significance of DEA outputs such as the weights, peer groups and performance targets associated with each DMU.

O3: To provide an implementation framework with real data — this objective involves the development of an implementation algorithm.

O4: To explore a reporting mechanism to transform the DEA results into clear and meaningful recommendations for decision-making.

1.6 Conceptual Research Model

The importance of including efficiency, effectiveness and service quality in performance measurement systems for social service organizations is well documented in the nonprofit literature (e.g. Gore, 1993; Kravchuk & Schack, 1996; Kunst & Lemmink, 2000; Martin & Kettner, 1996, 1997; Newman et al., 1987; Nyhan & Martin,
1999; NPR, 1997; Patti, 1988). Sink and Tuttle (1989) state that for the white-collar arena, the organization must concentrate its efforts on effectiveness, efficiency and quality (Sink’s definition of efficiency²).

Despite of the well documented importance of all of these three dimensions, the causal relationships between efficiency, effectiveness and customer satisfaction or service quality has not clearly been analyzed in the literature. There are several studies measuring each one of the possible relationships among those dimensions with contradictory reports (Kunst and Lemmink, 2000). For the purpose of this study, I have based my hypotheses both, on the findings reported in the literature and also on my own observations at the agency where the performance measurement system has been implemented. Table 1-2 summarizes the relationships between effectiveness, service quality and technical efficiency found in the human service field. I have depicted these relationships in the model of Figure 1.4.

Table 1-2. Hypothesized Relationships between Effectiveness, Technical Efficiency and Customer Satisfaction

<table>
<thead>
<tr>
<th>Performance Dimensions</th>
<th>Effect Cause</th>
<th>Quality of Service</th>
<th>Technical Efficiency</th>
<th>Customer Satisfaction</th>
<th>Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of Service</td>
<td></td>
<td>Positive relationship</td>
<td>Positive relationship</td>
<td>Positive relationship</td>
<td></td>
</tr>
<tr>
<td>Customer satisfaction</td>
<td></td>
<td>Positive relationship</td>
<td>Positive relationship</td>
<td>Positive relationship</td>
<td></td>
</tr>
<tr>
<td>Technical Efficiency</td>
<td>Positive relationship if including other service quality measures, such as timeliness and errors.</td>
<td>Negative relationship</td>
<td>Positive relationship</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

² Outcomes/cost
Other influential factors for evaluating social service programs have also been reported in the literature. According to Yates (1996) a wrong assumption in evaluating social service programs is that process measures are the sole cause of program outcomes. Outcomes of human services are usually more complex in their etiology. Factors exogenous to the service system may inhibit or facilitate the effect of services. Because of this influence, the results of several research studies exploring the relationships among different performance variables are still inconclusive since it appears that the direction of relationships among variables depends a lot on these factors (Ruggiero, 1998). Recognizing that these exogenous factors are important influential elements for social service organizations, the formulations developed here have attempted to include them in this research. Potential exogenous environmental factors are depicted in Figure 1-4, namely:

- Organizational factors such as size and nature of organization and program being measures, employee satisfaction and employee’s characteristics, etc.

- Environmental factors such as community involvement (degree of), community wealth, regulations, politics and the economic conditions, etc.

- Client characteristics and culture (a major influential factor in social services).
Figure 1-4. Hypothesized Model of the Relationships between Dimensions of Performance in Social Service Organizations
1.7 Research Questions

To achieve the research objectives previously introduced, the following questions will be addressed:

R1:  How can social service organizations measure and evaluate the overall performance of their services and programs?

R2:  What dimensions of performance are the most appropriate to represent the overall “health” of social service organizations?

R3:  How do these dimensions impact program/service effectiveness?

R4:  How can one represent and integrate these performance dimensions into an overall measure of performance (performance score)?

R5:  What modifications to the traditional DEA technique need to be made so that the measure defined in R4 is computed by establishing a relevant “performance” frontier?

R6:  How will this “performance frontier” identify best practices and eventually allow organizations to benchmark their services to other social service organizations of a similar kind?

1.8 Hypotheses

The following hypothesis is related to research objective O4 and research questions R4, and R5.

**Hypothesis 1:** A single index of performance that includes efficiency, quality and effectiveness can be defined and computed.

Ho: A single index of performance that includes efficiency, quality and effectiveness cannot be defined and computed.
1.9 Operational Research Model

The operational research model related to the overall performance measurement model that calculates an overall index of performance is depicted in Figure 1-5. Figure 1-6 depicts the hypothesized relationships among these variables.

**SERVICE QUALITY, Q**
Service Quality Characteristics
(Parasuraman et al. 1988):
- Tangibles: equipment and facilities
- Reliability: ability to perform the promised service reliably and accurately
- Responsiveness: willingness to help customers and provide prompt service.
- Assurance: knowledge and courtesy of employees and their ability to inspire trust and confidence
- Empathy: caring, individualized attention provided to customers

**Customer satisfaction:**
- Customer satisfaction index

**TECHNICAL EFFICIENCY, TE:**
Aggregated measure of Outputs/Inputs
- Inputs: financial and human resources, materials and information
- Outputs: number of services provided, number of clients served, number of courses taught, etc.

**EFFECTIVENESS, E:**
- Achievement of immediate outcomes measured through outcome indicators (variables).

**OVERALL PERFORMANCE SCORE**
(f(Q, TE, E))

Figure 1-5. Overall Operational Research Model
Service Quality:
- Tangibles
- Reliability
- Responsiveness
- Assurance
- Empathy

Customer Satisfaction
- Customer satisfaction index

Effectiveness:
- Achievement of outcomes measured through outcome indicators (variables)

Technical Efficiency:
- Aggregated measure of Outputs/inputs
- Inputs: financial and human resources, materials and information
- Outputs: number of services provided, number of clients served, etc.

The arrows indicate the direction of the relationship

Figure 1-6. Operational Research Model: Relationship among Variables
1.10 Contribution of this Research

This tri-dimensional performance measurement system would allow social services agencies to evaluate their overall performance taking into account not only process measures but the impact that their services are having on their beneficiaries.

Also, this research represents one of the few attempts to incorporate outcome data into a DEA model. While DEA has been used to evaluate service quality, customer satisfaction and technical efficiency of public sector organizations, few studies have combined these dimensions to compute DEA performance scores. Also, the creation of tools to elicit outcome data, and the actual construction of the DEA framework is per se, a valuable contribution to the field. Finally, an attempt to create management reports for decision-making obtained from an interpretation of the DEA results complements the performance measurement system proposed in this research and gives light to the ultimate usefulness and applicability of this research.

1.10.1 Outputs and Desired Outcomes of this Research

Outputs:

♦ The definition of the elements of a performance measurement system for social service organizations (i.e. variables, data collection mechanisms, information systems, DEA evaluation tool, reporting and information visualization systems, etc.);

♦ A DEA model that computes performance scores and recommended performance targets based on service quality, effectiveness and technical efficiency performance;

♦ A better understanding of how these three dimensions of performance are interrelated for nonprofit settings;
A series of reports and visibility boards to aid the understanding and interpretation of the DEA results

Impact of this research:

- Other private social service organizations could potentially adopt this performance measurement approach for improved decision-making.

- Nonprofits in other areas of service such as advocacy, health, environmental, etc. could also adopt this approach adapting the model to their financial realities.

- This approach could be modified in a number of ways to include other interesting mathematical approaches such as fuzzy math and integer programming to explore/improve the approach used to include particular sections of the formulations developed, such as the ones that deal with environmental variables, categorical or integer indices results, etc.

- This approach can be extended to include medical treatment outcomes. Health services organizations could then adopt this approach to calculate their performance scores including the treatment of outcomes.

- Public sector and other white-collar performance measurement systems could also adopt this approach.

1.11 Assumptions and Delimitations

The following are the basic assumptions of this research approach:

- That program outcomes could be translated into a suitable numerical format such as categorical data (such as different levels of stress relief or improvement of the quality of life), and therefore, be treated as variables of a linear programming model.

3 The outcomes observed on patients after medical treatment has been provided
• That all the technical assumptions for DEA hold for the treatment of this problem. These are described in the DEA section of Chapter 2 of this document.

• That the services provided by the agency chosen to apply the research approach is an appropriate and suitable environment to test the measurement framework and draw conclusions

Delimitations:

♦ The methodology to elicit outcomes does not prove causality. We cannot prove that the outcome is a result of the service or program, only that the outcome exists.

♦ The collection of service quality (including customer satisfaction) and outcome data went through the agency’s volunteer field offices to reach clients. Therefore, some DMUs did not have enough data points due to low response rate.

♦ While several environmental variables have been included in the model, many other influential variables affecting program performance such as organizational design, technology available, etc. will not be considered in the model.

The following section provides an overview of how this dissertation is organized.

1.12 Organization of this Document

Chapter 1 introduces the scope of the research for this study.

Chapter 2 presents the review of the literature. It starts describing the latest trends in the performance measurement field, especially those pertinent to nonprofit organizations in general and social services in particular. Next, it presents a review of the theoretical concepts behind the field of evaluation research. Finally, Chapter 2 introduces
Data Envelopment Analysis (DEA), the different modifications to the basic model to address different data and modeling issues and discusses all the studies in this field relevant to the development of a DEA formulation for nonprofit organizations.

Chapter 3 develops a conceptual model for the measurement and evaluation of performance in nonprofit organizations. Also, based on that conceptual model presents a detailed explanation of the methodology used to develop the DEA formulations suitable for this framework. There are two formulations developed. One that links all production nodes with the output measures of the previous one and another that uses the projections or proposed target level of outputs from the previous node as inputs to the following node.

Chapter 4 describes the methodology used for data collection as well as prescribes a methodology for variable selection for the DEA formulations.

Chapter 5 illustrates the implementation of the two formulations developed in Chapter 3 for the performance measurement system proposed. I will do this by applying them to two data sets, one obtained for the purpose of this research from a pilot study conducted at the agency and the second one with two types of data, real data from the agency and simulated data for variables that were not available. Input and output variables were obtained from the current and past archival data from the same agency while measures for service quality and effectiveness were simulated.

Chapter 6 presents the conclusions of this research effort and makes recommendations for future research. The first section summarizes the major findings for this study. The second section describes the major contribution emanating from this research and the third section outlines some recommendations for future research.

Finally, this document contains several Appendices. Appendix A presents the Outcome-Based development process. Appendix B presents the guidelines created for surveying agency clients and one of the survey instruments developed. Appendix C contains the actual data sets, used in Chapter 5, for the application of the models introduced in Chapter 3. Appendix D displays the programming code used to run the
models in MS Excel Solver and Visual BASIC for Excel.
CHAPTER 2 Review of Literature

This chapter covers three basic bodies of knowledge:

♦ Performance Measurement of nonprofit organizations in general and social services in particular;

♦ Evaluation Research and

♦ Data Envelopment Analysis (DEA)

My objective is to first provide an overview of what has been done to measure the performance of social service organizations, particularly the current situation regarding outcome measures. In addition, I review different approaches found in the literature for developing performance measurement systems for social and government agencies. Also, I try to address the following question: why customer satisfaction, technical efficiency and effectiveness are valid constructs to measure the performance of social service agencies? Second, I review the field of program evaluation for human services, and present its relationship to performance measurement, from the vantage point of Industrial Engineering. Last, I present a review of Data Envelopment Analysis and introduce studies that have used this technique to measure the performance dimensions pertinent to this research as well as applications of DEA in nonprofit organizations.

2.1 Background

Both human service programs and the techniques we use to evaluate their efforts and outcomes are currently undergoing a “paradigm shift” characterized by a re-conceptualization of such terms as empowerment, equity, stewardship, support, inclusion, accountability and pragmatic evaluation (Schallock, 1995). As part of this shift, outcomes have begun to be included as metrics of performance along with more traditional variables, such as service quality and productivity. However, the use of outcomes to evaluate programs is not as new as it seems, outcomes have been used as parameters for program effectiveness in the education and health services fields for more
than 15 years (e.g. Boschee & Baron, 1993; Kiresuk & Lund, 1976; Ruggiero, 2000).
What is relatively recent is the use of outcome-based evaluations to measure performance
(or efficacy/effectiveness) of social programs and other service-related applications (e.g.
Schalock, 1995; Todd and Ramanathan, 1994; Ruggiero and Duncombe, 1995; Yates,
1996) and the proposition of doing it on a regular reporting basis. Moreover, as a result
of this paradigm shift and all the reasons pointed out in the introduction, the measurement
and identification of causes of inefficacy and technical inefficiency in the provision of
social services has become of great importance for agencies and local governments
during the 1990s (e.g Light, 2000; Ruggiero and Duncombe, 1995; Schalock, 1995).

Recently, stakeholders of social service programs have demanded more
than just the provision of service. This appeal is forcing program administrators to
evaluate their programs’ effectiveness and efficiency (Light, 2000; Schalock, 1995). Yet,
there is no common agreement as to which variables should be measured, what are
acceptable benchmarks to which one should compare against, and which tools might be
used to measure them. The only common agreement seems to suggest that program
evaluation should include a variable that takes into account the after-math of the service
provided —i.e. the benefit or improvement perceived on the quality-of-life of the client
as a direct result from the service, its so called “outcome” (Light, 2000). At the center of
all these notions is the belief that the primary reason for providing the service is to benefit
the clients (Shaughnessy and Crisler et al., 1994).

An attempt to link outcome-based evaluation to some kind of technical
performance measurement model has been widely suggested during the last five years
(Nyhan & Martin, 1999; Ruggiero and Duncombe, 1995; Todd and Ramanathan, 1994).
Shaughnessy and Crisler (1994) presented conceptual approaches for assessing
effectiveness through the measurement of the outcomes of home health care, but came
short in suggesting a formal measurement model. On the other hand, the measurement of
efficiency as the relationship of outputs to inputs of the service delivery process has been
discouraged by some donor agencies. Donor agencies are stressing the fact that
productivity is merely measuring the efficiency of the agency’s process and not the
results. This is because productivity measurement as is most widely understood, is really
measuring technical efficiency or the relationship between physical inputs and outputs. Donors want a focus on outcome measurement, i.e. results. However, in the long run, the measurement of an agency’s efficiency will also have an important role in an agency’s evaluation. It is not desirable to spend more donors’ dollars to achieve an outcome that other agencies are achieving with fewer resources. While outcomes are important, the way money is spent to accomplish them must be an undeniable necessary consideration. State and Federal initiatives have begun to recognize the importance of making that connection (GASB, 1994; GPRA, Public Law 103-62; Nyhan & Martin, 1999). As a result, we see donors requiring measures of the agency’s efficiency as part of a second stage in the agency evaluation process (Triantis and Medina-Borja, 1996). In the next section, I will review what has been done in the field of performance measurement for the nonprofit sector, particularly in social service organizations.

2.2 Performance Measurement Systems for Nonprofit Organizations

The importance of including efficiency, effectiveness and customer satisfaction or a measure of service quality in the performance measurement systems of social service organizations is well documented in the literature. Several authors (Lett et al., 1999; Light, 2000; Martin & Kettner, 1996, 1997; Newman et al., 1987; Nyhan & Martin, 1999; Patti, 1988; Sola and Prior, 2001) have advocated the importance of evaluating agencies against at least two (e.g. effectiveness and customer satisfaction) if not all three of these dimensions.

Sola and Prior (2001) presented a taxonomy for the public sector that incorporates economy, efficiency, efficacy and effectiveness. Their criteria incorporate the concept of capacity building and value creation advocated by Moore (1995) that embodies service quality dimensions. According to these authors, it is important not to confuse efficacy with effectiveness. Efficacy depends on the achievement of fixed targets while effectiveness depends on the given impact promoted in social welfare. Sink and Tuttle (1989) state that for the white-collar arena, the organization must concentrate its efforts on effectiveness, efficiency and quality (Sink’s definition of efficiency).
In an argumentative article, Patti (1988) proposed a performance model for managing service effectiveness in social welfare. Patti’s definition of effectiveness is based and supported by Patti, Poertner & Rapp (1983) and accounted for three dimensions of effectiveness including outcomes. A brief description of these three dimensions of effectiveness follows:

The first type of effectiveness measures the extent to which the agency is successful in bringing about desired changes in or for the client systems it serves —i.e. outcome achievement.

The second aspect of effectiveness is service quality or the extent to which the organization is competently implementing methods and techniques that are thought necessary for achieving service objectives. Accessibility, timeliness, consistency, humanness, and technical proficiency of services, are examples of service quality.

For Patti, client satisfaction is the third dimension of effectiveness. It is concerned with how consumers assess the quality and/or impact of the services received. In addition to direct feedback, client satisfaction can sometimes be inferred from attendance rates, premature terminations, and related types of data.

This definition raises a number of issues including the extent to which it accounts for the “effectiveness” construct as it is normally defined in the literature. Particularly, whether service quality should be included with service outcomes, since it is an indicator of how well services are delivered and not the results achieved. Patti, however, recognizes these contradictions and justifies them by the fact that these three measures may or may not be related. Patti suggests that for any particular dimension of effectiveness, substantive criteria will tend to vary across different types of social organizations —i.e. what is a valid outcome for one type may not be for others. Service quality or customer satisfaction may be a measure of outcome achievement for certain types of social service organizations. This author further suggests that any management model must also address the interaction between this and other dimensions of performance such as output, efficiency, resource acquisition, and worker satisfaction.
Osigweh (1988) presented a "value-added" model for evaluating performance in training-oriented human service organizations. For this author, value-added emphasizes creating a positive impact and in that, the model is more a tool for program evaluation than for organizational performance. However, it has the merit of providing a holistic view of the program. It searches for successful outcomes, points to the need for program effectiveness, and, as a result, underscores the need for performance measurement. In short, it involves assessing the "positive difference" introduced by an existing program (Astin, 1982). Osigweh's model is a seven-stage approach to identify status, enrollee's needs, use of pre-post testing, the final measurement of specialized skills and the final tracking of added general value through a series of specialized surveys for assessing effectiveness data. The intent is to offer program managers some insights as to the success or failure of their training interventions.

Martin & Kettner (1996, 1997) propose a performance measurement system that combines three major “accountability” perspectives into one: the efficiency perspective; the quality perspective, and the effectiveness perspective. Nyhan & Martin (1999) later support this system. According to Martin & Kettner’s model, in respect to the first dimension, an accountable human service program —according to the efficiency perspective— is one that strives to maximize outputs in relation to inputs. An accountable human service program according to the quality perspective is one that strives to maximize quality outputs in relation to inputs. From the effectiveness perspective, performance measurement incorporates a focus on outcomes (i.e. the results, impacts, and accomplishments) of human service programs. Effectiveness in this sense is concerned with the ratio of outcomes in relation to inputs.

The measurement of efficiency and effectiveness in producing certain outcomes requires that financial and non-financial measures be combined for either inputs or outputs. The current most widely used nonprofit model reports receipts of contributions and expenditures for budgetary stewardship reporting only and ignores non-financial outcomes measurement (Todd and Ramanathan, 1994). These authors propose the use of a model combining efficiency and outcome measures that will provide agencies with powerful information to make their service delivery more “effective.”
this case, effective is being interpreted as producing results with the minimum possible amount of resources.

A government based large-scale initiative is the GPRA, which moves agency managers' focus away from simply measuring inputs, activities and outputs to measuring outcomes or in other words, program effectiveness through results. The provisions of the GPRA emphasize on

"improving service delivery through the provision of information about program results and service quality" and on

"improving congressional decision-making by providing more objective information on achieving statutory objectives and on the relative effectiveness and efficiency of Federal programs and spending" (GPRA, 1993: 2[a][6]).

This means that the focus on results is accompanied by new rigor in other important areas of service quality and customer satisfaction (Laurent, 1996). In 2001, the President linked GPRA to budgets, requiring agencies to submit performance-based budgets for their programs and demonstrate results in quantifiable ways (The President’s Management Agenda, 2001). In 2001 most federal agencies presented the report “Status of Achieving Key Outcomes and Addressing Major Management Challenges” pertinent to their agencies. Congress now uses this GPRA information to support its decisions on appropriation levels and reauthorization of programs. Many state and local governments are requiring similar systems of accountability to document results and justify funding (Administration on Aging, 2002).

The “service efforts and accomplishments” (SEA) reporting strategy of the Governmental Accounting Standards Board (GASB) is another government initiative promoting performance measurement. The idea behind SEA reporting is to encourage state and local governments to routinely collect and report information about efficiency, quality and effectiveness of their programs in addition to financial information that is already required. The GASB’s intention is not to prescribe specific performance measures for government programs but to rather specify the categories that state and local governments should use (Nyhan & Martin, 1999). Although currently optional, SEA's
officials expected it would become mandatory before the year 2000. A research study with 12 different government entities was done and this research work strongly suggests that the time has come for major experimentation in implementing measurement and reporting SEA indicators to elected officials and the public (Hatry et al., 2001). The project team believes *it is now important for governmental entities to experiment widely with external reporting of SEA indicators*. These experiments could start with the sets of indicators recommended for the twelve services researched. But it is important that the SEA indicators not be limited to those.

In fact, as several authors are beginning to note, the integrative nature of such a performance measurement system could make it become the de facto standard or benchmark against which all other performance measurement systems are judged (Epstein, 1992; Nyhan & Martin, 1999).

The key issues examined by GASB research study included:

♦ What types of performance measures should be considered? What performance indicators should be candidates for reporting? To what extent are these indicators measurable, valid, and comprehensive?

♦ How, and to what extent, should individual performance measures be disaggregated?

♦ What comparisons should be reported for the various performance measures?

♦ What explanatory data should be included with performance measures, and how should these data be presented?

♦ To what extent are these performance measures verifiable?

♦ How should performance information be communicated and displayed? In what types of documents should performance indicators be reported?
What are the likely added costs and the feasibility of obtaining and reporting performance measures?

What are the uses for, and who are the users of, performance information?

Likewise, the Roberts Enterprise Development Fund (2001) has created the SROI framework (social return on investment) to measure the social value created by nonprofit organizations. This foundation has researched and focused on six metrics that attempt to capture the impact of the social enterprise on individual social outcomes that are difficult to ascribe a monetary value. This is an attempt to understand the connection between the funds provided to nonprofit organizations and the results those organizations achieve.

I have followed the theory supported by most of the authors presented in this section that supports efficiency, service quality and effectiveness as the right drivers of nonprofit performance.

2.2.1.1. The effects of the Environment in organizational performance

Several authors have emphasized the impact that the operating environment have on public and nonprofit production above other types of private production systems (Blank and Valdmanis, 2001; Fried et al., 1999; Ruggiero, 1996).

The recognition that production is affected greatly by the environment is not new. Production in the public sector was modeled by Bradford, Malt and Oates (1969) as a two stage process in which in the first stage inputs are used to produce intermediate outputs and the second stage final outcomes are determined by the level of these intermediate outputs and by non-discretionary environmental variables. Empirical studies of public sector production support this theory —i.e. environmental variables not under the control of the management of the firm have a substantial impact on the outcomes that are provided (Ruggiero, 1996). In fact, one can assume that environmental exogenous factors affect both, intermediate outputs and outcomes in social service production.
An example of this double effect on outputs and outcomes can be given by training courses targeting disadvantaged communities. Service delivery is impacted by demographic and geographic characteristics of the site where the course will be taught. Certain demographic groups present different levels of interest to certain types of training. The distance between the training center and the homes of the target population is also a factor that affects enrollment. One can expect that the number of enrollees will be typically lower in highly dispersed areas, for example, with families at the lower brackets of income level, where public transportation is an issue. Furthermore, the level of education, the age of the participants, and other environmental factors have been proved to affect outcomes, or the level of observed change on the community.

In presenting a performance evaluation framework for social services, one should include, therefore, the effect of the environment on the organization.

2.2.2 Performance Measurement Dimensions

As it was argued in the introduction of this study, the exact definitions of effectiveness and efficiency have been matter of extensive debate. This section further examines this debate.

2.2.2.1. Effectiveness

Effectiveness measures the extent to which the program obtains its goals and objectives fulfilling its mission statement (Epstein, 1992; Hatry et al., 1992; Kirchhoff, 1997; Schalock, 1995). A program is effective to the extent that it is accomplishing what it was designed to accomplish. Effectiveness is defined in light of the goals and objectives of the organization (Cooper, Seiford and Tone, 2000; Chalos and Cherian, 1995). It has an external focus incorporating judgments of relevant stakeholders (Epstein, 1992) and is measured according to the level of social welfare or social capital it generates (Sola and Prior, 2001).

In the program evaluation field there are several measures of effectiveness. Several articles have been published assessing the effectiveness of nonprofit units. A careful review of these studies, however, reveals two different
interpretations of the effectiveness dimension of social, educational and health service units. A brief discussion of these two approaches follows.

One would think that the sole measurement of outcome indicators would represent the effectiveness dimension of performance. For several authors, however, outcome measurement per se is not enough. Effectiveness is being interpreted as producing results with the minimum possible amount of resources. A definition similar to the one provided by Sink and Tuttle (1989):

\[ \text{Effectiveness: outcomes/inputs}. \]

The second approach to effectiveness is the one adopted by the United Way of America (1996) and others, who measure the degree of change in clients’ attitudes, knowledge, behavior, quality of life, etc. For these authors the sole demonstration of program outcomes is a measure of how effective the program was. They leave to other dimensions of performance to account for the amount of resources consumed to obtain those outcomes.

Magura and Moses (1986) define program effectiveness as the ratio of units of client outcome to some standard unit; this standard may represent the maximum improvement or deterioration possible in any given problem area or may represent some minimal level of adequacy as legally or culturally defined. A similar approach is followed by Blank and Lovell (2000) who state that effectiveness refers to the concordance of the service mix with the stated objectives of the agency.

According to Kaplan (2001) several authors have reported difficulties in defining exact metrics for organizational effectiveness (Goodman and Pennings, 1977; Cameron and Whetten, 1983). These last authors state that given that there is no one universal model for organizational effectiveness, it is more worthwhile to develop frameworks for assessing this metric than to attempt to develop theories of effectiveness. In general, a framework to assess social organization's effectiveness will include the identification of the most relevant outcomes pursued and proxy indicators to measure them (Cameron, 1986; Cameron and Whetten, 1983).
As part of this research, I investigate and propose a definition as well as an index of effectiveness suitable for human service programs that is used in the DEA formulation of Chapter 3. The index of effectiveness proposed in this research may come from a percentage of clients who acknowledge the presence of an outcome or it may also come from the mean response to a Likert scale depending on the type of survey used as well as the needs of the organization being assessed.

2.2.2.2. Productivity and Technical Efficiency

Traditionally, technical efficiency used to be considered a measure of labor productivity (Farrell, 1957). It was expressed, for example, as a ratio of units produced divided by hours worked. However, in many cases this measure is inadequate because it is more common that organizational units have multiple incommensurate inputs and outputs (i.e. they are not in the same proportionate scale).

According to Kirchhoff (1997), the notion of productivity remains problematic, with some making it equivalent to efficiency (Bouckaert, 1993) and others making it synonymous with performance (Epstein, 1992). For several authors (Hatry & Fisk, 1992; Kirchhoff, 1997; Magura and Moses, 1986) productivity is a combination of efficiency (outputs to inputs) and effectiveness (accomplishments relative to goals and purposes). According to these authors, performance measurement then, consists on operationalizing performance by identifying, selecting, measuring and analyzing indicators of productivity construed as efficiency and effectiveness. Outcomes and technical efficiency ratios can be related by defining “efficiency” as the ratio units of client outcome either to units of program output or to units of resource input. Program output would be used in the definition when interest is focused on the relationship between work accomplished and client outcomes. Resource input would be used when interest is focused on the relationship between the amount of money spent for services and client outcomes (Hatry, 1978).

Blank and Lovell (2000) state that efficiency in the public sector refers to the degree to which service provision is maximized, given the resources at hand. Under
different circumstances, according to these authors, efficiency could also be measured from the opposite orientation, as the degree to which resource consumption is minimized in order to satisfy service demand.

In the economics and operations research literature, however, technical efficiency is a measure of outputs (aggregated outputs) over inputs (aggregated inputs), which in a way can be compared to the traditional productivity definition. I will make use of this approach when referring to TE or technical efficiency for the DEA formulation.

Production Function

Farrell (1957) introduced his relative measure of technical efficiency. Conceptually, Farrell’s approach involves comparing the performance of production units against each other, using an empirical enveloping hull constructed from observed input-output values. This hull is called the efficient production function. The concept of production may be somewhat challenging to understand in general terms but it becomes simpler when explained in terms of real systems. I will explain this concept further in Chapter 2 of this document.

Unfortunately, Farrell (1957) showed that in most cases the efficient production function could not be determined exactly because in essence, production systems involved complex processes and the more complex the process the less accurate the production function will be. Nonetheless, he ascertained that the production function could be approximated by an empirical construct representing the “best observed results in practice” (Girod, 1996).

2.2.2.3. Service Quality and Customer Satisfaction

As Deming stated (Deming, 1994; Walton, 1986), quality comes first and customers must be the focus of everything an organization does. The quality revolution, evident in both industry and human services, stresses that quality is essential to both the processes used in service delivery and the outputs and outcomes obtained from those services. It is agreed that quality processes and quality outcomes should be guiding principles around which service’s outcomes are selected and interpreted (Schalock, 1995).
Unlike goods, however, which can be measured objectively by such indicators as durability and number of defects (Crosby, 1979; Garvin, 1983), service quality is an abstract construct because of three features unique to services: intangibility, heterogeneity and inseparability of production and consumption (Gronroos, 1988; Parasuraman et al., 1985). In the absence of objective measures, an appropriate approach for assessing the quality of a firm’s service is to measure consumer’s perceptions of quality. Perceived quality is the consumer’s judgement about an entity’s overall excellence or superiority (Zeithaml, 1987).

Service quality performance measures generally take one of two forms: (a) consumer (client) satisfaction as perceived by the clients and (b) outputs or process measures related to some quality standard (e.g. timeliness, cleanliness, error free rates, etc.). A service is perceived of as having high quality or not, to the extent that the service organization provides its client with a product that meets his/her needs in the most appropriate way. (Triantis and Medina-Borja, 1996). In some organizations, however, the constructs used to measure customer satisfaction include items in process related quality measures such as timeliness, and kindness of the service deliverer. Human services agencies are now more than ever concentrating in client-centered service delivery (Rapp & Poertner, 1992). This is the case of the Agency customer satisfaction survey. For my research I will use both approaches to service quality: customer satisfaction including process related quality measures.

Models to Measure Service Quality

Service quality is a perceived judgement resulting from an evaluation process where customers compare their expectations with the service they have received (Gronroos, 1984). Due to this subjective nature, measuring service quality poses difficulties for the service provider because of the unique characteristics of service: intangibility, heterogeneity, inseparability and perishability (Lamb & Woo, 1997). It is indeed and elusive and indistinct construct  (Parasuraman, Zeithaml & Berry, 1985).

Service quality research has already gone through two phases of development from building up a conceptual foundation for understanding service quality
in the early 1980s to designing models to measure service quality in the late 1990s (Gronroos, 1993). Lam and Woo (1997) point out that the third phase has involved from refining a static model to a dynamic model of service quality.

However, as Tankersley (2000) points out, the usefulness of this information, once gathered, through surveys or other means, remains problematic. The theoretical guidance available is still very limited, mainly because there are a number of significant questions still unanswered regarding service quality. For example, it is yet to be determined how the different dimensions interrelate, whether or not high performance in one dimension compensates for low performance on another. As state, the question as to the level of quality (high, moderate, low) that a service firm should offer is still undetermined by the literature (Parasuraman and Zeithaml., 1990; Parasuraman et al., 1994).

Customer surveys have been used widely to measure customers’ opinions and/or perceptions of service quality (e.g. Brown, 1996; Carman, 1990; Hatry et al., 1998). One very well known way to measure service quality is by using the SERVQUAL scale developed by Parasuraman, Zeithaml & Berry (1988 & 1991a). This scale has been used by both academics and practicing managers in a variety of settings to measure the quality of service (e.g. Babakus & Boller, 1992; Carman, 1990; Lam, 1997; Lam & Woo, 1997). SERVQUAL is based on the concept developed by Gronroos in 1982 of perceptions minus expectations (P-E), referred to as “gap theory”. (Hill & McCrory, 1997). The SERVQUAL scale aims to compare customers’ expectations and their perceptions of actual performance and is used to measure a specific long-term attitude at a single point in time. The operationalization of the service quality construct represents the difference between two 7-point SA/SD (strongly agree to strongly disagree) Likert scales — one to measure customers’ expectations about companies in general within the service sector being investigated and the other to measure customers’ perceptions about the particular company whose service quality is to be assessed (Lamb & Woo, 1997). If a consumer’s expectations are met, service quality is perceived to be satisfactory; if they are not met, it is perceived to be less than satisfactory; and if they are exceeded, it is perceived to be more than satisfactory, delighting the customer (Hill & McCrory, 1997).
In SERVQUAL there are 22-pair statements to measure expectations and perceptions across five generic service quality dimensions:

**Tangibles**: Appearance of physical facilities, equipment, personnel and communication materials.

**Reliability**: ability to perform the promised service dependably and accurately.

**Responsiveness**: Willingness to help customers and provide prompt service.

**Assurance**: knowledge and courtesy of employees and their ability to convey trust and confidence, and

**Empathy**: Caring, individualized attention provided to customers (Lamb & Woo, 1997; Orwig, Pearson & Cochran, 1997).

According to Parasuraman, Zeithaml & Berry (1988, 1991a and b) the difference between the rating of what the customer expects and what he actually experienced represent the measure of perceived service quality.

The literature reports both, advantages and problems with the instrument (Lamb & Woo, 1997; Orwig, Pearson & Cochram, 1997) particularly with regard to its conceptual foundation and methodological limitations. Particularly Carman (1990) and subsequently others (e.g. Babkus & Boller, 1992; Brown, Churchill & Peter, 1993; Cronin & Taylor, 1992) criticized the concept of service quality in SERVQUAL from a theoretical perspective as confounding the concept of customer satisfaction with service quality. This is based on the SERVQUAL scale being based on the same disconfirmation model, which is widely adopted in the customer satisfaction literature. Another criticism arises from the fact that the operationalization of service quality as performance minus expectation (P-E) make it difficult to reconcile with general attitudinal models even though service quality is conceptualized as an attitude by the SERVQUAL developers. Also, Carman in particular (1990), discussed the value of measuring ex-post expectations. This is consistent with some empirical findings (Clow & Vorhies, 1993) that customers
who have a negative experience with the service tend to overstate their expectations, creating a large gap, while customers who have a positive experience tend to do the opposite (Lam & Woo, 1997). Other criticisms to the SERVQUAL scale have been presented, some of which have been compiled by Smith (1995) and Buttle (1996).

As a consequence of SERVQUAL’s limitations, alternative models have been proposed in the literature (e.g. Carman, 1990; Hill & McCrory, 1997; Johnson, Tsiros and Lancioni, 1995; Rao & Mayuresh, 1997; McDougall & Levesque, 1994; Mels et al., 1997; Rust & Oliver, 1994). All these approaches utilize surveys to elicit customers’ perceptions and/or expectations of service quality. Hill & McCrory (1997), for example, conceptualized service quality as perceptions minus importance (P-I) for a study done in a maternity hospital. Cronin and Taylor (1992) disagree with the P-E scale and suggest the superiority of their performance-only SERVPERF scale in terms of construct validity and operational efficacy.

Johnson, Tsiros and Lancioni (1995) for example adopt a "systems approach" to modeling service quality, conceptualizing the construct as a function of inputs, process and output. McDougall and Levesque (1994) presented a similar approach but they measure outcomes of the service delivery instead of outputs.

A second approach to measure service quality is that of measure perceptions of performance (Hurley & Estelami, 1998). One of the developers of SERVQUAL (Parasuraman, 1995) has come to the conclusion that a perceptions-only approach to measuring quality is even more acceptable from a predictive validity point of view. De Ruyter et al. (1997) also found that the perception of the service performance is the most important indicator of service quality than expectations and other measures. Some other researchers have opted to develop industry specific conceptualizations of service quality (e.g. Dabholkar et al., 1996). In fact, Asubonteg et al. (1996) suggested that due to the great variability of potential elements of service quality from one service and industry to another, it does make sense to develop industry-specific constructs.

I will use a scale for service quality performance that includes timeliness, empathy and other constructs comparable to those on SERVQUAL and one item in
customer satisfaction. All items will measure perceptions and not expectations. While I recognize both constructs —service quality and customer satisfaction— as different, I do believe they are complementary and that service quality causes customer satisfaction. I will combine them on a single index, though, as a measurement artifact to avoid multi-collinearity problems (with both variables being highly correlated with each other.) If an index is created then each variable will contribute to it with weights set up by the decision maker, and we’ll have only one variable instead of two explaining the same performance dimension.

2.2.3 Relationships among performance dimensions

The literature that investigates the relationships among different dimensions of performance lies in different organizational realms. The next sections review studies pertinent to this research.

2.2.3.1. Studies in the Performance Measurement Field

Garvin (1988) provides an analysis of how quality is related to other performance measures. According to this author, the argument for a positive correlation between quality and productivity is usually oversimplified. For manufacturing for example, less rework means more time devoted to manufacturing acceptable products (more output) and less scrap means fewer wasted materials (less input). Garvin presents an argument saying that this view takes into account partial definitions of productivity (such as outputs over labor inputs) and does not take into consideration the fact that in reality, a good measure of productivity has aggregated measures of inputs and outputs (such as the total factor productivity or Farrell’s measure of technical efficiency). According to Garvin, it appears that in the short run both variables move in opposite directions, since as quality programs get underway there are inevitable disruptions. Nevertheless, once the programs have been institutionalized, productivity should increase. Garvin presents several research studies that corroborate a positive correlation between quality and productivity but acknowledges that more systematic studies are needed in this area. Several authors, (Kunst and Lemmink, 2000; Martin & Kettner, 1996, 1997; Stamatis, 1996) corroborate these finding stating that according to the TQM theory,
productivity and efficiency increase when programs provide high-quality services and decrease when low-quality services are provided. Stamatis (1996) goes further stating that increased service quality results in increased customer satisfaction. Brown (1996) has also addressed the issue about customer satisfaction and service quality. According to this author, service or process quality drives customer satisfaction.

Kunst and Lemmink (2000) present results of a European study targeted at identifying and evaluating success parameters of high (quality) performance and their possible interrelationships. To do so they adapted the SERVQUAL instrument to hospitals. They concluded that environmental variables play an important role in how performance dimensions are correlated and that perceived service quality by customers are positively, but only to a limited degree, linked to business performance (efficiency). They also concluded that progress in TQM implementation leads to higher business performance of hospitals (efficiency/cost effectiveness) and to a higher perceived service quality by patients that in turn, increases effectiveness. Figure 2-1 below describes Kunst and Lemmink results.

![Figure 2-1. Performance Relationships in Hospitals (from Kunst and Lemmink, 2000).](image-url)
Sink & Tuttle (1989) present a model that depicts the interrelationships between organizational performance criteria. According to this model, there is an overlap among the different criteria. For Sink, effectiveness incorporates quality attributes; productivity includes quality, effectiveness and his definition of efficiency, and quality is quite pervasive. For these authors, productivity will follow if the organization manages to achieve a balance between effectiveness, quality and efficiency. I have adapted Sink and Tuttle’s model in Figure 2-2 to show only the criteria of interest for my research.

![Figure 2-2. Interrelationship between Organizational Performance Criteria (adapted from Sink & Tuttle, 1989)](image)
Kaplan and Norton (1992; 1996) presented the concept of the balanced scorecard that allows organizations to look their operations from four important perspectives:

- The customer perspective (how do customers see us?)
- The internal perspective (what must we excel at?)
- The innovation and learning perspective (can we continue to improve and create value?)
- The financial perspective (how do we look to shareholders?)

They advocate that organizations should include measures in these four perspectives and try to achieve a balance to avoid sub-optimization. Each company develops its own scorecard with its few vital measures analyzing their interrelationships and causation according to their particular view of the world. For the purpose of my research, I will analyze measures that fall into the internal perspective (technical efficiency); customer perspective (customer satisfaction and service quality). Effectiveness, as the accomplishment of objectives can also be viewed as an internal business perspective measure, but it also can be part of the customer perspective, since it is the customer the one that presents the changes or outcomes.

Rouse et al. (1997) presented a framework for performance measurement of highways in New Zealand based in the Balanced Scorecard concept. This framework makes use of measures of inputs, outputs and outcomes. The framework encompasses a performance pyramid based on the Cross and Lynch approach and informed by the balanced scorecard approach of Kaplan and Norton (1992). This framework embodies multiple perspectives of the organization with a structure of measures linking critical success factors to process drivers. The schematic model is depicted in Figure 2-3. Distinctions between measures of outcome, output, and input enable finer partitioned of analyses into managerial notions of efficiency, effectiveness and economy. Data
Envelopment Analysis was used to perform the analysis. The impact of environmental factors on efficiency is stressed. The Rouse et al. study concludes that efficiency (technical efficiency) and effectiveness are not correlated.

Rodgers and Hunter (1991) performed a meta-analysis of studies of management by objectives successful implementations as a measure of organizational effectiveness. This study indicated that 68 out of 70 studies showed productivity gains. There is evidence that when an organization accomplishes its planned objectives, therefore increasing effectiveness, then productivity increases.

Figure 2-3. The integration of the Balanced Scorecard with the Performance Pyramid (adapted from Rouse et al., 1997).
Sudit (1996) states that efficiency and quality drive effectiveness. Organizations normally strive for higher levels of effectiveness and rarely do efficiency goals serve as end objectives. This author also states that quality is different since certain quality characteristics (zero defects, customer satisfaction) may also logically feature in the set of effectiveness goals.

Due to the unique characteristics of the nonprofit sector, I include now research in the fields of government and nonprofit performance measurement and program evaluation. The next section presents findings regarding the relationships of these three performance variables found in these areas.

2.2.3.2. Studies in the Nonprofit and Program Evaluation Fields

This section begins by defining nonprofit organizations in order to clarify performance relationships pertinent to this sector. The National Taxonomy of Exempt Entities (NTEE) offers a definitive classification system for nonprofit organizations recognized as tax exempt under the Internal Revenue Code:

These charitable organizations are exempt from federal taxes because of their religious, educational, scientific, and public purposes. The broad range of their activities includes health, human services, arts and culture, education, research, and advocacy. The use of NTEE is key to illuminating the diversity of the nonprofit sector. This review of literature addresses in most part performance issues in these type of organizations.


This study focuses in the measurement of performance in two types of nonprofit organizations dedicated to the delivery of social services (Major groups V and VII of human services, public and societal benefit). These organizations might collect charitable donations from the general public, but not exclusively. Most also get grants and contracts from other agencies, particularly from the government to perform their services or put in place a program.
Several research studies related to social services have explored the relationship between productivity or technical efficiency, effectiveness and service quality. The results are still inconclusive. The following is a summary of some of the most influential results in the field.

The establishment of quality management systems that control and assess customer satisfaction and other dimensions of service quality has been widely explored in literature. Still, a timely service, and/or a satisfied customer are only telling us that the service was provided in the best possible way. We do not know if the service is attaining its goals and if it has any impact in the life of the client.

Sola and Prior (2001) state that for the public sector, production is a process of "value generation" just as in any other process. Products or services delivered can be described as promoters of certain desired consequences or changes on customers. They, therefore, claim to have initial evidence that the production of goods and services, and their corresponding level of quality, shape the final outcome and, therefore, the degree of effectiveness of a given production process. These authors analyze public health care organizations and say that in certain cases the increase on productivity or efficiency might be promoted, although the rate of success (effectiveness) might be reduced. On the other hand, quality objectives can be considered in some cases as final objectives because in the health care industry, total quality management (zero defects) is a desirable objective that defines the effectiveness of these services. According to Sola and Prior, the relationship between efficiency, efficacy, quality and effectiveness can follow two patterns. The first follows a traditional approach, which represents a negative trade-off between service quality and efficiency. In health care this pattern might signify to improve productivity and reduce cost by lowering the quality in the provision of services, such as reducing the time given to each patient, reduction of technical requirements for carrying out operations, etc. The second approach is the "total quality management" approach. It focuses on quality as the final target of the organization, and therefore, the relationship between quality and total cost changes. Under this approach,
costs and quality exhibit complementarily, so that maximizing quality helps minimizing costs. The authors explain this concept by saying:

"...the [effectiveness] objective for hospital A is to reach the maximal quality. In doing this, the hospital is optimizing quality and efficiency at the same time, and, most importantly, optimizes the effectiveness of the hospitals." pp. 222

Martin & Kettner (1996 & 1997) provided a list of relationships between quality and productivity for the social service sector based on TQM interventions: higher quality of service should result in lower error rates (higher output) less paperwork (less inputs), less processing time (less inputs), happier funding sources (more outcomes), more satisfied clients, lower costs (less inputs) and a better public image (higher outcomes). This statement implies that higher service quality drives higher productivity or technical efficiency and higher effectiveness. However, it is also true that higher customer satisfaction might imply a greater dedication of time to each client, therefore reducing productivity.

Measures of productivity and quality in the public sector have their unique approach. Rosen (1993) advocates that quality of service needs to be considered in the productivity picture. This author says that for “public sector practitioners” how much they do is far less important than how well they do it. Just how to “consider” quality is a question still unsettled. This author presents four alternatives to consider quality:

♦ Disregard quality

♦ Present quality information separately from efficiency measures. Presumably decision-makers internalized both sets of information and somehow integrate both factors into their actions.

♦ Use of quality measures as a screening device. One or more quality criteria may be established for a given kind of output. When the output is measured, any unit of work that does not meet the standards can simply be excluded from the output count. This is the strategy adopted by Hatry (1980).
Use quality measures to discount output. This is the fourth approach adopted by Rosen (1993) that does not eliminate output but rather discounts those units that fall below the desired standard. This is slightly different than the previous alternative in that it is possible to eliminate fractions of the product according to the rule set.

The approaches advocated by Rosen and Hatry assume that somewhat, quality and productivity are positively correlated.

Health and social scientists do not normally include technical efficiency measures in process evaluation. They do include, however, measures of process quality and recognize that quality influence the outcome of the service. In particular, McGraw et al. (1996) presents a study that illustrates the complementary evaluation strategies of process and outcome evaluation. These authors examine the relationship between teacher characteristics, measures of curriculum implementation and competing influences (all process evaluation elements) and student outcomes. In this model, process quality measures such as teacher characteristics and staff training were hypothesized to have direct effects on the degree of implementation of the classroom curricula, which, in turn, would have a direct effect on student outcomes. The observation lasted three years. The results of the analyses did not support the model. Teacher's characteristics did not affect implementation. The McGraw et al. study has a number of implications for the field of program evaluation. In the context of my proposed research, the most important is that these results support the importance of process measures (conceptual model, quality of service, implementation factors) and its relationship with program outcomes (all discussed in the present Chapter 2).

Plante et al. (1995) present a study that measures treatment outcome and client satisfaction among families and children that underwent mental health treatment.

Several studies have found an inverse relationship between worker to client ratios (which can be treated as an inverse measure of efficiency if the client is considered a measure of output) and various indices of service effectiveness (Holland, 1973; Linn, 1970; Moos, 1974; Martin & Segal, 1977). The Moos and Martin and Segal
studies both found that clients in facilities with lower staff-client ratios were more likely to present better outcomes than those in agencies with higher caseloads. However, most of the work on the relationships between productivity as reflected in staff-client ratios and service outcomes appears to have been done in inpatient facilities. Patti (1987) suggests that there is some evidence that similar trade-offs may occur in outpatient settings. For example, there is research that indicates a positive relationship between caseload size and error in the determination of eligibility for public assistance (a service quality measure) (Baker & Vosburgh, 1977). In another study of several state welfare departments, Newman & Ryder (1978) found that agencies with lower error rates had higher administrative costs. They also concluded that some aspects of service quality, such as providing assistance to clients in the completion of applications, were negatively associated with output. Whetten (1978) provides further evidence that staff members’ perceptions of the effectiveness of the service provided in manpower training agencies were inversely associated with agency output. Patti, however, asserts that this type of research findings does not necessarily suggest that agencies must necessarily be inefficient in order to be effective.

As a result, one can conclude that a definite direction and sign of the relationship between efficiency and effectiveness has not yet been confirmed. This will be taken into account for modeling purposes in Chapter 3.

2.2.4 The Concept of Capacity Creation

Organizational capabilities have been recognized as the drivers of service quality and performance in several areas, such as banking (Athassanopoulos, 1999), public sector (Kirchhoff, 1997) and social services (Eisinger, 2002). Roth and Van der Velde (1989, 1991) stated that technology, human resource development, business process integration and factor productivity are capabilities of firms. Their degree of importance within an organization is affected by the nature of the service environment. This is also true for nonprofit organizations. According to Eisinger (2002), the ability to acquire resources is a measure of organizational capacity. However, this author qualifies as organizational effectiveness the ability of organizations to acquire enough resources
without over spending to perform their jobs. The present research study will call that ability as organizational efficiency instead, since it is a measure of resources used to create capabilities and not a measure of the level of outcome achievement only (although Eisinger’s operationalization does include perceived goal attainment).

In the provision of social services, organizational units collect or receive resources (inputs) from the community, donors or government, to generate value by providing services to clients. They do this through the acquisition of facilities, equipment, staffing and training, and the preparation of programs and services, the so-called "capabilities." Several authors have advocated the importance of including capacity creation in the evaluation of nonprofit organizations (Eisinger, 2002; Letts et al., 1999; Kaplan, 2001; Kirchhoff, 1997; Sola and Prior, 2001). In particular, Letts et al. (1999) claim that capacity is the ability of nonprofits to perform and this is the foundation for "lasting social benefits." They say that the missing ingredient in the prevalent, program-centered conception of social impact is organizational capacity.

Sola and Prior (2001) present the following definition of capacity. Capacity implies "the ability to organize centres of responsibility in order to orientate their management by following the achievement of global organizational targets." P. 219

In simple terms, as Kirchhoff (1997) explains: a nonprofit organization selects the site; hires counseling staff; develops and implements management and control procedures for ensuring contract performance; compensates, manages, trains, and develops staff members; markets services to eligible populations; and ensures the availability of counselors for service delivery during business hours. Once the agency is "open for business" capacity is available. However, the capacity is not necessarily being delivered. Service deliver occurs when eligible clients seek and receive services from the capacity being made available. Capacity building yields fixed, measurable, direct service capacities for the given level of inputs.

The process of value generation (Moore, 1995; Sola and Prior, 2001), requires, according to these authors, the joint accomplishment of economy, capacity, efficiency and effectiveness according to the graphic shown in Figure 2-4. Unfortunately,
as Letts et al. (1999) put it, building organizational capacity in the nonprofit sector is not rewarded or encouraged by the market. These authors establish a research study that included roundtable discussion with nonprofit managers to address the following two questions: "How do nonprofit organizations perform?" and "What enables them to perform?" The results obtained outlined organizational capacity as an essential resource to perform, and therefore, achieve social impact.

![Diagram of the Process of Value Generation](image)

According to Letts et al. (1999), effective nonprofit organizations rely on three types of organizational capacity: capacities for program delivery, which grows out from a specialized field of practice (health, education, environmental etc.) Staff is recruited according to programmatic skills and experience. This type of capacity provides the organization the ability to accomplish outcomes and where effectiveness is taken as a crucial dimension of performance. The second one is the capacity for program expansion,
or the capacity to grow, which requires more resources and capabilities to respond to issues it once handled more informally. Performance is a matter of organization, management and program expertise. The third type of capacity is what is called "adaptive capacity" or the capacity of delivering on its mission, knowing that the program is still relevant given the changing needs of its beneficiaries. Performance here encompasses organizational effectiveness, efficiency and mission impact.

One study (Eisinger, 2002) found that capacity attributes do not appear to contribute to effectiveness among street-level food programs. Eisinger concludes that effectiveness (efficiency) is a function of individual effort and skills rather than rules, routines, support networks, or planning. Others would argue that individual effort and staff skills are of course, part of the construct “capacity”.

The concept of capacity creation encompasses the acquisition of inputs to efficiently produce outputs, which matches the basic production theory central to DEA. The concept of capacity creation will be taken into consideration in the model formulation in Chapter 3.

2.2.5 Emergency Management in Social services

Services provided to people that have been affected by an emergency have unique characteristics. Consequently, how clients react during and after an emergency and their expectations of service greatly affect the way they perceive the service.

Researchers in the social sciences as well as in the management fields are beginning to gain a better understanding of disasters or emergencies and their effects on people’s lives, properties, society and the environment. Researchers and practitioners are developing and assessing methods of responding to emergencies using technology and other managerial and logistic approaches to develop new ways for “managing” emergencies (Tufekci & Wallace, 1998). According to these authors, the field of emergency management encompasses two stages: i) intervening to reduce the impact of emergencies before the event becomes catastrophic and ii) intervening to deal with the impact of the event after it occurred. The services to be studied at the Agency fall into the second category.
In fact, it has been suggested that the nature of the emergency response at the agency and the emergency reaction of the victim follow the laws of chaos theory, particularly fractals and path dependent processes (Koehler, 1995). The structural forms that a chaotic process leaves in its wake are fractals (like the different branches and parts of a cauliflower, all identical in different sizes). In summary, chaos is the process and fractals are the resulting structure (Peitgen, Jurgens & Saupe, 1993). At the same time, the way the service is delivered follows very defined paths prone to errors and mistakes. These conditions require substantial changes in the way emergency care is delivered.

In the case of emergency communications for the military, fractal structures and processes affect the agency at the micro level. How does the emergence of dynamic microstructures —i.e. military, family, chapter, calling center— effect the macro pattern, and what can be done to improve this dynamic so that the most efficient emergency response macrostructure emerges? The errors or delays in the path-dependent processes perform until a message is delivered and some action is taken increase the error rate driving the system even further into chaos (Koehler, 1995).

Further, the unique nature of emergency services can make that the way the client or beneficiary of the service, who is at the same time the victim of the emergency, perceive the way the service was delivered and the outcome of it depending of the path-dependant error rate. Several authors (Furlong & Scheberle, 1998; Schneider, 1992) have recognized that there is a gap between the way emergency clients perceive the availability, usability and effectiveness of the agency providing support and the way the agency itself perceive their operations and ability to respond. This is basically motivated by client’s expectations of the role of the agency and the help that they are entitled to receive. When this help is not immediately forthcoming, clients may think that the agency has been unsuccessful delivering its service.

Research on customer satisfaction and customer’s perceptions of effectiveness in these types of cases must consider this gap. Furlong & Scheberle (1998) for example, included in their questionnaire a question about the expectations of clients of the service and correlated them with the satisfaction and effectiveness index. The path-
dependent processes and error rate must also be considered. Furthermore, it is clear that not all individuals respond to traumatic events with the same pattern of adjustment (Freedy et al., 1992). This latter fact suggests that individual differences with regard to mediating variables (e.g. social support, coping behavior, etc.) may be very important in determining the reasons for unrealistic expectations of service. Post-emergency factors, such as ongoing experiences, have also been noted as influential of clients’ perceptions of the emergency service.

These notions are considered in the course of this research by adding a question about prior expectations of the client in all emergency-related questionnaires so that the effect of expectations is considered in the analysis.

### 2.3 Outcome Measurement Approaches

The measurement of efficiency and effectiveness requires that financial and non-financial measures be combined for inputs and/or outputs. The current most widely used nonprofit model reports receipts of contributions and expenditures for budgetary stewardship reporting only and ignores non-financial outcomes measurement (Todd and Ramanathan, 1994). These authors report that measurement of outcomes becomes all the more important in order to achieve effective use of scarce funds, particularly in those cases in which social organizations have been submitted to production functions that do not take into account outcomes. In those cases, it has been shown that their resource levels have dropped below those required to fulfill the organization’s mandate as a consequence of the evaluation. In such cases, where funds are scarce, the use of a model combining technical efficiency and outcome measures will provide agencies with powerful information to make their service delivery more effective. Effective being interpreted as producing results with the minimum possible amount of resources.

Todd and Ramanathan presented an accounting-based empirical model to examine the responsiveness of budgetary allocations to public demand for services and the resulting outcome-generating activities in the New York City Police Department (NYPD). The analysis combines non-financial measures of the results of operations with
budgetary and financial measures. The model explicitly incorporates efficiency and effectiveness measures and reflects as well the interactive nature of outcome-generating activities in the NYPD. Todd and Ramanathan (1994) state that the impact of a program resulting from changes in budgetary responsiveness to changes in social needs may be evaluated by examining successively:

1. the responsiveness of budgetary provisions to prior changes in demand for services;
2. the changes in outputs generated from changes in budgetary provisions; and
3. the changes in outcomes produced as a function of outputs.

In fact, as Todd and Ramanathan (1994) stated, the nature of the outcome-generating activities of an organization is illustrated in the responsiveness of outputs and outcomes to increased inputs.

Moore (1995) presented the outcomes-to-costs efficiency matrix as a tool that allows for simple comparison of program outcomes to program costs. Moore classifies programs with high outcome and low cost as “bright spots.” Programs that are high in both costs and outcomes are “rising suns.” “Space” programs cost little and, in turn, produce few real outcomes. Finally, programs that produce low outcome yet involve high cost are “black holes”. Moore presents an interesting practitioner point of view of what to do to produce more program outcomes and be more efficient. However, Moore does not present a way in which those outcomes will be measured and evaluated.

Schalock (1995) presented an outcome-based evaluation model that at the heart combines three types of program analyses: effectiveness, impact and benefit cost analyses. Each of them provide formative feedback, information that can be used to determine the following:

- The extent to which the program obtained its goals and objectives (effectiveness analysis).
• Whether the program made a difference compared to either no program or an alternate program (impact analysis)

• Whether the program’s benefits outweigh its costs (cost-benefit analysis).

The most common approach to outcome measurement has been to collect outcome-related data from the clients and statistically analyze the results. Some agencies have begun to implement variations of this trend. United Way of America (UW) (1996), for example, initiated an evaluation program that will monitor the outcomes achieved by the services provided by its funded agencies. The UW approach utilizes the concept of causal model in which as a result of the provision of the service some change or benefit is achieved in the community. The causal model, called by United Way “program’s logic model”, is the “If-then” sequence of changes that the program intends to set in motion through its inputs, activities and outputs. It is “a description of how the program theoretically works to achieve benefits for the participants.” Finally, the UW’s definition of outcome measurement states that outcome measurement does not prove that clients change as a result of the program. It simply shows that clients change. The necessity of spending resources to measure change without being sure of its causal relationship is subject to controversy and debate.

2.3.1 Outcome Measurement in Health Care

Historically, health professionals were able to treat patients as they wished, maintaining legal and professional standards of care as dictated by state laws and discipline-specific ethical principles. Fee-for-service policies and generous insurance reimbursements were assumed, and rarely did insurance carriers question the activities of the treating professionals. The industry has changed dramatically over the past several years. The demand has grown, but professional and financial resources have diminished steadily. Extremely high health care costs and a lack of commensurate increases in health have combined to convince the public they are not getting their medical care money’s worth (Wetzler, 1994).
As the delivery and reimbursement methods for mental and other health services change rapidly, measuring treatment outcome and client satisfaction has become critical (Plante et al., 1995). Some of these service delivery changes are the result of pharmacological advances, but much of the change is driven by pressure to contain health care costs and by payment systems that reward managed care companies and providers for doing so. Little is known, however, about the effect these changes have had on treatment effectiveness or on the longer-range costs that may be incurred if treatment is inadequate (Cacciola and Durrel et al., 1995).

To maintain preferred provider status and to establish mental health contracts with private insurance carriers as well as federal, state, and local government agencies, measuring outcome and satisfaction has become mandatory (Plante et al., 1995). Intense competition for funds and support has developed among health service agencies, which have been required to prove their effectiveness and their ability to use these limited resources efficiently. Legislative bodies continue to set more demanding standards of accountability for agencies that receive government funds (Coursey, 1977).

After studying hospitalization practices, healthcare researcher Noralou Roos concluded that there is “a need to revise our assumption that a logical model underlies medical practice. If one is to build a logical model, one needs an improved decision making framework to help us optimize treatments and lessen the pain of rationing“ (Wetzler, 1994). In short, we need a way to measure value. Wetzler provides the following definition of value in the health services sector:

\[ Value = \frac{Outcome}{Cost}. \]

This definition highlights the importance of measuring both, outcomes and financial measures such as cost.

Outcome measurement can satisfy numerous objectives for the different players in the behavioral and health care arena. Demonstrating treatment effectiveness, identifying which treatments at what level of intensity and duration are effective and appropriate for which types of patients and problems, and permitting active tailoring of treatment based on feedback from database (Cacciola and Durrel et al., 1995). There is
still, however, a lot of resistance among health professionals (particularly mental health professionals) to begin their own assessment programs, particularly because the notion of being held more accountable by insurance agencies results very unappealing. The medical establishment does not want to be required to demonstrate actual treatment results before a payment is authorized.

Nonetheless, many agencies and health providers have begun to design their own evaluation systems. Usually, the measurement of treatment outcomes in the health field is linked to customer satisfaction, and as such, follow-up questionnaires are sent to patients inquiring for their input in both areas, customer satisfaction and treatment outcome.

In response to the growing recognition of the limitation of traditional outcome measures in the health services, newer, more patient-oriented measures of outcome are receiving greater attention. One measure that has become quite popular is the Health Status Questionnaire. Developed at the Rand Corporation by John Ware, this tool, in addition to an overall evaluation of health (general health perception), includes seven separate dimensions of functional status and well-being as well as a depression screener (Wetzler, 1994). Because it includes such categories as physical and social functioning, mental health, energy/fatigue, and pain, the Health Status Questionnaire assesses the effects of treatments on those outcomes that patients have said to be important — providing the customer oriented vision of outcome evaluation.

Plante et al. (1995) present a study that measures treatment outcome and client satisfaction among families and children that underwent mental health treatment. These authors report four types of instruments used by agencies providing these type of mental health services that measure outcomes: Child and Adolescent Adjustment Profile (CAAP); a Brief Psychiatric Rating Scale for Children (BPRS); Customer Satisfaction questionnaire (CSQ) and a Demographic questionnaire (DQ). The BPRS and the DQ are completed by the clinicians. Parent/guardians provide their assessment filling out the CAAP scale at the first session and every 3 months during treatment and during a 6-month follow-up to termination. The CSQ is administered with the CAAP Scale at the
same intervals except the client’s first session. However, the implementation of such type of assessment system does not go smoothly. As Cacciola et al. (1995) points out, to begin the work of outcome measurement, a program must commit itself to the process. There must be sufficient institutional support for the process to succeed. Also, researchers in the health care field are developing groups of outcome measures and patient status indicators that reflect change in health status (Shaughnessy et al., 1994). This is done because each individual patient is a composite of many different attributes that reflect health status and aggregate outcome measures of care are not appropriate. Other health administrator groups are developing outcome measurement frameworks that try to relate the quality of the service provided to the outcome of the treatment. Some of these frameworks follow a P-D-S-A cycle such as the Total Quality Outcomes Management Model and are being implemented in hospitals and health-care agencies throughout the country (Joint Commission Mission, 1996; Melum and Bartleson, 1995; Spath, 1996). These approaches are the response of the health care industry to the quality revolution.

This research work has learned from these approaches. The frameworks followed by these researchers resemble the framework for development of outcome-based objectives. It is clear that both the outcome-based objectives development process (see Appendix A) and the performance measurement system proposed for social services can be translated and adapted for health care programs.

2.3.2 Outcome Measurement in Education

The process of outcome-based education is offered as a means of fulfilling the demands placed on education today. Outcome-based education is a student-centered, results-oriented design premised on the belief that all individuals can learn. This new paradigm in education attempts to meet the complex needs required of learners for the twenty-first century. (Boschee and Baron, 1993). Outcome-based education is:

- a commitment to the success of every learner
- a philosophy that focuses educational choices on the needs of each learner
• a process for continuous improvement

Outcomes are:

• future-oriented
• publicly defined
• learner-centered
• focused on life skills and context
• characterized by high expectations of and for all learners
• sources from which all other educational decisions flow

Because outcome-based education is centered on the customer (the learner), it is results-oriented and designed to meet the needs of all individual learners in achieving outcomes of life significance (Boschee and Baron, 1993). The assessment of such type of education is done through the assessment of what is called by the authors “authentic tasks”. Authentic tasks consist of student-centered activities that focus on content and skills that are useful in real life. All authentic tasks should contain the following elements:

1. Learner outcomes --specific content knowledge and skills that relate to the topic being studied.

2. Learner exit outcomes –outcomes developed by the district or school that particularly lend students to the knowledge and skills related to the topic being studied.

3. Complex thinking processes –students will use one or more complex thinking processes as they relate to the specific knowledge and skills associated with the topic being studied.
The assessment of authentic tasks is usually made in multiple validations over time in assessment sheets created by the teacher and school rather than by outside auditors.

2.3.3 Outcome Measurement in Social Services

Dissatisfied with well-intentioned but often unavailing programs, society is asking social service organizations to prove that their work is worth supporting (Magura and Moses, 1986). Many social services agencies have tried to identify service objectives and to develop accurate measures of progress toward them. Unfortunately, consensus on objectives has been hard to attain, and precise measurement of well being and the suitability of services have faced formidable conceptual and technical obstacles. In 1976, Shyne stated that after at least two decades of effort, the child welfare field, for example, still lacked generally agreed upon, valid, useful, and readily obtainable client outcome measures (Shyne, 1976). In fact, in the late 1970’s, many social service agencies tried to develop their in-house measures but the results were rarely disseminated. The outcome measures were criticized most frequently for being overly subjective, too general or simplistic; insensitive to case progress or changes over time; and of undetermined reliability and validity (Magura and Moses, 1980).

It is a fact that agencies would like to measure client status or outcome of their programs. The problem and weakness facing agencies is the development of actual instrumentation and evaluation methodology. As Magura and Moses state (1980) specifying objectives for clients is not identical to measuring the extent that objectives are achieved. Most social service departments have a file drawer filled with discarded “quick and dirty” instrumentation.

An in-depth study of seven states was conducted in, 1982 to determine the criteria and means used to measure the efficiency and effectiveness of the states’ social service programs. The study found that effectiveness information, such as “change in client status” is neither readily available nor is it easy to understood once collected. (Majchrzak et al., 1982).
Some noteworthy progress was made in the eighties in state-level implementation of case status measurement (performance indicators). Several coherent, representative systems for child welfare services are part of the Integrated Performance Management Reporting System in Illinois and the Voluntary Cooperative Information System (Magura and Moses, 1986).

Magura and Moses (1986) present a general model for program evaluation in social services. This model is consistent with that developed by Triantis and Medina-Borja (1996) to measure outcomes at the Agency. This model has the following components:

- Program objectives are statements of what a program is trying to accomplish. These statements may refer to processes (for example, improving efficiency of service delivery) or outcomes (such as improving service effectiveness).

- Program inputs are generally used in terms of money, facilities, staffing, equipment, and materials.

- Program operation variables are the structural arrangements of the agency and the procedures followed by practitioners that facilitate the provision of services. They determine the process by which resources are converted into program outputs. For instance, in the case of child protective services, a structural variable would be the existence of a specialized 24-hour intake unit, and a procedural variable would be the set of instructions for social workers regarding the initial family contact. Desirable program operations characteristics in social work are often termed “best practices”.

- Program outputs are the direct product of the system, usually represented by unit of service measures such as number of intake studies completed, hours of counseling provided, and so forth. The United Way of America and other state and local agencies have made several attempts to standardize service units in social services, but yet, there is no generally
accepted unit of service classification for most social services (Magura and Moses, 1986).

- Social impacts may be defined as the consequences of individual client outcomes for the wider community. To have social impact, client outcomes must be stable over time and pertain to social goals.

Methodological Concerns of Outcome Measurement

McKinlay (1996) advocates that different design approaches, measurement and data collection techniques must be employed for the evaluation of social services. Generally speaking, traditional evaluation methods as developed by behavioral scientists and epidemiologists (such as randomized controlled trials or case control studies) tend to be employed exclusively to measure outcomes of downstream interventions, in which individuals are the unit of analysis. As one moves upstream, the utility of traditional quantitative methods becomes problematic, not because they are intrinsically defective or flawed but because the phenomena to which they are applied (the units of investigation) are of a qualitatively different type (McKinley, 1993, 1994). Rigorous experimental control and manipulation are not always possible, especially when change is unexpected or unplanned.

Quite often, egregious methodological errors result from confusing upstream unit of random assignment (such as community, school, factory etc.) with a downstream unit of analysis (such as individual student, client or worker). When an intervention program is applied to an aggregate unit (e.g. community, school, work site) and the analysis is based on individual-level observations, the residual error is deflated by intra-cluster correlation and leads to overstatement of the statistical significance, not to mention the more important problem of measuring the wrong outcome. Intra-cluster correlations measure the strength of the relationships within a cluster. When one does cluster analysis there is a creation of similarity or dissimilarity groups. This, the intra-
cluster correlation may seriously overstate the Pearson correlations between the same variables overall.

Any reorientation to organizations, communities and national policies requires the development of measurement and indicators appropriate to that new level of focus. In contrast to the traditional measurements of individuals, systemic interventions must be assessed through the use of systemic outcomes —for example, how have you improved the community independent of individuals and their risky behaviors? In other words, QOL (quality of life) becomes QOC (quality of community) or QORG (Quality of organizational environment) (McKinley, 1993, 1994). The interest is not in whether an individual quits smoking but in whether there is improvement in the quality of organizational environment (how many workplaces are designated not smoking). What proportions of schools change the way school meals are prepared? Is there a reduction in the overall rate of avoidable death? The list of system outcomes is extensive, and the appropriateness of any is largely a function of the problem being addressed.

This section has outlined the need and importance of outcome assessment identified in different service areas, from health to educational services. It has also emphasized the methodologies used to measure outcomes and the methodological and practical concerns related to outcome measurement. All these have been considered to develop the instruments to elicit outcome data described in Chapter 4.

2.4 Program Evaluation

This section outlines the relationship between traditional program evaluation and the performance measurement approach proposed for this research. I will first present a description of terminology in the field, as well as the evolution that program evaluation has had in the last decades. I will also describe the types of evaluations available and the different paradigms related to them. Then I will show how measures of customer satisfaction and technical efficiency can be compared or included into what traditional evaluators call monitoring/process evaluation, and how the outcome-based evaluation approach for measuring effectiveness could be catalogued as part of what evaluators call the “evaluation of program utility.”
Program evaluation is concerned with collecting, analyzing and interpreting information on the need for, implementation of, and effectiveness and efficiency of intervention efforts to better humankind (Rossi and Freeman, 1993).

Cronbach et al. (1980) argued that program evaluation was initiated in the 1600s, when Thomas Hobbes and his contemporaries were concerned with devising numerical measures to assess social conditions and to identify the causes of mortality, morbidity, and social disorganization. The origins of formal program evaluation, however, can be traced back to the 1960’s, when we first began to see a major increase in social programs and requests to “evaluate” them. Throughout the ensuing years, definitions of program evaluation have remained fairly consistent (Schallock, 1995).

One good definition of program evaluation is the one provided by Patton (1986):

“Program evaluation is the systematic collection of information about activities, characteristics, and outcomes of programs for use by specific people to reduce uncertainties, improve effectiveness, and make decisions with regard to what those programs are doing and effecting.”

What is good about this definition in particular, is that it acknowledges the practical side of program evaluation as a decision making tool. Previous definitions are similar to this one but stressed the formal research side of this field, to such extent that the term “evaluation research” is used indistinctly as a synonymous of program evaluation.

Schallock (1995) points out two commonalities found in program evaluation definitions:

♦ Using research procedures to systematically collect information about activities, characteristics, and outcomes of social programs.

♦ Using program evaluation data (information) for decision-making and/or program improvement.
These two characteristics are also characteristics of what would be a good performance measurement system for social, educational and health services organizations and in that, a program evaluation system can be treated as part of a performance measurement system. If the PDSA cycle is closed, as Schallock suggests, and managers of service organizations make decisions regarding the way the services are provided based on the information obtained from evaluation, then this performance system resembles an organizational performance measurement and control system.

There have been a lot of changes in the approach to program evaluation and the methods used to perform it. In the 1960s, findings of limited effectiveness and serious questioning of the benefits-to-costs ratios of federal initiatives demonstrated the importance of undertaking program evaluation both before and after implementing programs and on a permanent and widespread basis and whenever modifications are made to them (Haveman, 1987). As a consequence of the apparent ineffectiveness of programs, the decade of the 1970s saw a lot of criticism of the continued expansion of government programs giving rise to fiscal conservatism and “sunset laws” -- the shutdown of ineffective programs (Freeman and Solomon, 1979; Adams and Sherman, 1978). The result was a change in emphasis in the evaluation field. In particular, there was increased concern with documenting the worth of social program expenditures in comparison to their costs, and increased attention to fiscal accountability, and the management effectiveness of programs (Rossi and Freeman, 1993). The 1980s therefore, saw increased concern with the balance of costs and benefits of programs, especially after the federal government experienced pressure for the curtailment of domestic federal expenditures. The emergence of new social problems during that decade, such as homelessness and AIDS also influenced new evaluation trends. In the 1990s the evaluation of programs and agencies continue to be a priority. According to Rossi and Freeman (1993), restraints on resources will continue to require choosing which social problem areas to concentrate upon and which programs should be given priority. Also, an intensive scrutiny of existing programs will continue because of the pressure to curtail or dismantle those for which there is limited evidence of program effectiveness and efficient delivery of services. Moreover, both because dissatisfaction with existing
programs and because of shifts in political and policy ideas, evaluations will continue to be needed to assess new and modified programs that hold the promise of being more effective and less costly. Also, in developing nations throughout the world, there is intense pressure to develop and evaluate social programs virtually overnight.

In fact, there is evidence of innovative re-conceptualization leading to promising new approaches and more sophisticated analytic approaches are emerging (Feldman and McKinlay, 1994). The potential for qualitative techniques is increasingly recognized, although there is, unfortunately, resistance from some quantitative groups. Scientists with vision are urging a re-orientation of efforts away from people (individuals) to places (communities) (McKinlay, 1996). Sufficient comparable studies have now been conducted to permit meta-analyses to identify specific characteristics of an intervention that produce particular results (Sellers et al., 1997). There is vigorous debate concerning the interpretation of results (Susser, 1995).

While the emphasis of program evaluation is on the evaluation of health, education, welfare and other human service policies and programs, it is not restricted to those areas. In fact, evaluation procedures have been applied to a variety of fields, including the procurement and testing of military hardware, research programs in the repair of major highways, and other areas a field from social interventions (Rossi and Freeman, 1993).

2.4.1 Paradigms in Program Evaluation

There are at least two different positions regarding the program evaluation paradigm. One, and perhaps the more widely disseminated is the influenced and adopted by the field of social psychology. Within this paradigm, the experimental model is the only way to perform program evaluation. According to Rossi and Freeman (1993), the single most influential article in the evaluation field was the 1969 paper by Donald Cambell. This paper outlined an ideological position that this author has advanced over several decades. His perspective was that policy and program decisions should emerge from the continual testing of ways to improve the social condition and that social change efforts should be rooted in social experimentation. The community and the nation should
be seen as a laboratory for social experiments. Campbell and his colleagues have since then sought to promote this position while refining and improving the methodology of social research in order to convert our society into and “experimenting society.”

In contrast, is the pragmatic viewpoint of Lee Cronbach (1982), another very well respected statistician and researcher. While acknowledging that scientific investigations and evaluation efforts may use the same logic of inquiry and the same research procedures, Cronbach argues that the purpose and intent of evaluations differentiate them from scientific investigations. According to him, long-term programs of scientific research are ill-suited to evaluation. Evaluation is an art, and every evaluation represents, or should represent, an idiosyncratic effort to meet the needs of program sponsors and stakeholders. In some cases, for example, evaluations may justifiably be undertaken that are “good enough” for answering relevant policy and program questions even though from a scientific standpoint they are not the best possible designs (Rossi and Freeman, 1993). This does not mean that many evaluations that take into consideration the needs of stakeholders and sponsors were not conducted in such a rigorous way. In many instances, evaluations can be maximally useful to decision-makers and also meet the requirements of scientific investigation.

Outcome evaluation can be catalogued in any of the two evaluation approaches, although most recent measurement initiatives are classifying it as part of the practical paradigm defended by Cronbach. In that, the measurement system proposed in this research is best-suited to the managerial evaluation of ongoing long-term programs and follows Cronbach’s view.

The substantial difference for many evaluation researchers would be the fact that most evaluation studies are controlled studies, some longitudinal over many years, others come to measure specific interventions effects, in two or more points in time. They are formal research that accounts for more or less all the variables involved in the intervention. Some measure outcomes when others measure outputs of the program and process characteristics. Generally, these formal evaluation studies are indeed very expensive, require expert knowledge from the beginning to the end of the study and data
generated by them is not easily interpreted by decision-makers. All these facts have made of program evaluation a separate field from performance assessment.

The approach taken by this research maintains that agencies need to have on hand a type of program evaluation that is continuous, internal to the organization and that provides decision-makers with up-to-date information on program performance. Not all agencies are capable of hiring research experts to evaluate their programs and not all interventions last enough time to make controlled experiments. So, it is highly beneficial to design an evaluation system that controls all important elements of organizational performance and that is part of the day to day operations of the organization. In this way, all programs will undergo a periodic evaluation and revision against benchmarks and goals and objectives facilitating performance improvement over the long run.

The approach advanced by this research does not, however, substitute formal evaluation research studies, but is complementary in those cases where impact evaluations are necessary from time to time. The evaluation proposed here will not only perform a type of outcome-based assessment, continuously collecting data from their customers on the benefits of the service received, but will allow agencies to gather data on the quality of the service delivery process and/or customer satisfaction. Other process measures that are generally administratively collected, such as outputs and inputs will permit the calculation of technical efficiency scores that can then be included into this measurement system.

2.4.2 Types of Evaluation Activities

Evaluations are undertaken for a variety of reasons, from assessing the utility or benefit of new programs to judge the worth of ongoing or past programs. This research approach concentrates in the evaluation of ongoing programs.

Evaluation research, then, is

“The systematic application of formal research procedures for assessing a program at different stages, from conceptualization, design, implementation, and utility of the intervention.” (Rossi and Freeman, 1993 p. ?).
These authors distinguish three major classes of evaluation research: (1) analysis related to the conceptualization and design of interventions; (2) monitoring of program implementation; and (3) assessment of program effectiveness and efficiency.

The first class, the *analysis of conceptualization and design of interventions* is closely related to the analysis of problems and needs in the community. Programs are justified as long as the problem that originated them persists. Diagnostic evaluation tools and activities are used for defining and specifying a social problem in ways that enhance both the design of appropriate interventions and their evaluation.

The second class of evaluation is involved with *monitoring of program* and is used to check that program managers conduct their day-to-day activities utilizing their resources as efficiently as possible. This type of evaluation is also called “*process evaluation*” because it evaluates process characteristics, such as cost, quality and resource utilization (McKinlay, 1996). Inefficiencies in this area are known to jeopardize the future of their programs themselves. A second objective of this type of evaluation is based on the fact that program sponsors and stakeholders require evidence that what was paid for and deemed desirable was actually undertaken.

Rossi and Freeman stress the importance of process evaluation as a management tool. These authors acknowledge that in many programs, regular feedback of evaluation information is one of the most powerful tools the program manager has for documenting the operational effectiveness of the organization. This type of evaluation can help justify the ways staff is employed, requesting further support, and defending the program’s performance compared with that of programs undertaken by other organizations in the same social program sector. Also, evaluating the impact or outcome of a particular project is important only if the delivery of the service has indeed taken place and served the appropriate participants in the way intended. In summary, monitoring provides a systematic assessment of whether or not a program is operating in conformity with its design and reaching its specified target population.

The third class of program evaluation according to Rossi and Freeman (1993), is the assessment of program utility. In this class, is critical to know both the
degrees to which a program produces the desired outcomes and its benefits in relation to its costs. These authors call the former as the program’s impact the latter as its efficiency. Together, these qualities are referred to as the program’s utility.

For evaluation researchers the term “effectiveness” describes “…how successful the programs are providing their intended target populations with the resources, services and benefits envisioned by their sponsors and designers” (Rossi and Freeman, 1993).

The performance measurement system proposed in this research would be a hybrid of the second and third class, monitoring the service quality and technical efficiency on a constant basis while evaluating outcomes or benefit/cost. The type of evaluation proposed here can be performed at each stage of a program, but will be implemented more frequently as a more managerial oriented type of monitoring and evaluation of program utility activities. The next section describes program monitoring.

2.4.3 Program Monitoring for Evaluation and Management

According to Rossi and Freeman (1993), program monitoring is said to be comparable with a set of “quality assurance” activities directed at maximizing a program’s conformity with its design. It is defined as the “…systematic examination of program coverage and delivery.”

There are two possible uses of program monitoring: i) assessing program coverage consists of estimating the extent to which a program is reaching its intended target population; and ii) evaluating program delivery consists of measuring the degree of congruence between the plan for providing services and treatments (program elements) and the ways they actually are provided. Aspects such as the compliance with regulations and laws, and the evaluation of the amount of inputs or resources being used to deliver the program are included here.
Of the two possible uses of program monitoring defined by Rossi and Freeman, this research effort will be concerned with the monitoring of the delivery of services.

Rossi and Freeman state that the monitoring of programs is directed at three key questions:

(1) the extent to which a program is reaching the appropriate target population,

(2) whether or not its delivery of services is consistent with program design specifications, and

(3) what resources are being or have been expended in the conduct of the program.

2.4.3.1. Process Evaluation as a Management Tool

Process evaluation is part of program monitoring and both terms have been interchangeably used. The term *process evaluation* is used to differentiate the research activities explained above from those directly involved in assessing program impact. These activities are also undertaken during the design and testing stages. Process evaluation has been an essential evaluation activity that oftentimes helped explain why outcome evaluations reveal little or no impact. Further, to estimate the benefits or effectiveness of a program relative to its costs, the evaluator must have information on resources expended by it.

Process evaluation studies how an intervention was implemented, including the study protocol and the conceptual (behavioral/cognitive) model employed, the fidelity with which the intervention was carried out; and the presence of competing programs or other influential elements. Process data are also useful for examining the relationships among program components to identify critical intervention elements for future practitioners (McGraw, 1996).
Increasingly, social and health researchers are recognizing that process evaluation may be as important as outcome measurement, because outcomes of health or other social programs become misinterpreted without it (McKinlay, 1996). Process evaluation data can be useful for understanding how a program operated and might help explain obtained outcomes. This is because defects in implementation might explain the extent of impact of a program. Without adequate monitoring or process evaluation, it is impossible to estimate if a program’s treatments are efficacious (Rossi and Freeman, 1993). Without information about what actually occurred during the implementation of an intervention, evaluators may not be adequately interpreting study results (Mcgraw, 1996).

Process evaluation permits the systematic exclusion of competing explanations for an observed experimental result. Most of the disappointing large-scale and costly community interventions reported in recent years had no process evaluation, so it is impossible to know why they failed or whether perhaps they succeeded on some other level. According to McKinlay (1996), when no effect (outcome) is observed for an intervention, process evaluation can answer the following crucial questions:

Is there no effect because the program was not properly implemented (implementation was not fully effective)?

Is there no effect because the program could not be fully implemented for some subjects (compliance was variable)?

Is there no effect because of barriers to program access?

On the other hand, if a beneficial effect is observed during the outcome measurement step, then process evaluation can answer the following questions:

Is the effect actually tied to the program, or is it due to the receptivity of selected subjects or target groups?
Is the effect actually due to the experimental program, or is it due to other competing, uncontrolled interventions (recognizing the reduction in experimental control at the aggregate level)?

Process evaluation or monitoring can be used as a management tool. Monitoring informs policymakers, program sponsors and other stakeholders. Most of the information provided to these stakeholders and necessary to efficiently manage is the same required by evaluators to assess the implementation of programs. Nonetheless, this has hardly happened. Historically, evaluators have undertaken monitoring activities as an adjunct to impact evaluations in order to estimate the extent to which the right treatment was being received by the appropriate targets. Such process analyses are usually undertaken independently of the efforts of program managers and staff to track and report on program activities. Indeed, evaluators, (whether insiders or outsiders) were expected to maintain as much independence as possible from program staff. These tacit requirements often meant that they collected their own data, performed their own analyses, and limited the feedback given to project staff until the evaluation was completed. In some cases, persons engaged in outcome evaluations insisted on initiating their own monitoring even though a program’s management staff was expending significant resources on similar monitoring activities (Rossi and Freeman, 1993).

Over time, the distinction between process evaluations and monitoring for management purposes has blurred and both terms are used interchangeably. Increasingly the trend is to plan and undertake monitoring in ways that both provide information relevant to explaining impact findings and to meet the needs of program managers and stakeholders. As Rossi and Freeman point out, the increasing use of management information systems, MIS, has helped both evaluators and managers, to share data and make their work conjoint.

The importance of process measures has been demonstrated in a variety of studies performed in all kinds of fields. A good example is provided by McGraw et al. (1996) who illustrate the complementary evaluation strategies of process and outcome evaluation as used in large multi-site trial called the Child and Adolescent Trial for
Cardiovascular Health (CATCH). The authors present the classroom curriculum component of the CATCH intervention to examine the relationship between teacher characteristics, measures of curriculum implementation and competing influences (all process evaluation elements) and student outcomes. CATCH was a collaborative-field trial, the purpose of which was to access the effectiveness of a school-based intervention. The design was based on a theoretical, causal model of behavior change in which modification of psychosocial factors would lead to changes in risk-factor behaviors, which in turn would result in improvement in physiologic endpoints (specifically, serum cholesterol). Grounded in social cognitive theory, the intervention model for this trial addressed organizational change and multiple individual student behaviors –specifically diet, exercise, and smoking. This longitudinal study was implemented in four study centers located in California, Louisiana, Minnesota and Texas. It encompassed 96 schools with 1908 school staff (Luepker et al., 1996). Process evaluation was developed to fulfill three functions during the trial: (a) to describe program implementation, (b) to provide information for quality control and monitoring, and (c) to explain program effects (understood as the result of variation in program implementation characteristics). In this model, teacher characteristics, staff training and competing programs were hypothesized to have direct effects on the degree of implementation of the classroom curricula, which, in turn, would have a direct effect on student outcomes. The observation lasted 3 years. In their conceptual model, the authors expected that implementation measures had a direct effect on student outcomes and mediated the effect of teacher characteristics.

The results of the analyses did not support the model. Instead of teacher characteristics moderating the effect of implementation, both, teacher characteristics and implementation measures had direct and independent effects on student outcomes. Teacher characteristics did not affect implementation. Measures of program implementation, such as fidelity (the percentage of classroom sessions modified by the teacher), were associated with changes in self-efficacy and knowledge. It appears that the teacher modification of the sessions had beneficial effect on the outcomes. Also, the modifications of the classroom curriculum made by the teachers might be the result of the teachers tailoring the specific lessons to the needs and interests of their students. It may
be that these modifications made the program more effective, but this last assertion was not tested in this study.

The McGraw et al. study has a number of implications for the field of program evaluation. The importance of the McGraw study in the context of this research is that these results support the significance of process measures (conceptual model, quality of service, implementation factors). It also highlights the importance of investigating the relationship between process characteristics and the actual benefits generated by such process (process measures vs. outcomes).

The following Section describes what Rossi and Freeman call the “program’s utility”, what are impact and efficiency assessments.

### 2.4.4 Evaluation of Program’s Utility

Two concepts are commonly used to assess program performance: effectiveness and efficiency. Magura and Moses (1986) provide definitions for both concepts: program effectiveness can be defined as the ratio of units of client outcome to some standard unit; this standard may represent the maximum improvement or deterioration possible in any given problem area or may represent some minimal level of adequacy as legally or culturally defined. For Magura and Moses, outcomes and technical efficiency ratios can be related by defining “efficiency” as the ratio units of client outcome either to units of program output or to units of resource input. Program output would be used in the definition when interest is focused on the relationship between work accomplished and client outcomes. Resource input would be used when interest is focused on the relationship between the amount of money spent for services and client outcomes (Hatry, 1978).

This section briefly reviews approaches to evaluate program’s utility. These include more traditional evaluation approaches to measuring outcomes, the so-called “impact assessments.” One could compare impact assessment methodologies to the approach proposed for this research to measure effectiveness in the sense that both estimate the benefits of a service or program on the target population. However, the similarities stop there. The kind of outcome-based evaluation that will be adopted for
this research is based on the approach proposed by the United Way of America (1995) and is not based on a strict experimental design for collecting the data. In that sense, it does not prove that the program causes the benefits, it only links the benefits to the service in a logic way.

Ordinarily, evaluators assess the outcomes of social programs by comparing information about participants and non-participants, or by making repeated measurements on participants, most commonly before and after the intervention, or by other methods that attempt to achieve the equivalent of such comparisons. The objective is to produce an estimate of the impact of an intervention uncontaminated by the influence of other processes and events that also may affect the behavior or conditions at which the social program being evaluated is directed (Rossi and Freeman, 1993). Rossi and Freeman emphasize the technical and managerial difficulties involved in undertaking impact evaluations, particularly due to the difficulty in reaching the subjects and/or in collecting reliable data.

According to Rossi and Freeman (1993) establishing impact is identical to the problem of establishing causality. Factors extraneous to the service or program can act as confounding factors, so that while the evaluation might be capable of proving that “A is a cause of B”, being A the program being evaluated and B the effect or outcome on the population, other factors may be influential. In social programs, the environmental conditions, the selection of participants and the existence of other social programs targeting the same population might be confounding factors among others.

2.4.4.1. **Impact Assessments**

For traditional program evaluation, the effects of the intervention or assessment of the outcomes of social programs are generally estimated through a type of evaluation called impact assessment. On the other hand, the costs of intervention are linked to benefits in what is called a cost-effectiveness evaluation.

An impact assessment weighs the extent to which a program causes change in the desired direction. The starting point for impact assessment is the identification of one or more outcome measures that represent the objectives of the
program. These are a set of operationally defined objectives and criteria of success. The objectives may be social-behavioral ones, such as lowering drug addiction or nutritional deficiencies among children; they may be community-related, such as reducing the frequency of certain crimes; or they may be physical, such as decreasing water pollution (Rossi and Freeman, 1993). They are used either to compare different programs or to test the utility of one program to the community. In that, impact assessment and outcome-based evaluation are comparable.

Because of practical problems, impact assessments are generally conducted using statistical approaches, such as non-randomized quasi experiments although sometimes it is possible to use classic experimental designs, which is the ideal situation.

**Design Issues**

In many circumstances, it is difficult to conduct impact evaluation using an ideal or best possible research design. Authors such as Rossi and Freeman are advocates of what they call the “good enough” rule. Following this rule, evaluators should choose the best possible design from a methodological standpoint, taking into consideration the potential importance of the program, the practicality and feasibility of each design, and the probability that the design chosen will produce useful and credible results.

2.4.4.2. **Efficiency Assessments**

Efficiency assessments in the evaluation field have a definition that differs from the engineering definition of efficiency. For evaluators it is generated by the relationship of costs to effectiveness (Rossi and Freeman, 1993). Efficiency assessments—cost-benefit and cost-effectiveness analyses—provide a frame of reference for relating costs to program results. This is a useful concept because in most cases, knowledge of
the impact a program had is not sufficient to make decisions regarding its implementation or maintenance. The results produced by a program must be judged against its cost, especially in this era of resource constraints. Programs with a high cost relative to their impact might not be supportable. Because of this, most commonly efficiency analyses in the social field take place after the completion of an impact evaluation, when the net impact of a program is known.

According to Rossi and Freeman (1993), estimating impact in comparison with costs can be tricky and arguable because it often requires both making assumptions about the dollar value of program-related activities and imputing monetary value to program benefits. The techniques to perform this type of evaluations are found in two closely related approaches: cost-benefit and cost-effectiveness analyses.

Cost-benefit analysis

A cost-benefit analysis requires estimates of the benefits of a program, both tangible and intangible, and estimates of the costs of undertaking the program, both direct and indirect. Once specified, evaluators need to adopt a specific economic approach to translate the benefits and costs into a common measure, usually a monetary unit (Rossi and Freeman, 1993).

The underlying principle is that cost-benefit analysis attempts to value both inputs and outputs at what is referred to as their marginal social values. The situation is most difficult when the good or service is not even traded, as is the case for most human-services benefits such as human lives, human health or quality of life, etc. For many items, such as the cost of providing a certain medicine or the monetary benefit of outfitting new cars with engines that burn less gasoline, market prices perform this task quiet well (Mishan, 1982).

Cost-effectiveness analyses

Because of the controversial nature of valuing outcomes into monetary values, in many cases cost-effectiveness analysis is seen as a more appropriate technique (Rossi and Freeman, 1993). Cost-effectiveness analysis requires providing monetary
value for the program’s costs; its benefits are expressed in outcome units. For example, the cost-effectiveness of distributing free polio immunizations to children could be expressed as: each 1,000 project dollars reduced the number of children affected with the illness in 10. So, when utilizing this approach the outcome of a program needs to have a numerical equivalent or be expressed as a numerical indicator.

In the section that follows I provide a thorough description of the DEA literature that deals with DEA and the measurement of service performance.

### 2.5 Data Envelopment Analysis and the Measurement of Service Performance

Data Envelopment Analysis is a non-parametric technique based on the principles of linear programming (LP) used to estimate measures of performance of a set of comparable producing units based on production characteristics such as technical efficiency and productivity (Charnes, Cooper and Rhodes, 1978; Banker, Charnes and Cooper, 1984). For DEA, entities responsible for the transformation of inputs into outputs are referred to as decision-making units, or DMUs. Originally, DEA was designed to assess how efficiently a firm, organization, agency, program site, or other such decision-making unit produces the outputs that it has been charged to produce.

DEA uses the set of producing units in the sample to construct a production or efficiency frontier consisting of all possible linear combinations of efficient producing units (the so called isoquant). By definition, then, any point on the frontier represents a feasible way (or "technique") for efficiently combining the set of inputs to produce a given set of outputs (Silkman, 1986). A complete DEA requires that one such linear program be solved for each DMU. The result is both the set of weights for the DMU and the measure of its relative efficiency or performance.

In brief, each producing unit is compared with all other producing units in the sample. A producing unit is said to be efficient if the ratio of its weighted outputs to its weighted inputs is greater than or equal to a similar ratio of outputs to inputs for every other producing unit in the sample (Silkman, 1986). An appealing property of the DEA
approach is that multiple-input, multiple-output technologies readily can be modeled without revenue or cost data. This is important in efficiency and productivity measurement in public production, where price or cost data are often unavailable (Lothgren and Tambour, 1996).

The following section presents the theory and mathematical formulations of the DEA technique.

The Table 2-1 below describes the notation that will be used throughout this document.

Table 2-1. Mathematical Notation Used in this document for DEA

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>n:</td>
<td>Number of [comparable] organizations or DMUs. n could also represent the number of time period for a single DMU; in this case, the analysis is time-based.</td>
</tr>
<tr>
<td>T:</td>
<td>Number of outputs</td>
</tr>
<tr>
<td>M:</td>
<td>Number of inputs</td>
</tr>
<tr>
<td>yrj:</td>
<td>Amount of output r produced by DMU j</td>
</tr>
<tr>
<td>xij:</td>
<td>Amount of input i used by DMU j</td>
</tr>
<tr>
<td>θ, h, φ, β, γ, ρ, π</td>
<td>Efficiency score of DMUj in subsequent formulation throughout this document</td>
</tr>
<tr>
<td>vi:</td>
<td>Weight given to input i (activity parameter)</td>
</tr>
<tr>
<td>ur:</td>
<td>Weight given to output r.</td>
</tr>
<tr>
<td>λj:</td>
<td>Weight given to DMU j (activity parameter) Also η, α, μ, σ, φ are used as weights in subsequent formulations of Chapter 3</td>
</tr>
<tr>
<td>ε:</td>
<td>A small positive number</td>
</tr>
</tbody>
</table>
2.5.1 DEA Review

When a process or unit has a single input and a single output, efficiency is defined simply as:

\[ \text{Efficiency} = \frac{\text{output}}{\text{input}} \]

However, in many cases this measure is inadequate because it is more common that organizational units have multiple incommensurate inputs and outputs. For these variables to be considered, they had to be combined into an aggregated form. A general formulation was proposed by Farrell (1957), so that this complexity can be incorporated in an efficiency measure by defining (relative) efficiency, as:

\[ \text{TE} = \frac{\text{Aggregate Measure of Output Performance}}{\text{Aggregate Measure of Input Performance}} \]

This definition requires a set of weights that can be difficult to define (Boussofiane, Dyson & Thanassoulis, 1991). Farrell’s approach tries to find the “best-practice” frontier. The “best practice” frontier is usually found by using ex-post data on inputs/outputs of decision-making units; this exercise is much easier as compared to finding the absolute frontier. The distinction between the two types of frontiers does not have much practical significance, since the two frontiers converge as the sample size tends to infinity (Triantis, 1997).

Therefore, as Girod explains it (1996) technical efficiency relates to an organization’s ability to either:

- Produce as much output as possible given a constant set of inputs (output-maximizing efficiency) or,
• Utilize as little input as possible given a constant set of outputs (input-minimizing efficiency).

Thus, one can study efficiency from an input point of view or from an output point of view. Output maximizing isoquants denote the maximum output level producers can achieve from a given set of inputs $X^0$.

Let’s assume that the DMUs are schools and that we are in charge of calculating the technical efficiency of school districts. In the simplest case, let’s consider only two inputs and two outputs. Inputs: i) number of students and ii) number of teachers. Outputs: i) number of students graduating and ii) standardized test scores in mathematics and English. The production system can be then, represented by Figure 2-5.

![Figure 2-5. Production System for District Schools](image)

The output-maximizing efficient isoquant in the case of a one-input producer set (A, B, C, D, E) all producing two outputs ($y_1$ and $y_2$) making use of a unit of single input $x$ is shown in Figure 2.6. The line called isoquant links all points representing the efficient production units or DMUs and for this case (output correspondence) represents one constant level of input. Output $y_1$ represents the number of students graduating and $y_2$ represents the average test scores. Schools B, C, and D are all efficient, but all display a different level of outputs. The outputs mix of DMUs A, B, and C and D differ. However, this not necessarily indicates inefficiency, it may be simply a reflection of market trends.
For school D, for example, it is more important to graduate less students though with higher test scores. For school B, on the other hand, even though their students score less on the tests, the unit has chosen to graduate more students. However, School E produces less graduating students with lower average test scores than those of school C, and yet, they both utilize the same amount of input (let’s say, number of teachers). Inefficient output bundles are separated from efficient bundles by joining adjacent pairs of bundles with a line segment; if the line segment has a non-positive slope and none of the other bundles lies to the north-east of it, the chosen bundles are declared efficient, otherwise, they are not.

By taking weighted averages of DMUs (for example, D and C), DEA generates a set of output bundles of hypothetical DMUs that envelope inefficient DMUs such as A and E.

This envelope, which constitutes a production possibilities frontier (PPF), serves as a benchmark against which the bundle of the DMU in question is compared. For DMU A, its technical efficiency is OA’/OA. The production possibilities frontier indicates the maximum amount of one output that can be produced given the amount of the other output. In that, the output increasing technical efficiency is the level by which the outputs produced by an organization can be increased without altering the level of inputs used.

The corresponding case for evaluating the DMUs from an input perspective (i.e. rewarding DMUs that use less resources to produce a constant level of output) is called "input-reducing" formulation. The concepts behind the "input-reducing" model are similar to the output increasing case. The difference is that in the input reducing case the frontier is convex to the origin, enveloping the data as shown in Figure 2-6.
2.5.1.1. Math Programming Formulation

The algebraic formulation follows:
Max $h_o = \frac{\sum_{r=1}^{t} u_{r}y_{rj}0}{\sum_{r=1}^{m} \frac{\sum_{i=1} v_{i}x_{ij}0}{\sum_{i=1}^{t} \sum_{j=1}^{n} y_{rj}0}}$ (2.1)

subject to

$\sum_{r=1}^{t} \frac{\sum_{i=1}^{m} v_{i}x_{ij}0}{u_{r}y_{rj}0} \leq 1, \ j = 1, \ldots, n$

$u_{r}, v_{i} \geq \varepsilon \ \forall r \text{ and } i$
The variables of the above problem are the weights and its solution produces the weights most favorable to unit j and also produces a measure of efficiency.

The DEA model in equation 2.1 is a fractional linear program but it may be converted into a linear form so that the methods of linear programming can be applied. For example, for output maximizing orientation, let’s assume a fixed use of input $x = 100$.

DEA models are usually not solved in their primal, but in their dual forms, which require less computational effort. In the primal form, there are as many constraints as decision-making units. In the dual form, there are as many constraints as inputs and outputs. Since it is more common to have more decision-making units than inputs and outputs, the decision-maker is better off computing the dual. Also, the primal will give the efficiency score, the projections of inputs and outputs to the efficient frontier and the weights for the input and output variables. The dual gives the peers for each DMU and their relative importance for determining the "virtual" DMU for comparison.

The resultant linear program is as follows:
Max \( h_o = \sum_{r=1}^{t} u_{r} y_{rj0} \) \hspace{1cm} (2.2)

subject to

\[
\sum_{i=1}^{m} v_{i} x_{ij0} = 100
\]

\[
\sum_{r=1}^{t} u_{r} y_{rj0} - \sum_{i=1}^{m} v_{i} x_{ij0} \leq 1, \quad j = 1, \ldots, n
\]

\( u_{r}, v_{i} \geq \varepsilon \quad \forall r \) and \( i \)

Or

Max \( \theta \) \hspace{1cm} (2.3)

\[ z \]

subject to

\[
\sum_{j=1}^{n} \lambda_{j} x_{ij} \leq x_{ij0} \quad i = 1, \ldots, m \quad (2.3 \text{ a})
\]

\[
\sum_{j=1}^{n} \lambda_{j} y_{rj} \geq \theta_{rj0} \quad r = 1, \ldots, t \quad (2.3 \text{ b})
\]

\[
\sum_{j=1}^{n} \lambda_{j} = 1 \quad (2.3 \text{ c})
\]
Where,

- The right hand side of (2.3 a) gives the potential output production of the DMU whose efficiency is being measured which, if the DMU is inefficient, will be greater than one;

- The left-hand sides of (2.3 a) and (2.3 b) represent the hypothetical DMU formed by taking weighted averages of the real DMUs for each input and output.

- The first set of constraints (2.3a) indicates that the weights must be such so that the hypothetical DMU for each input uses not more than what is being used by the DMU whose efficiency is being analyzed;

- The second set of constraints (2.3 b) indicates that the weights will generally be chosen so that the hypothetical DMU produces outputs in the same proportion as the DMU whose efficiency is being chosen. The inequality is to allow for the possibility that the DMU whose efficiency is being measured is on the flat segment of the production possibilities frontier;

- The third constraint (2.3 c) ensures that the hypothetical DMU is operating at a scale similar to the one under which the DMU for which technical efficiency is being measured.

One of the main characteristics of the DEA routine is that the u and v weights are unknown and will be picked to maximize the ratio of the weighted sums. So, the program will find the set of weights that make the DMU look better. In the classical DEA formulation, the weights are allowed to vary from DMU to DMU (Girod, 1996). Although in several new formulation the concept of weights restriction has been explored. To illustrate and further explain this formulation, let’s go back to our school
district example in Figure 2.6. Let’s calculate the efficiency score for school A in the graphic, which produces outputs:

\[ x_1 = 19 \text{ graduating seniors and} \]
\[ x_2 = \text{average test score of 3} \]

We had assumed a fixed input level of 100 teachers. The formulation will be as follows:

\[ \begin{align*}
    \text{Max} & \quad \theta \\
    z & \\
    \text{Subject to} : \\
    19 \lambda_A + 50 \lambda_B + 30 \lambda_c + 10 \lambda_D + 20 \lambda_E \geq 19 \theta & \quad (\text{for } y_1 = \text{number of graduating students}) \\
    3 \lambda_A + \lambda_B + 2 \lambda_D + 2 \lambda_E \geq 3 \theta & \quad (\text{for } y_2 = \text{average test scores}) \\
    \lambda_A + \lambda_B + \lambda_C + \lambda_D + \lambda_E = 1 & \\
\end{align*} \]

In the case of A, we know that the hypothetical DMU is a linear combination of points C & D, so that:

\[ \begin{align*}
    \lambda_B = \lambda_E = \lambda_A = 0 \\
    30 \lambda_C + 10 \lambda_D = 19 \theta \\
    3 \lambda_C + 4 \lambda_D = 3 \theta \\
    \lambda_C + \lambda_D = 1 \\
\end{align*} \]

Solving algebraically, we obtain:

\[ \theta = \frac{90}{79} = 1.14 \quad \text{meaning that the school can increase its outputs in 1.14 to be efficient with the same level of inputs} \]

The results for the corresponding weights are:
\[ \lambda_c = 0.42 \]
\[ \lambda_D = 0.58 \]

The section that follows presents the DEA assumptions and discusses how these assumptions can be verified with a real data set.

2.5.1.2. Assumptions and DEA Postulates

The DEA analysis is based in six basic production technology assumptions. If the production function is the abstract mathematical equation that describes how quantities of the inputs \( X_1 \) and \( X_2 \) are combined to produce quantities of output \( Q \), represented by:

\[ q = f(x_1, x_2) \], then,

Production technology is understood as all the technical information necessary for the production of output \( Q \), including all physical possibilities to combine the inputs \( X_1 \) and \( X_2 \).

The six postulates are:

1. *Free production is not feasible, i.e. one cannot have positive outputs without the usage of some inputs.*

2. *Infinite production is not feasible, i.e. finite inputs cannot produce infinite outputs.*

3. *Proportional increases in inputs do not decrease outputs (“weak” disposability of inputs).*

4. *Proportional decrease in outputs remains producible with no change in inputs (“weak” disposability of outputs). Another way of stating the same assumption is that a proportional increase in outputs cannot be obtained if inputs are reduced.*

5. *The correspondence between inputs and outputs is closed. This assumption allows for the definition and the existence of an isoquant.*
6. The isoquant is convex to the origin; in a two dimensional input space this indicates that if two input bundles can each produce one unit of output, then so can the weighted average of them.

The verification of the applicability of the above basic DEA assumptions to a specific case —i.e. specific production system or unit— is a combination of analytical and mathematical tests. One can analyze the production system as a “blackbox” to which inputs are introduced and transformed into outputs. An analytical test can be performed. The verification of assumptions 1 and 2 can be done in this way. To verify assumptions 3, 4 and 5, non-parametric tests have been developed (Afriat 1972; Hanoch & Rothschild, 1972; Diewert & Parkan, 1983; Varian, 1984; Banker & Maindiratta, 1985; and Fare, Grosskopf, & Lovell, 1987).

When DEA was first presented by Charnes, Cooper and Rhodes (1978) the assumption was that the technology presented constant returns to scale. Later, Banker, Charnes and Cooper developed a modification to that formulation to account for the many cases in which the technology does not behave as constant regarding the returns to scale. Descriptions of the two original basic formulations (the CCR and BCC models) are presented here.

CCR Model

Charnes Cooper and Rhodes (1978) recognized the difficulty in seeking Farrell’s common set of weights to determine relative efficiency. They proposed that each unit should be allowed to adopt the most favorable set of weights in comparison to other units so that the efficiency of a target unit $j_o$ can be obtained by solving the following model:

\[
\text{Maximize the efficiency of unit } j_o \\
\text{Subject to the efficiency of all units being } \leq 1.
\]

The variables of the above problem are the weights and the solution produces the most favorable weights to unit $j_o$ and also produces a measure of efficiency (Boussofiane et al., 1991). The algebraic model is as follows:
Considering the $j_0$ decision making unit, $DMU_{j_0}$, for which production plan $(X_{j_0}, Y_{j_0})$ was observed.

\[
TE_{CCR} (X_{j_0}, Y_{j_0}) = \max_h \frac{\sum_{r=1}^t u_r y_{rj_0}}{\sum_{i=1}^m v_i x_{ij_0}} \quad (2.4)
\]

subject to

\[
\frac{\sum_{r=1}^t u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \leq 1 \quad j = 1, \ldots, n
\]

\[
u_r, v_i \geq \varepsilon, \quad \forall \ r \ and \ i
\]

where

$TE_{CCR} =$ Charnes, Cooper and Rhodes interpretation of Farrell’s (1957) efficiency measure for $DMU_p$

$y_{rj} =$ amount of output $r$ from unit $j$

$x_{ij} =$ amount of input $i$ to unit $j$,

$u_r =$ the weight given to output $r$,

$v_i =$ the weight given to input $i$,

$n =$ the number of units

$t =$ the number of outputs

$m =$ the number of inputs

$\varepsilon =$ a small positive number
In the solution to this model the efficiency of unit \( j_0 \) (representing the unit under analysis) is maximized subject to efficiencies of all units in the set having an upper bound of 1.

The dual form of this linear problem follows:

\[
\min \quad \theta - \varepsilon \sum_{r=1}^{t} s^+_r - \varepsilon \sum_{i=1}^{m} s^-_i \quad \text{for } DMU_0
\]

subject to

\[
x_{j0} \theta = \sum_{j=1}^{n} x_{j} \lambda_j + s^-_i \quad i = 1,2,3,\ldots m
\]

\[
\sum_{j=1}^{n} y_{j} \lambda_j = y_{r} + s^+_r \quad r = 1,2,3,\ldots t
\]

\[
\lambda_j \geq 0 \quad j = 1,2,3,\ldots n
\]

\[
s^+_r \geq 0 \quad r = 1,2,3,\ldots t
\]

\[
s^-_i \geq 0 \quad i = 1,2,3,\ldots m
\]

\[
\theta \quad \text{unconstrained}
\]

Charnes, Cooper and Rhodes (1978) model assumes constant returns to scale: if a producer proportionally increases all inputs by some amount, outputs will be proportionally increased by the same amount. Banker, Charnes and Cooper (1984) observed that the constant returns to scale model gave distorted results when comparing decision-making units of significantly different sizes. Thus, these authors proposed a new formulation of DEA capable of remedying all size-related problems. This formulation is commonly referred to as the BCC model.

**BCC Model**

Banker, Charnes and Cooper (1984) modified the original CCR formulation to account for the variable returns to scale. The envelopment form of the BCC model would be the same as the dual for the CCR model but with an additional constraint:

\[
e\lambda = \sum_{j=1}^{n} \lambda_j = 1 \quad (2.6)
\]
The constraint above tightly envelops all the data and in doing so, accounts for variable returns to scale. Usually, under variable return to scale assumption more DMU's are deemed efficient. This variation to the original formulation allowed all sorts of production technologies to be analyzed that prior to this did not fit the basic CCR assumption.

The BCC responds to an axiomatic framework. This is written in terms of four postulates:

The estimated production possibility set $T$ must satisfy:

1. Convexity postulate.
   
   If $(X_1, Y_1)$ and $(X_2, Y_2) \in T$ and $\lambda \geq 0$ then $(\lambda X_1 + (1 - \lambda)X_2, \lambda Y_1 + (1 - \lambda)Y_2) \in T$

2. Monotonicity postulate.
   
   a) If $(X_0, Y_0) \in T$ and $Y \leq Y_0$ then $(X_0, Y) \in T$
   
   b) If $(X_0, Y_0) \in T$ and $X \geq X_0$ then $(X, Y_0) \in T$

3. Inclusion of observations postulate.

Each of the observed vectors $(X_j, Y_j) \in T (j=1, 2, ..., n)$

4. Minimum extrapolation postulate. $T$ is the intersection set of all $T^*$ satisfying the earlier three postulates

   The above four postulates determine a unique production possibility set based on the observed input-output data

### 2.5.2 DEA formulations for the Treatment of Special Cases

This section reviews the available DEA formulations for the treatment of special cases of variables in the input/output set, such as categorical variables, ordinal variables, controllable and non-controllable variables, the inclusion of qualitative attributes of inputs and outputs, and the inclusion of outcome data, etc.
2.5.2.1. **DEA and Non-Discretionary (non-controllable) Variables: Categorical and Continuous Approaches**

Several authors in the DEA community have addressed the problem of incorporating the effect of non-controllable, environmental characteristics in the formulation. Examples of external variables include the form of ownership, location characteristics, demographic characteristics, geographic factors, labor relations, government regulations, etc. This is particularly important for public and nonprofit production systems, as explained in section 2.2.1. These formulations assume that the characteristics of the external environment could influence the ability of management to transform inputs into outputs.

According to Ruggiero (2000), public sector production —which can be compared to nonprofit production— is characterized by a strong influence of environmental factors on service outcomes, leading to the existence of multiple frontiers.

Several examples of these different formulations have been applied to homes for the disabled (Blank and Valdmanis, 2001), nursing homes (Fried et al., 1999), school districts and public production in general (Ruggiero and Duncombe 1995; Ruggiero, 1996). These authors follow the tenet that omitting environmental effects is particularly detrimental in public production systems of social nature, since the omission may lead to incorrect conclusions about the potential of certain units and about the overall best-practice in the production of services. According to Blank and Valdmanis (2001), without adjustments or modifications to the formulations DEA scores may also contain inefficiencies due to the external influences that are beyond the firm's managerial control, and cause a biased predetermination of actual performance.

In addition, several of these environmental characteristics are usually represented as variables of "categorical" nature. Categorical variables, or variables that use a finite number of integer numbers to represent different ranges or levels of a variable, have been studied in the DEA literature (Banker & Morey, 1986). Likert type scales, and other scales used in surveys can be classified as being categorical. According to Fried et al. (1999), previous work on external environment can be broadly classified
into three categories: the frontier separation approach, the all-in-one approach and the

two-stage approach.

**All-in-one Approach**

Fried et al. (1999) describe what they call the "all-in-one" approach to
incorporate environmental variables in the DEA model as normal inputs and/or outputs,
directly in the LP formulation along with the traditional inputs and outputs. An advantage
of this approach is that it allows for the inclusion of more than one environmental
variable and that it is not restricted to categorical variables. The resulting efficiency score
takes into account the environment by treating these variables exactly as conventional
inputs or outputs. One "disadvantage" stated by these authors is that it requires that the
environmental variable be classified a priori as either an input or an output, when it might
be difficult to do so. They explain:

"An additional input enables production of more outputs, holding
efficiency constant. If the operating environment enters the LP formulation as an input,
this implies that more outputs could be produced, suggesting that the operating
environment is favorable. If the operating environment enters the LP formulation as an
output, this implies that more inputs are required, suggesting that the operating
environment is unfavorable." Fried et al. p. 250

This classification into inputs or outputs is not suitable if the objective is
to test whether a determined environmental factor is favorable or unfavorable to
production. In addition, the efficiency score is based on the assumption that inputs are
shrinkable and outputs expandable (if output oriented), which might make no sense for
certain environmental variables but not for others. Furthermore, the inclusion of
environment in the normal formulation will give rise to an inappropriate selection of the
reference set since the projected value will be the weighted average of the actual values
of the peers also for the environmental variable. This means that if the level of an
environmental variable treated as input for unit A is 20, and for unit B is 10, representing
two different operating environments, chances are that DEA would calculate the
projected target value for an inefficient unit C —operating at a level of environment 25—as being 15, which in practical terms may or may not make sense at all.

All the above limitations of this approach make it not suitable for this research work.

Alternatively, in order to overcome the above difficulties, Banker and Morey (1984) developed a formulation for non-controllable characteristics of continuous nature. This problem is solved in two stages as described in Banker and Morey:

\[ \text{Minimize } h_0 = \theta_0 - \varepsilon (\sum_{i=1}^{m'} s_i^- + \sum_{r=1}^{t} s_r^+) \quad (2.7) \]

Subject to:

\[ \sum_{j=1}^{N} \lambda_j x_{ij} + s_i^- = \theta_{ij} x_{ij} \quad (i = 1,2,\ldots,m'), \quad (2.8) \text{ discretionay inputs (controllable)} \]

\[ \sum_{j=1}^{N} \lambda_j z_{ij}^e + s_i^- = z_{ij}^e \quad i = (m'+1, m'+2,\ldots,m), \quad (2.9) \text{ non-discretionary inputs (uncontrollable)} \]

\[ \sum_{j=1}^{N} \lambda_j y_{ij} - s_r^+ = y_{ij} \quad (r = 1,2,\ldots,t), \quad (2.10) \text{ outputs} \]

\[ \sum_{j=1}^{N} \lambda_j = 1, \quad (2.11) \]

\[ \lambda_j \geq 0 \quad (j = 1,2,\ldots,j_o,\ldots,n), \quad (2.12) \]

\[ s_i^- \geq 0; \quad s_r^+ \geq 0 \quad (i = 1,2,\ldots,m; \quad r = 1,2,\ldots,t), \]

2nd stage:

\[ X'_{ij} = \theta^*_o X_{ij} - s_i^- \quad (i = 1,2,\ldots,m') \quad (2.13) \]

This formulation differs of the standard DEA formulation in that it separates the input variables in controllable and non-controllable and for the second type. These authors indexed the input so that the first m’ inputs are controllable and the other m-m’ are non-controllable based on the formulation provided by the same authors (1984) with one restriction for controllable inputs and another for uncontrollable.
Also, the formulation does not link the efficiency score to the non-controllable inputs' constraint in 2.9. The non-controllable variables are used to construct the reference set only and do not affect the efficiency score obtained by the DMU being evaluated.

The optimal level of the variable \( \theta_0 \), denoted \( \theta_0^* \) from the first stage LP, (2.7)-(2.12) provides the so called Farrell score (a radial measure of inefficiency) in which each controllable resource can be simultaneously “contracted” by the factor \( \theta_0^* \).

The equation (2.13) on the second stage provides the maximum resource conservation potential for the \( i_{th} \) controllable resource.

The above formulation is consistent with the BCC axiomatic framework for DEA (1984).

\[
T = \{ (X, Y) \mid X \geq \sum_{j=1}^{N} \mu_j X_j, Y \leq \sum_{j=1}^{N} \mu_j Y_j, \sum_{j=1}^{N} \mu_j = 1, \mu_j \geq 0 \}
\]

This formulation is useful and valid to assure that the value of the projection to the frontier is the weighted average of the values of the composite set of the non-controllable variable. However, whenever the variable represents a level of environmental harshness that can be used to group the variables, so that the peers can only be obtained from units in the same or worst level, then this formulation is neither useful nor recommended.

**Frontier Separation Approach: Mathematical Formulation for the Case of Non-Controllable Categorical Variables**

Banker and Morey (1986) modified DEA to incorporate on-off characteristics (categorical variables of the nature 0-1) that represent the presence or not of an attribute. They modified the DEA formulation to include categorical variables instead of continuous data in an input reducing orientation. They have concluded that by using this approach, DEA loses relatively little discriminating power. They approach the
problem in cases when categorical variables are either controllable or uncontrollable by the manager.

Banker and Morey (1986) developed models to accomplish the above goals, i.e. for selected uncontrollable variables that are, at the same time categorical representing factors in scales, the authors relaxed the requirement in DEA that a constant marginal productivity situation applies. Such as certain demographic characteristics that can be grouped in scales representing ranges or degrees of intensity of the variable, such as cost of living, competition, level of wealth, etc. The intuition behind this formulation is that when the value of a variable is not continuous but categorical, an increase in inputs cannot expect a corresponding increase in the value of outputs.

The number of DMUs identified as technically inefficient will not increase under the categorical treatment, since the reference groups are more restrictive than in the case of continuous data. They observed, however, that the efficiency score is slightly smaller when the continuous treatment is used instead of the categorical approach, since it is less constraining.

The logic behind it follows. When a given characteristic is present in some DMUs and not in others (such as the existence of a drive-in capability in some banks and not in others) it is desired that the composite reference members (the peer group) be constructed from DMUs which are in the same category or possibly from those which are operating in an even more unfavorable situation. For example, suppose that the probability of a Agency field office geographical jurisdiction to suffer natural disaster is either low, medium or high. Then a field office with a high natural disaster probability should only be compared to other field offices with the same rating. To do otherwise, say based on some sort of scale, would imply that a “constant marginal productivity” applies.

Now, considering the case where one of the non-controllable inputs, say $i = m$ is a categorical variable (a variable that has a finite number of values that the variable can take on). The problem with this is that the convexity postulate provides that all convex combinations of the $n$ referent points may be used as points to evaluate the relative performance of the $j_{th}$ DMU. However, for a categorical variable, a convex
combination needs not to belong to the categorical scale and it may not even have a meaningful interpretation.

Further, the convexity axiom, along with the minimum extrapolation axiom implies a piecewise linear surface for the production possibility set and, therefore, the marginal productivity is constant on each such linear segment. However, for a categorical scale, this assumption of constant marginal productivity may not be valid. Hence, Banker and Morey modified the BCC postulates by deleting the operability of the convexity postulate for the categorical variable. In this situation, the production possibility set \( T \) will then be given by:

\[
T = \{(X, Y) \text{ where } X = (X_1, X_2, \ldots, X_m) \text{ and } Y = (Y_1, Y_2, \ldots, Y_R) \\
\text{where there exists a set } \{\lambda_j, j = 1, 2, \ldots, N\} \text{ characterized by } \lambda_j \geq 0 \\
(j = 1, 2, \ldots, N) \text{ and } : \sum_{j=1}^{N} \lambda_j = 1 \text{ such that } : Y_r \leq \sum_{j=1}^{N} \lambda_j Y_{rj} (r = 1, 2, \ldots, R), \\
X_i \geq \sum_{j=1}^{N} \lambda_j X_{ij} (i = 1, 2, \ldots, m - 1) \text{ and } X_m = X_{ml} \\
\text{for all those indices } l \text{ for which } \lambda_i > 0\}
\]

For the LP formulation, they add \( k \) new variables where \( k+1 \) is the number of values the categorical variable \( X_m \) can take on. For example, if the level of severity can be described as high, medium, and low (3 levels), then three descriptor binary variables \( (d_{mj1} \ldots d_{mj3}) \) need to be defined for each of the DMUs to replace the continuous variable \( X_{mj} \) that describes severity. If the jth DMU is actually in the “none” level for the \( m_{th} \) input, one can capture this by setting like this

<table>
<thead>
<tr>
<th></th>
<th>( d_{mj1} )</th>
<th>( d_{mj2} )</th>
<th>( d_{mj3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Medium</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>High</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
The constraint set (2.9) in the Banker and Morey formulation is deleted and is replaced by three new constraints, namely

\[ \sum_{j=1}^{N} \lambda_j d_{m,j}^{(1)} \leq d_{m,j_0}^{(1)}, \quad (2.14) \]
\[ \sum_{j=1}^{N} \lambda_j d_{m,j}^{(2)} \leq d_{m,j_0}^{(2)}, \]
\[ \sum_{j=1}^{N} \lambda_j d_{m,j}^{(3)} \leq d_{m,j_0}^{(3)} \]

The approach for non-controllable categorical variables is proven to assure that the referent composite DMUs consist only of DMUs matched more carefully with the DMUs being evaluated in regard to the categorical variable characteristics (for instance, population categories, market size, etc.).

Charnes et al. (1994) however, proposed a different approach. Instead of focusing on changing the LP model formulation, a better and more easily applied approach is to adapt the LP solution algorithm. This has the additional advantage, which can be significant in extensions, of allowing multiple categorical variables. The proposed modification rests on the assumption that there is natural nesting or hierarchy of the categories. It is then desirable that each DMU be compared only with DMUs in categories that represent equal or more disadvantageous operating conditions. If the categories are not comparable, the authors state (e.g. public universities vs. private universities) then a separate analysis should be performed for each category.

Take an input variable that can assume one of L levels (1,2,…,L) which partition the set of DMUs into categories. Specifically, the set of DMUs \( D = \{1,2,...,n\} = D_1 \cup D_2 \cup ... \cup D_L \), where \( D_k = \{i : i \in D \text{ and input value is } k\} \) and \( D_j \cap D_k = 0, j \neq k \).

The following model specification allows decision-making unit \( DMU_o \in D_k, K \in \{1,...,L\} \) to be evaluated with respect to the units in \( \bigcup_{k=1}^{L} D_k \).
Two-stage Approaches to Incorporate the Environment.

This formulation includes inputs and outputs in the traditional LP formulation to compute radial technical efficiency. The technical efficiency score is then used as the dependent variable in a second stage regression, where the explanatory variables measure the external environment. Some studies use ordinary least squares (OLS) to estimate the second stage equation (Ray, 1991), others use a Tobit model (MCCarty and Yaisawarng, 1993). An advantage of the two-stage approach is that the influence of the external variables on the production process can be tested in terms of both sign and significance. However, a disadvantage is that most models ignore the information contained in the slacks and surpluses. This might give misleading conclusions regarding the impact of each external variable on efficiency.

Fried et al. (1999) presented a model to separate the management component of inefficiency from the other components, defined to be outside the control of management within the timeframe of the analysis. This approach utilizes the input slacks or the output surpluses (depend upon the orientation of the model) and re-calculates a measure of technical efficiency that adjusts for differences in the operating conditions.
environment across production units in a four-stage procedure. The main contribution is to generate a pure measure of managerial inefficiency.

The first stage is to calculate a DEA frontier using the traditional inputs and outputs according to standard production theory. The external variables are excluded. Efficiency scores, slacks and surpluses are computed for each observation.

In the second stage, a system of equations is specified in which the dependent variable for each equation is the sum of radial and non-radial input slack for an input-oriented model or radial plus non-radial output surplus for an output oriented model. The independent variables measure features of the external environment and are not restricted to be identical across equations. The authors state that this system of equations identifies the variation in total by variable measures of inefficiency attributable to factors outside the control of management.

The third stage is to use the parameter estimates (coefficients) from the second stage to predict the total input slack or output surplus, depending upon model orientation. These predicted values represent the "allowable" slack or surplus, due to the operating environment, and are used to calculate adjusted values for the primary inputs or outputs.

In the fourth stage, Fried et al (1999) re-run the DEA model under the initial output specification, using the adjusted data set. The new radial efficiency measures incorporate the influences of the external variables on the production process, and isolate the managerial component of inefficiency.

Separate Frontier Approaches in the Case on Non-controllable Continuous Environmental Variables

To account for public production technology, characterized as receiving a greater impact from the environment than private firms, Ruggiero (1996) presented an extension to the Banker and Morey (1986) model for environmental exogenous factors. Ruggiero's model targets environmental factors that are continuous instead of categorical.
Characterizing $Z$ as the vector of exogenous inputs $Z_t$ for which the lesser values imply a harsher operating environment for the DMU, we have the production possibility set:

$$T(Z) \equiv \{(X, Y) : Y \leq f(X \mid Z)\}$$

Where $Y$ represents the outputs produced as a function of the inputs $X$ given the environment level $Z$.

According to this author, the Banker and Morey model for continuous variables that includes non-controllable variables (1984) would underestimate the level of technical efficiency if for the composite reference group determined $(X^*_o, Y^*_o) \notin T(Z_o)$. This happens because the restrictions are placed on the level of $Z_o$ for the composite reference group but not for the DMUs that should be included in the reference set based on the level of environmental harshness. Thus, an optimal solution for $DMU_o$ may include in the reference set $DMU_j$ where $Z_j > Z_o$. Consequently, without further restricting the reference set, a solution $(X^*_o, Y^*_o)$ may not be feasible in real and practical sense.

To account for this environmental impact on production,—i.e. DMUs with a more favorable environment are able to produce more with less inputs— Ruggiero extends the production postulates described before for DEA to include the effect of the environment by introducing an adding postulate in order to assert the relationship between the environment and the production frontier.

"Postulate 5 (Environmental effect) $T(Z_1) \subseteq T(Z_2)$ for all $Z_1 \leq Z_2$." P. 557
Recognizing that not always the optimal solution might be feasible if the peer groups used are not operating at the same environment, Ruggiero states the modified DEA model as:

\[ \begin{align*}
\text{Min } & \quad \eta_o - \varepsilon \left( \sum_{i=1}^{M} s_{io}^+ + \sum_{r=1}^{S} s_{ro}^- \right) \\
\text{s.t. } & \quad \sum_{j=1}^{n} \lambda_j y_j - s_o^- = y_o, \\
& \quad \sum_{j=1}^{n} \lambda_j x_j + s_o^+ = \eta_o x_o, \\
& \quad \sum_{j=1}^{n} \lambda_j = 1, \\
& \quad \lambda_j \geq 0 \text{ } \forall j \in Z_j \leq Z_o, \\
& \quad \lambda_j = 0 \text{ } \forall j \in Z_j > Z_o, \\
& \quad s_o^-, s_o^+ \geq 0 \\
& \quad \varepsilon > 0 \text{ is non-Archimedean}
\end{align*} \]

The model above is based on the belief that to assure that the resultant composite reference group is feasible, the reference set needs to be constructed ONLY from those DMUs with at least as harsh an environment as the one that DMU \( o \) faces. Ruggiero's model achieves this by constraining the multipliers to zero for DMUs with a more favorable environment. In that, the efficiency score determined from this model is greater than or equal to the score determined from the Banker and Morey model.

The same author (1998) acknowledged later than the performance of the above model declines as the number of non-discretionary inputs increases (i.e. the discriminatory power of the model is greatly reduced). Ruggiero then developed a new model to handle multiple exogenous factors to account for the fact that in his previous model as the number of continuous fixed factors increases, the possibility of identifying a DMU as efficient by default also increases. This is because the model is unable to properly weigh the importance of each non-discretionary variable in production and thus also ignores comparisons between two DMUs that overall can be classified as having the same or worse environment. For example, let's say a field office can be classified under 2
different environmental variables: aggregated household income and % minority in the community. Field office A presents an aggregated household income of $1000 and 34% minority. Field office B presents an aggregated household income of $900 and a % minority of 36%. Less income is considered harsher environment to operate as well as a higher % of minority in the population. An expert observer would consider both field offices as comparable while Ruggiero's model would not put them in the same group, and as such could not be compared. This is particularly true if two DMUs present different values only in one particular non-discretionary input and are comparable in all others. If you reduce the number of DMU's in your set, then the likelihood of more DMU's been considered efficient increases. Hence, Ruggiero's model will more likely overstate efficiency.

To overcome these difficulties, Ruggiero proposes a three-stage procedure in which he runs a normal DEA excluding the environmental variables, then he uses Ray's procedure regressing the environmental variables against the efficiency scores to obtain an equation with parameters $\beta_i$ for each environmental variable $z_i$ and in the third stage he makes use of Ray's second stage regression model (1991) to obtain an overall index of environmental harshness from Ray's regression's parameters as:

$$z = \sum_{i=1}^{g} \beta_i z_i \quad (2.17)$$

After the construction of $z$ the following third-stage linear program is solved:

$$\text{Min} \quad \theta \quad (2.18)$$

s.t.

$$\sum_{j=1}^{J} \lambda_j y_i^j \geq y_i^o \quad \forall i = 1, \ldots, S,$$

$$\sum_{j=1}^{J} \lambda_j x_k^j \leq \lambda x_k^o \quad \forall k = 1, \ldots, M,$$

$$\lambda_j = 0 \quad \text{if } z^i > z^o,$$

$$\sum_{j=1}^{J} \lambda_j = 1,$$

$$\lambda_j \geq 0 \quad \forall j = 1, \ldots, J,$$
The problem with this three-stage approach is that it requires functional form specification of the regression equation. Mis-specification of the form could lead to distorted measurement.

The original model formulated by Ruggiero in equation 2.16 will be used in the formulations developed in Chapter 3 for continuous environmental variables that can classify the DMUs by the harshness of the environment. The Banker and Morey formulation for non-controllable variables described in equations 2.7 to 2.13, will also be used to develop those formulations for nonprofit production in Chapter 2. The objective is to include more than one environmental variable besides the one who classifies the DMUs in categories of environmental harshness.

2.5.2.2. Mathematical formulation for the Case of Controllable (Discretionary) Categorical Variables

In the case that the categorical variables are controllable, Banker and Morey (1986) proposed a mixed-integer programming approach to deal with categorical outputs, such as quality ranking like in a Likert scale or level of service.

Suppose that a qualitative characteristic can be described in L different categories from category 1 to category L ordered in an increasing level of such characteristic (from poor to excellent service, for example). Then, one defines parameters \( \{w_{ij}: l=1,2,..L-1\} \) as a service descriptor vector, one for each DMU \( j \), where the lowest service category is represented by the vector 0’s, the second lowest category is represented by the vector (1,0,…0), the third lowest by the vector (1,1,0,0) and so on.

<table>
<thead>
<tr>
<th>Category</th>
<th>( W_1 )</th>
<th>( W_2 )</th>
<th>( W_3 )</th>
<th>( W_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor (0,0,0,0)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fair (1,0,0,0)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Good (1,1,0,0)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Very Good (1,1,1,0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Excellent (1,1,1,1)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
The problem is formulated as the following mixed-integer LP model as follows:

$$\text{Maximize } \sum_{i=1}^{L-1} t_i \quad (2.19)$$

Subject to:

$$\sum_{j=1}^{N} \lambda_{ij}x_{ij} \leq x_{ij}, \quad \forall i = 1, 2, 3, 4,$$

$$\sum_{j=1}^{N} \lambda_{ij}Y_{ij} \geq Y_{ij}, \quad \forall r = 1, 2,$$

$$\sum_{j=1}^{N} \lambda_{ij}w_{ij} - t_l = w_{ij}, \quad \forall l = 1, \ldots, L - 1,$$

$$t_{l-1} \geq t_l, \quad \forall l = 2, \ldots, L - 1,$$

$$\sum_{j=1}^{N} \lambda_{ij} = 1, \quad \lambda_{ij} \geq 0, \quad \forall j = 1, 2, \ldots, N,$$

where $t_i$’s are 0–1 variables.

The third and fourth constraints are the additions where $t_l$s are 0-1 variables. The third constraint ensures that DMUs included in the composite referent point are necessarily of the same or higher service orientation. Additionally, one must ensure that the levels of the categorical variable (improvements) for the $j_0$th DMU can only occur by sequentially converting the elements of the service vector. For instance, if the $(l-1)^{th}$ slot is 0, then the element in the $l^{th}$ slot must also be zero. The fourth constraint together with the fact that the $t_1$s are 0-1 variables, insures that if $t_{l-1}$ is 0, then $t_1 = 0$.

However, Kamakura (1988) criticized the work of Banker and Morey (1986) in reference to controllable categorical variables only. This author demonstrated that the mixed-integer LP model above specified could lead to incorrect evaluations. Kamakura states that the last set of constraints in the BM model

$$t_{l-1} - t_l \geq 0 \quad l = 2, \ldots, L - 1$$
leads to an incorrect evaluation of the DMUs. This is because for certain cases, when this formulation is used for categorical controllable outputs, the prescribed optimal solution may not make sense, since the projection could pertain to a category of less value than the original. For example, if the categorical variable is service quality level, say 4, the target can be a service quality of 3, which does not make sense. This is because in certain cases, the potential solutions generated by certain set of peers have not been well studied in the formulation. Kamakura proposes the following modification to solve this problem: replacing the equation above with

\[ t_l - t_{l-1} \leq W_{t-1,jo} - W_{t,jo} \quad (2.20) \]

where the \( W_{ij} \) – “dummy” variable representing the level of the categorical output for DMU \( j \). This constraint pertains to the level of categorical outputs and not to binary stacks as in the previous one. This new constraint assures that the 0-1 descriptors are consistent with the ordered categories after improvements in the discrete output. Incorporating this, the model will identify possible improvements in the categorical output as intended. The author recognizes that this model will not be capable of identifying inefficiencies that result from the continuous input-output variables. To solve this problem he develops the following model:

\[
\begin{align*}
\max \quad & \theta + \epsilon \left( \sum_{r=1}^{K} \bar{s}_r + \sum_{l=1}^{N} \bar{s}_l^+ + \sum_{l=1}^{L} t_l \right) \\
\text{s.t.} \quad & \sum_{j=1}^{N} \lambda_j Y_{rj} - s_r - \theta Y_{ro} = 0 \quad (r = 1,2,\ldots,K) \\
& \sum_{j=1}^{N} \lambda_j X_{ij} + s_i = X_{ijo} \quad (i = 1,2,\ldots,M) \\
& \sum_{j=1}^{N} \lambda_j W_{ij} - t_l = W_{ijo} \quad (l = 1,2,\ldots,L-1) \\
& \sum_{j=1}^{N} \lambda_j = 1 \quad (2.\) \\
& t_l - t_{l-1} \leq W_{t-1,jo} - W_{t,jo} \quad (l = 2,\ldots,L) \\
& s_r, s_l^+, \lambda_j, \theta \geq 0
\end{align*}
\]
2.5.2.3. *DEA Approaches to Measure Service Quality*

Service quality and customer satisfaction barometers can often be regarded as complementary to productivity measures since the latter seldom take customer perceived quality into consideration. If customer satisfaction is an explicit objective for a firm, the inclusion of such measures into productivity indices leads to more valid performance measures (Lothgren & Tambour, 1996). A unit running at high levels of output but of low quality may be efficient in providing output volume for its resource level but it is not indicative of high overall performance (Thanassoulis et al., 1995). In fact, several authors have incorporated quality and/or customer satisfaction into the DEA formulations utilizing different approaches (Burgess and Wilson, 1998; Fare, Grosskopf and Roos, 1995; Lothgren & Tambour, 1996; Olesen & Petersen, 1995; Sola and Prior, 2001; Thanassoulis, Boussofiane & Dyson, 1995).

In particular, Thanassoulis et al. (1995) proposed a model for perinatal care in England in which the output set incorporates both activity levels and quality measures. They extended the model to service quality measures (number of very satisfied mothers and number of satisfied mothers with the service) and quality of medical outcomes measures (survival rate of babies at risk as a surrogate measure). They assumed constant returns to scale but developed a model that could account for differences in the importance of outputs (very satisfied mothers are more important than dissatisfied mothers). Three extensions to the basic DEA model were used by these authors (1995):

i) a weights based DEA model, incorporating weight restrictions;

ii) a weights based target DEA model;

iii) an ideal input-output levels based DEA approach.

Olesen and Petersen (1995) developed three models for incorporating information on quality differences into DEA. A relationship between the models and the microeconomic production theory are established. In the models developed, the quality
of a unit of output is measured on an ordinal scale. Two representations of quality of production are suggested: i) the cumulative quantities of outputs produced at or above any given level of quality and ii) the cumulative probabilities that units of a given output can be produced at or above any given level of quality. The three approaches used are:

1. Each output is desegregated into different types according to the cumulative quantities of the output produced at or above any given level of quality.

2. The cumulative probabilities for each output to be produced at or above any given level of quality are included as a characteristic of the production along with quantities produced.

3. The cumulative probabilities for each output to be produced at or above any given level of quality can be seen as the only characteristic of the production, i.e. the quantities produced are ignored.

To compute the probabilities, the following is done. Each of the outputs is measured on a number of quality dimensions. Then scores are assigned in ordinal intervals. Then the output measure is aggregated over the quality levels to give the cumulative quantity of each output produced in each dimension in the interval. Then the probability is the cumulative probability of producing one unit of the same output in the above quality interval in the same dimension.

Fare, Grosskopf and Roos (1995) incorporated qualitative attributes into the production model using a Malmquist approach. This approach utilizes the attributes together with distance functions to measure service quality and then calculates productivity. By further imposing a separability assumption on the distance functions, these authors decomposed the Malmquist productivity index into three components, namely, quality change, technical change and efficiency change. The technology constitutes all outputs and attributes that are feasible for some input vector.

\[ S' = \{(x^i, y^i, a^i): x^i \text{ can produce } y^i \text{ and } a^i\} \]

Fare and Grosskopf treat attributes as outputs and state that if one so wants, one may think of the vector \((y^i, a^i)\) as the output vector. Assuming data from
inputs and outputs is available for two different periods in time $t$ and $t+1$, they define two different input distance functions that serve as components in the quality index. Their input-based measure of quality is based on the retail industry (pharmacies) for which they developed this application that requires pharmacies to supply drugs at the lowest possible cost.

The first distance functions measure the maximal feasible reduction in $x^i$ given $(y^i, a^i)$ using information from one period. Information from mixed periods is also considered in the index. The quality (or quality change) index for the technology between periods 0 and 1 is defined as:

$$Q^{01}(y^1, a^1, x^1, y^0, a^0, x^0) = \sqrt[2]{\frac{D_i^0(y^0, a^1, x^0)D_i^1(y^1, a^0, x^1)}{D_i^0(y^0, a^0, x^0)D_i^1(y^1, a^0, x^1)}} \quad (2.22)$$

This index depends on three things in addition to the attribute vectors $a^0$ and $a^1$, namely the technology as expressed by $D_i$, $t=0,1$, the resource $x$, $t=0,1$, and the output vectors $y$, $t=0,1$. Thus, changes in these will affect the quality index, which makes the index

“…an integrated part of the production process, and clearly shows that the index measures the (change in) quality of the production process.” p.138

A fourth approach is that developed by Lothgren and Tambour (1996, 1999) called the Network Model. The network model is based on the work of Fare et al (1991) and Fare and Grosskopf (1996) regarding the incorporation of intermediate products in DEA by the way of a network activity analysis model. The network activity analysis model explicitly recognizes that some inputs are produced and consumed within the production technology before reaching the final output node. The Network model for customer satisfaction will be explained in the next section.

The Network Model for Customer Satisfaction and Service Quality

The fundamental idea of the network model is that production and consumption processes can be represented as separate nodes. This network approach is
flexible in the sense that production and consumption can be jointly modeled along with a broad representation of customer satisfaction. Output characteristics or attributes, such as service quality characteristics are intermediate products that generate customer satisfaction. Figure 2-7 depicts this network.

The network model allows an allocation of resources between customer oriented activities ($x^c$) and traditional production ($x^p$). A sub vector of the outputs ($y^f$) representing characteristics and quality attributes of the service is treated as an intermediate input in a consumption technology. Moreover, quality assessments ($q$) (interpreted as customer reported satisfaction) as well as the “ordinary” outputs ($y^p$) are considered as final exogenous production from the consumption and production node respectively.
The technology is then represented by distance functions, which also define the efficiency and productivity measures. The distance functions are estimated by DEA.

The following section presents additional approaches that deal with service quality and/or customer satisfaction and DEA.

Two stage Approaches to Incorporate Service Quality in DEA

Soteriou and Zenios (1999) considered the production process of branches of a bank as one that takes as input one or more links of the service-profit chain and delivers as output the next link(s). The service-profit chain is the characterization of a bank's production technology from the service to the actual realization of profits as the final output. According to the authors, there are three nodes (or models as the authors call them) that capture the components of the service-profit chain and benchmark: (i) operational strategies and service delivery systems, (ii) service delivery concepts, and (iii) customer satisfaction. These concepts can be used to validate the chain.

Supported by this framework, the operating strategy and service delivery system model consists of inputs such as workplace design, job design, procedures, policies, customers and the outputs of this node are services transactions, work, etc. In the service delivery concepts (results for customer model) inputs are the same as above and the outputs are quality and profits. In the customer satisfaction node, inputs are workplace design, job design, procedures, policies and customer perceptions (perceived quality), customer retention, referrals and loyalty. Outputs of this third node are revenue growth.

The authors assumed variable returns to scale so that non-proportional effects of the returns to scale are accounted for if they exist. At each stage of this cascade of models some or all the drivers of performance hypothesized in the service-profit chain can be included.

Homogeneous branches were then benchmarked against each other to identify those units that excel. According to these authors, it is important to compare
homogeneous units in order to isolate the effects of exogenous factors on the efficiency in the profit chain. The effects of exogenous factors can be assessed by comparing the results of two or more distinct groups of homogeneous units (Soterious and Zenios (1999); Zenios et al. (1999)).

Athanassopoulos (2000) presented an optimization framework that includes capabilities, service quality and performance of a firm in a two stage DEA model. He applied this framework to financial institutions.

According to this author, the literature shows that the concept of firm performance has taken the form of several variables, some of them being market share, repurchasing intentions, customer retention, profitability, and cost position. These elements of performance, however, cannot be seen as independent and thus, the issue of the definition of performance metrics should be emphasized along with the definition of service quality areas.

In this research, Athanassopoulos applied the capabilities-service quality-performance (C-SQ-P) paradigm to the banking industry to extend his (1997) and (1998) research. Athanassopoulos' former approach advocated the need to bridge service performance and quality towards a global view about service provision effectiveness. He proposes an effort-effectiveness function where the production and intermediation efficiency of bank branches is linked directly to the dimensions of service quality. In this research, however, service quality was treated as a control variable in order to facilitate the estimation of attainable efficiency targets for individual bank branches. The author used the projection of the service quality variable at their maximum attainable level for each DMU and as such, he eliminated the effects of quality—controlled for quality—while analyzing the efficiency of the bank branches.

In his 2000 work, the author uses four propositions in order to set the scheme of the proposed C-SQ-P framework:

1. In a retail-banking environment, firm capabilities will be given strong human and process orientation.
2. Service quality is considered to be an outcome of the strategic choices made by the firm to invest in human capital and also develop appropriate processes. This C-SQ-P framework will seek to estimate the best level of SQ that corresponds to given levels of firm capabilities.

3. Performance is a multidimensional construct and thus it will be assessed by using a multi-input-output representation of the operating process of individual branches. This concept is called Leibenstein’s X-Efficiency whereby branches will be considered efficient if for a given bundle of inputs they maximize their bundle of outputs and/or for a given bundle of outputs produced they minimize their bundle of inputs used. This is the equivalent to the traditional concept of technical efficiency in DEA.

4. The effect of firm capabilities on service quality is assumed to be supportive while there is no predetermined assumption as to whether service quality has positive or negative effect on performance (Schneider, 1991). Figure 2-8 below is a schematic representation of the above model.

However, the operationalization of the particular proposition requires specific modeling assumptions that predetermine the exact linkage between service quality and performance. This framework can be represented mathematically with the following equations:

\[ Service\ Quality = F (Capabilities, \ Size\ of\ Activity) \] (2.23)

\[ Performance_j = F (Service\ Quality, \ Branch\ Size, \ Cost, \ Outputs, \ Workload) \] (2.24)

Service quality and performance are represented in (2.23) and (2.24) as multi-attribute value functions. The multi-attribute value expressions about service quality and performance as illustrated in (2.23) and (2.24) constitute objective functions. The mechanism of this two-stage process of assessing the effectiveness of bank branch operations is depicted in Figure 2-9.
The set of equations implied in (2.23) seeks to estimate a branch-specific level of service quality, which is only "perceived" by an external or internal customer. This service quality dimension is a function of capabilities, or the number of employees and their level of training, education, experience and motivation, and of the size of the activity or branch size. Following that, service quality is given a direct role on the estimated overall branch performance in (2.24). The overall performance takes into consideration the operating cost, and outputs such as demand deposits, time deposits, savings deposits, etc plus the workload, for which the authors used mean length of queues as a proxy measure. The set of equations in (2.23) and (2.24) suggests that service quality is seen as a resource level, which should be capitalized in the form of expanding the value creation of individual branches.

Figure 2-8. Capabilities, Service Quality and Performance in the Banking Industry (Athanassopoulos (2000)
Figure 2-8 depicts the model that considers capabilities as the drivers of service quality. The identification of the maximum level of service quality provision, given the profile of the branch, is the first stage of this assessment. Having obtained the efficient levels of service quality for individual branches, the second step is to use these estimates to derive output augmentation targets.

In figure 2-9, the service quality performance of branch K is assessed by projecting it into the EB segment of the service quality frontier under variable returns to scale. The virtual DMU K’ (the projection to a point that will represent DMU K if it were in the best-practice service quality frontier) shows the service quality level had it fully utilized its corresponding capabilities. The efficient level K’ is used as an input into the assessment of the X-efficiency of branch K as illustrated on the right part of the frontier. Conversely, the X-efficiency of branch K will be estimated by projecting K’ to K^2 of the segment HD. According to Athassanopoulos, this artifact permits the
identification of a greater potential for improvement than if the maximum potential for
service quality would have not been considered in the first place. In the figure, the
underestimation would have been by an amount $K_1K_2$.

The DEA formulation that translates the above situation follows
(Athanassopoulos, 2000):

**FIRST STAGE:**

$$\text{Max} \sum_{\nu \in J} w_i \theta_{\nu \nu} \quad (2.25)$$

$$\sum_{j=1}^{n} \mu_j x_{ij} \leq x_{ik} \quad i \in I \quad \text{(internal factors)}$$

$$\sum_{j=1}^{n} \mu_j q_{ij} = \theta_{\nu \nu} q_{\nu \nu} \quad r \in Q \quad \text{(service quality features)}$$

$$\sum_{j=1}^{n} \mu_j = 1$$

$$\theta_{\nu \nu} \geq 1, \mu_j \geq 0$$

Where $q_{ij}$ is the service quality feature $r$ of DMU $j$ and $w_r$ is the
preferential weight assigned by the decision maker for the service quality feature $q_r$.

**SECOND STAGE:**

$$\text{Max} \sum_{\nu \nu \nu \nu} \omega_{\nu \nu} \phi_{\nu \nu} - \sum_{\nu \nu \nu \nu} \omega_i v_i \quad (2.26)$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} - v_i x_{ik} = 0 \quad i \in D \quad \text{(discretionary inputs)}$$

$$\sum_{j=1}^{n} \lambda_j y_{ij} - \phi_{r \nu} y_{\nu \nu} = 0 \quad r \in O \quad \text{(output quantities)}$$

$$\sum_{\nu \nu \nu \nu} \lambda_{\nu \nu \nu \nu} = y_{\nu \nu \nu \nu} \quad p \in F \quad \text{(non discretionarry outputs)}$$

$$\sum_{j=1}^{n} \lambda_j \theta_{\nu \nu} q_{ij} \leq \theta_{\nu \nu} q_{\nu \nu} \quad r \in Q \quad \text{(service quality level)}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\phi_{\nu \nu} \geq 1, v_i \leq 1, \lambda_j \geq 0$$
In the two stages presented above, the measure of efficiency is non-radial, allowing outputs and service quality attributes to vary in different proportions to reach the frontier. As the projections of service quality from stage one are carried to the second stage, the efficiency score of the second stage presumes the maximum level of quality for that DMU. A virtual behavior is considered in the sense that in reality, the projections that are input to the second stage are not the real observed level but the target recommended by DEA for that particular DMU. It assumes that the operations have undergone a re-engineering intervention so that the DMUs present the best possible picture. This attainment of the maximum possible projected level of output is not always possible in real life; thus, it is a virtual behavior instead of considering a real behavior. This is the main disadvantage of this model.

The overall idea of having a cascade of models and connected by making the outputs of one model the inputs to the next, is the main contribution of this study that is used in this research in the second model presented in Chapter 3. Also, the concept of using non-radial measures instead of radial measures to account for variables that not necessarily will increase in the same proportion as others in a preferential weights structure, where the weights are assigned by the decision maker to inputs and outputs or only to the output side is adopted in both models presented in Chapter 3.


Sola and Prior (2000) evaluated effectiveness, efficiency and quality in a sample of Spanish public hospitals. These authors adopted a definition of technical quality—the % of hospital acquired infections—as a proxy measure of the level of health care quality. They recognize that this is more of an "undesirable" output and that it represents more the "lack of quality", having negative consequences on levels of patient health and involving more resources. Even though the authors enter on a discussion about the relationship among effectiveness, quality and productivity or efficiency, the actual application of the model did not use any effectiveness measures per se. The limitation of this study accepted by the authors is that this measure evaluates only the
technical or operational level of quality setting aside important issues in health care service delivery such as behavior of personal care, the comfort of hospitalized patients, etc. This definition of quality makes it feasible to establish a rate of substitution between the number of cases treated and the level of technical quality in the cases treated. In order to evaluate this, they used the so called Malmquist index, which can relate the movements on productivity and quality between two time periods and establish the specific position corresponding to each hospital. DEA is generally a static approach, which provides a one-shot evaluation of the DMU’s performance. Malmquist productivity indexes (Malmquist, 1953), is one of the several dynamic approximations developed with the purpose of quantifying the evolution of productivity over a period of time. Ten years later, Fare, Grosskopf, and Roos (1994) broke down the Malmquist index recognizing two sources of productivity change: (i) efficiency change or "catching up" effect. Changes in productive efficiency (i.e. distance of an observation from the frontier), and (ii) technical change, or changes in production technology, which are normally referred to as "true technological change" because it requires a shift in best-practice technology.

The most known definition of Malmquist productivity index takes the geometric mean of two indexes related to the above concepts: efficiency change and technical change. Both calculate the change for years t and t+1. In contrast to the static DEA applications, the Malmquist productivity index decomposition requires the computation of six linear programs for each unit under analysis. This is because this index represents three nodes and for each node, variables in the present and in t+1 are considered, so a linear program for each point in time is solved (2 for each node) (Sola and Prior, 2000). These authors modified this index based on the work of Fare, Grosskopf and Roos (1995) who introduced quality in the Malmquist indexes and redefined the notion to solve ten linear programs for each DMU. This formulation can be decomposed into three factors, depicted in Figure 2-10 as described in Sola and Prior (2000):
A DEA Application to Measure Service Quality Construct in Marketing (Tankersley, 2000)

Tankersley (2000) presented an application of DEA as a managerial tool in the field of marketing since DEA has been reported as a more appropriate tool to aid managers in the provision of services, particularly for the management of service quality. In particular, this author attempts to use DEA to provide adequate levels of each of the various dimensions that comprise service quality. Particularly because DEA makes recommendations that attempt to optimize the level of service quality, not necessarily always "maximize" it. (What is the difference between maximization and optimization? I am confused.) This is particularly important in instances where service quality has a high cost and beyond a certain point, diminishing returns are encountered. This does not make sense.

This author does this by utilizing DEA as a technique where service quality measures serve as the inputs to be managed based on the model presented by Dabholkar, Thorpe and Rentz (1996) and the outputs included are measures a retail manager intends to maximize. Table 2-2 presents these inputs and outputs.
Table 2-2. Inputs and Outputs used by Tankersley (2000) to Determine Levels of Service Quality Dimensions

<table>
<thead>
<tr>
<th>INPUTS</th>
<th>OUTPUTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Physical Aspects</td>
<td>1. Net profit as a percentage of Sales</td>
</tr>
<tr>
<td>2. Reliability</td>
<td>2. Customers/labor hour</td>
</tr>
<tr>
<td>4. Problem Solving</td>
<td></td>
</tr>
<tr>
<td>5. Policy</td>
<td></td>
</tr>
</tbody>
</table>

2.5.2.4. *DEA Approaches to Measure Service Outcomes (Effectiveness)*

The inclusion of outcome measures to measure efficiency and effectiveness of social, health, educational and public programs using DEA invariably fall somewhere between three approaches found to include outcomes in the DEA formulations.

1. Some authors have reported the measurement of cost-effectiveness (Byrnes et al., 1997; Finkler & Wirtschafter, 1993; Nyhan & Martin, 1999). Cost-effectiveness requires that both cost and outcome terms be included. The Finkler & Wirtschafter (1993) study in particular, did not incorporate outcomes in the DEA formulations, but calculated the cost efficiency scores for the set of DMUs and then plotted those results against the outcome results for each unit, creating a cost-effectiveness frontier. This model has an input reducing orientation since the idea is to reduce the cost to produce a given level of output.

The Byrnes et al. (1997) study also calculates cost-effectiveness, but it does incorporate outcomes (as a change measure of functioning) in the DEA formulation. This study also gauges cost-efficiency (providing a given level of service for the least cost). Cost-effectiveness refers here to achieving a given level of outcome for the least cost. For example, achieving a given level of abstinence for a day of drug and alcohol treatment services. These authors establish that cost-effectiveness, then, refer to the
second linkage in the production process described in Figure 2-11—the linkage between activities and outcomes and as such, they measure only outputs and outcomes. This DEA model has an input reducing orientation for a fixed level of outcomes and calculates “cost”—effectiveness utilizing the units of service activity (outputs) as inputs. Outcomes then are treated as outputs in the formulation.

Considering outcomes as outputs in the formulation of the model is perhaps the most common approach to include outcomes in DEA formulations. The unique contribution of the Byrnes et al. study is the inclusion of activities (or outputs) as inputs. Outcomes are then treated as outputs. In another recent study, Nahra, (2000) applied DEA to drug treatment facilities in order to evaluate performance vs. differences in ownership. Her major contributions are two: one is to include measures of process effectiveness as an output of the unit as a proxy measure that represent treatment characteristics that have been shown to result in better outcomes. She does this recognizing that the best outcomes would be, in fact, measures of the degree of recovery of the patient. However, this author also recognizes that obtaining such level of metrics is very problematic. She also combines more traditional measures of output (such as therapy hours) with these indicators. The input side of Nahra's model is very traditional, including measures of resources used and cost treatment indicators. The study employs DEA with the intent of identifying the sources of inefficiency and getting efficiency scores as a basis for statistical comparison of units. The second major contribution is that Nahra utilizes a modification to the DEA formulation developed by Andersen and Petersen (1993) whereby the units along the frontier are no longer bound by the assigned value of one. The basic motivation for this formulation is to compare the unit under evaluation to a frontier defined by all units except itself. This is done so that peer groups and rates of improvements can be found even if the DMU in question is effective.

Rouse et al. (1997) took a different approach for evaluating effectiveness in highway performance in New Zealand. These authors run the DEA model three times, each time obtaining a different performance measure. The first run obtained the traditional TE measure (with inputs and outputs as variables). The second run obtained an effectiveness measure (with outputs and outcome as variables (this last one
operationalized as indexes of maintenance). The third run calculated a measure of “economy” with outcomes as outputs and inputs as variables.

Figure 2-11. The Production of Social Services (target population A). (adapted from Byrnes et al. (1997)

2. Another approach is that of researchers (e.g. Thanassoulis, 1995) who have included output measures [that could be also interpreted as outcomes] in their formulations. This has happened because most program outcomes need to be measured through indicators, or proxy metrics. Sometimes, a valid immediate-outcome indicator is a measure that could be considered output under another context. An example is the study reported by Thanassoulis (1995) on the assessment of British police forces. The output measures utilized are “crime clear-ups” defined as:

“...the ratio of violent crimes cleared up to violent crimes recorded, in both cases during 1991.” P.643.

One could argue that the number of violent crimes cleared-up could have served as an indicator of the outcome: crimes solved, a possible immediate outcome. Therefore, the rate would be an effectiveness measure (in the Byrnes et al. sense). The DEA assessment adopted the output orientation, consistent with the notion that crime
levels are largely outside the control of police and efficient operation should result in higher clear ups rather than lower crime level (a long-term outcome measure). The assessment was run assuming constant returns to scale.


In particular, Ruggiero and Duncombe (1995) presented a modified method for estimating technical efficiency differentiating discretionary inputs and socio economic variables affecting public production. The modified DEA model controls for environmental factors, comparing a producing unit only to other units that face a similar productive environment. This method is applied to the analysis of local school districts in New York State, where schools facing harsh environments may not be able to produce outcomes as high as those with more favorable environments (e.g. the influence of family background in academic achievement). They calculate what they call “technical efficiency” making use of students’ outcomes as outputs for a given level of inputs. Separate production frontiers are derived based on two levels of environmental harshness. The environmental factors are also treated as inputs in the formulation.

Chalos and Cherian (1995) presented a study of accountability in public schooling utilizing an input decreasing orientation of DEA. They also included socio-economic variables as uncontrollable inputs and controllable inputs (both controllable and uncontrollable? What differentiates the use of either category). Measures of students’ outcomes (standardized tests) where used as outputs. This is a study very similar to the one presented before by Ruggiero and Duncombe (1995).

Prieto and Zofio (2001) evaluated the effectiveness of public provision in Spanish municipalities. For these authors effectiveness is measure as the extent to which each DMU achieves its goals. To do that, effectiveness is assessed through a maximizing model relative to specific goals, i.e. standards or observed behavior. The authors defined their variables in per capita, per household or some relevant unit because the desired goals are stated in such form. In certain circumstances, goals and standards are
interchangeable terms. To evaluate individual service provision levels they introduced an additive model in which the sum of the slacks is minimized. Later they introduced standards for each output/outcome variable as benchmarks against which to measure provision effectiveness. To do so they partitioned the observed matrix of output observations departing from basic DEA models where the reference frontier is constructed as linear combinations of just the observed DMU's values. The new matrix is such that:

1. Observations that may exceed the reference standard are truncated so they define desirable facets in the provision frontier.
2. The standard values that are not present in the observed matrix because they are not reached by any observation are now introduced in the reference matrix.
3. The observed provision variables not subject to a standard still define the reference provision frontier.

With these, the authors formulated an additive model that has one equation for truncated values that exceeded the standard, values below the standard and all others.

Finally, Prieto and Zofio redefine an additive range adjusted measure of effectiveness, which corresponds to a "normalized" additive model. The great advantage of this model is that it not only provides an effectiveness score, but that through the break down of the DMU's score into partial and mutually exclusive elements, the partial projections to the frontier for each variable can be identified. The main disadvantage is that the model is computational intensive and very hard to understand and visualize.

Cooper, Seiford and Tone (2000) suggested that for including relevant outcomes in the provision of services, particularly public education, additional extensions could proceed to a "two-stage operation" in which outputs are transformed into outcomes, where the latter include things like success in securing employment after school has been completed.

Thanassoulis (1996) developed a DEA based method for identifying the extent and direction of differential effectiveness in schools using aggregate data at school level.
The formulation enables the determination of performance targets, which alter the bias and intensity of the differential effectiveness of a school. Environmental variables related to pupil performance are called contextual variables. Contextual variables, or non-controllable variables, are included in the model together with multi-outcomes of the education process such as employment, sporting, and musical achievements of pupils along with academic achievement. In Thanassoulis' model, DEA measures overall school effectiveness by setting pupil achievements as outputs, against the school's contextual variables such as pupil prior attainments and social backgrounds as inputs. The contribution of this model is mainly the inclusion of non-discretionary environmental variables that have been previously proved to influence outcomes as the only inputs and outcomes of pupil achievement aggregated at the school level as outputs.

2.5.2.5. **DEA Approaches to Measure More than Two Performance Dimensions**

More recently, some researchers in the nonprofit field have proposed an approach to evaluate service providers in more than one performance dimension (Nyhan & Martin, 1999). According to several authors (DeLancer, 1996; Lewin & Morey, 1986; Lovell et al., 1994; Thanassoulis, Boussofiane & Dyson, 1995) DEA can make efficiency assessments, quality assessments, effectiveness assessments, or any combination thereof. The formulation proposed is very similar to that discussed before for effectiveness and quality considerations alone, in that it considers *outputs* as a broader concept. According to Nyhan & Martin (1999), the term output in DEA can be interpreted to include not only output performance measures but also quality performance measures. Likewise, the term *efficiency* can be broadly interpreted to include not only an assessment of efficiency but also an assessment of quality and effectiveness (outcome).

In this sense, DEA can make simultaneous comparisons of multiple dependent performance measures (output, quality and outcome) and can provide a scalar measure of best overall practice.

Nyhan & Martin (1999?) present an example comparing different government agencies providing similar services. They evaluate cost ratios, namely cost
per outcome and cost per quality output but they state that exactly the same procedure can be performed for outputs interpreted only as outcomes and quality output without evaluating the relationship to cost.

2.5.3 Advantages & Disadvantages of the DEA approach

The following is a list of benefits of using the DEA approach to measure comparative performance of nonprofit or government providers of social services over other more traditional tools and techniques (e.g. ratio analysis, regression analysis):

- DEA provides considerable flexibility in data selection. Inputs and outputs can be continuous, ordinal or categorical variables, or a mix of different types (Banker & Morey, 1986).

- The outputs and inputs can also be measured in different units of analysis (e.g. dollars, FTEs, individual test scores, completion and hiring rates, etc.) (Nyhan & Martin, 1999).

- DEA is able to deal easily with large number of providers (DMUs).

- DEA can calculate the amount of resources that could be saved or, conversely, the amount of additional output that could be produced for any provider (DMU?) found to be inefficient (Nyhan & Martin, 1999).

- DEA generates objective data-based weights for inputs and outputs instead of relying on a priori weights assigned by evaluators and policy makers.

The following is a list of disadvantages or problems that performance analysis using DEA can face:

- DEA is most useful when large numbers of providers (DMUs) are involved. As the number of providers (DMUs) is smaller, more units tend to appear relatively efficient (Nyhan & Martin, 1999). Most authors suggest the need for between 4 and 15 DMUs for each input and output variable (Banker, 1993).
Because DEA is a nonparametric technique, it has no statistical indicators to measure error (noise) as other techniques do—such as regression. So, researchers using DEA must be well grounded in their data by utilizing a priori statistical techniques. However, there has been the work of Simar that studies the statistical properties of the DEA scores.

DEA does not measure the relative differences of efficient DMUs. Thus, it does not provide information to make comparisons based on a theoretical optimal performance standard. One can say that this is where the notions of super-efficiency and dominance attempt to provide a solution to.

Sound selection of input and output variables is fundamental. High face validity in the selection of the input and performance variables should increase the understanding and acceptance of DEA among evaluators and policy makers (Nyhan & Martin, 1999). Face validity is the type of validity attempting to assess the degree to which you accurately translated your construct into the operationalization. For DEA means how accurately the variables selected seem to be a valid representation of the DMU’s operations—whether "on its face" it seems like a good translation of the construct “production”.

DEA does not handle well large numbers of input and output variables. As more input or performance (output) variables are included in the analysis, the proportion of efficient, or best-practice providers tend to increase thereby decreasing the explanatory value of the solution (Chalos & Cherian, 1995).

### 2.5.4 DEA and the Nonprofit Sector

It is widely recognized that for nonprofit organizations the evaluation and control of operating activities is problematic. As these entities do not operate in
competitive output markets, measures such as net income and rates of return do not provide useful indicators of operating performance (Nunamaker, 1985). However, DEA facilitates a valid comparison of service provider performance in the nonprofit sector by modeling the provision of human services as a production process and providing a single summary measure of the relative efficiency or performance of each unit (Byrnes et al., 1997; Nunamaker, 1985). Thus, DEA has attracted the attention of researchers struggling with problems of productivity measurement, especially in the services and non-market sectors of economy (e.g. Bowlin, 1984; Lovell, 2000; Nyhan & Martin, 1999; Ozcan & Cotter, 1994; Ruggiero, 1996, 1998, 2000; Sexton, 1986). The main strength of DEA in this field is that it offers a method by which non-commensurate multiple inputs and outputs of an entity can be objectively combined into an overall measure of organizational performance (Greenberg & Nunamaker, 1987). The production function of nonprofit service providers is generally of that nature, multiple inputs and outputs of different dimensions and types. This methodology has widely been used to measure production (in)efficiencies of DMUs but it can be applied to the measurement of other performance indexes more relevant to the provision of services such as service quality and effectiveness. Indeed, DEA has been used to make provider comparisons in schools (Callen, 1991; Chalos & Cherian, 1995; Kao, 1994; Lovell, Walters & Wood, 1994, Ruggiero and Duncombe, 1995, Ruggiero 1996, 1998, 2000), human service agencies (Blank and Valdmanis, 2001; Byrnes et al., 1997; Ozcan & Cotter, 1994), court systems (Lewin, Morey & Cook, 1982), and health care providers (Cappetini, Dittman & Morey, 1985; Huang & McLaughlin, 1989; Juras, Kaspin & Martin, 1994; Nunamaker, 1983, Sherman, 1984, 1986), to name only a few.

An important factor to be considered is that DEA can be used to compare similar service providers or DMUs. The term similar means that all units must: (a) be involved in delivering services of some type; (b) use the same type of inputs; and (c) measure their performance using the same efficiency (output), quality and effectiveness (outcome) performance measures (Nyhan & Martin, 1999).

DEA not only leads to an identification of the most and least efficient units but also to potential targets for the unit to achieve in the future —measures of the
conservation of resources and/or augmentation of outputs levels that can be recommended to a unit to become efficient (Boussofiane et al., 1991).

According to Boussofiane et al. (1991), a DEA assessment yields other information, which proves useful in gaining a better insight into the performance of each unit and in guiding units to improve their performance. The following types of assessments of the individual units are obtained from DEA:

- A set of peers or reference group composed by units to which the organization under analysis can best be compared against.
- Identification of efficient operating practices.
- It sets performance targets (i.e. the levels of inputs and outputs that will improve performance).
- It identifies efficient strategies or production policies that efficient units follow, including insight into resource allocation strategies.
- It can monitor efficiency changes over time.

Also, standard linear programming shadow prices in DEA are used for different purposes. Nonzero shadow prices appear as output values only for relative inefficient DMUs. When they appear, they indicate which of the other DMUs form the reference set for that inefficient provider. The reference set is the number of other DMUs against which an individual DMU is found to be inefficient, or in other words, the set against which a DMU can be most meaningfully been compared (Ganley & Cubbin, 1992; Nyhan & Martin, 1999). In the nonprofit sector the so called reference set is important because qualitative information about the providers in the reference set may assist policymakers and policy evaluators to better understand the reasons for the relative inefficiency of a particular DMU (provider) (Callen, 1991). It can also be used to compute the amount of resources that could be saved or the additional amount of output, quality, or outcome that could be produced if an inefficient provider became efficient relative to the other providers in the reference set.
As Ruggiero (2000) points out, technical returns to scale represents the technical relationship between inputs and outputs holding environmental factors constant. Regarding the selection of appropriate returns to scale, Nyhan & Martin (1999) state that most public sector situations conform to the constant rates of return model (CCR model). But other models exist in case one would like to make no a priori assumptions (e.g. BCC model). In reality, when other types of variables are included, such as categorical, ordinal, of the qualitative or outcome type, the BCC model seems to be more adequate. In reality, it is recognized that constant returns to scale (CRS) is a special case of variables returns to scale (VRS) so why would one not assume VRS unless one can show that CRS is appropriate for the technology.

Furthermore, when non-discretionary environmental variables are present (and in most nonprofit and public production cases they are), the notion of returns to scale changes. Duncombe and Yinger (1993) provided a systematic analysis of returns to scale in the public sector assuming cost minimization. They identify three dimensions of scale: service quality, the level of governmental activity and population. At least two of these can be extrapolated to the nonprofit arena: service quality and population. Later, Ruggiero (2000) presented a method to estimate returns to environmental scale in the public sector, holding discretionary inputs constant without cost minimization assumptions. This author derives this method from his previous works (1998, 1996) in which develops a DEA formulation for public production that considers the level of environmental harshness. Ruggiero suggests that the environment may not only affect efficiency measurement but possibly the slope and consequently, the scale elasticity along the frontier for a given level of outcomes.

The following is the list of assumptions regarding the provision of social services in a nonprofit organization:

Assumptions:

- Variable returns to scale is assumed.
• Existence of multiple frontiers due to exogenous factors are prevalent. Since nonprofit production is greatly impacted by the DMU’s operating environment, different levels of operating environments generate different service delivery (production) conditions. Hence, DMU’s should be compared only to those DMU’s under the same or worse operating conditions. Thus, a fair analysis of nonprofits would use a DEA formulation that generates different frontiers from the corresponding DMUs at the same level for each DMU in the sample.

• Weak disposability of outputs/outcomes; outputs is assumed. Proportional increase in outputs or outcomes cannot be obtained utilizing less resources (inputs)

• Strong disposability of inputs is assumed. Outputs or services cannot decrease if inputs (resources) are increasing.

• There is no “free-lunch” – services and outcomes cannot be obtained without using some inputs.

• Infinite outputs and outcomes are not possible.

• The correspondence of inputs and output is closed

• The output-increasing measure of technical efficiency is adequate whenever the objective is not to minimize resources but rather to maximize outputs and outcomes.

From the review of literature in this chapter several theoretical implications regarding nonprofit operations are withdrawn. The inclusion of capacity creation as a fundamental piece of a nonprofit system is one of them. The approach to measure service quality as an index of perceived quality is also supported in the literature as well the notion of outcome achievement as the proportion of customers who experienced a change or benefit after receiving the service. Moreover, it is clear that the effect of the environment both in outputs and outcomes needs to be included in the
model, so that units operating under the same or worse environment are compared. This needs to be done in such a way as to avoid comparing units operating under favorable environments to be part of the reference set of units in harsh environments. All these implications for the present research will be discussed in Chapter 3.
CHAPTER 3 Model Development and DEA Formulation

In Chapter 2 existing theoretical models that measure performance in services—particularly services provided by nonprofit social service organizations were reviewed. The multi-dimensional character of performance was discussed in light of both traditional and most recent measurement models, each of them advocating a number of key performance areas that ought to be addressed to obtain a valid and overall picture of performance. In Chapter 2, I also reviewed the available formulations to handle the problem of incorporating service quality attributes and effectiveness indicators (variables) as well as external environmental factors out of the control of the firm into Data Envelopment Analysis, DEA. This review gave light to a conceptual model to measure performance in social services, which will be discussed hereinafter.

In Chapter 3, I present the development of the mathematical formulations suitable to handle the emerging conceptual performance measurement model for social service organizations. According to the United Nations’ International Atomic Energy Agency (IAEA) in their publications regarding technical standards (2002), a conceptual model is a set of qualitative assumptions used to describe a system (or part thereof). These assumptions would normally cover the dimensionality of the system, initial and boundary conditions, time dependence, and the nature of the relevant processes and phenomena. The mathematical formulations or mathematical model are a set of mathematical equations designed to represent a conceptual model. Both types of models are developed in this chapter.

Hence, this chapter is divided in two major sections. In section 3.1 the conceptual model that measures the performance of social services is developed. The aim of this section is to provide an answer to Research Questions R1 and R2, namely:
R1: How can Social service organizations measure and evaluate the overall performance of their services and programs?

R2: Which dimensions and respective measures of performance are the most appropriate to represent the overall “health” of social service organizations?

To answer those questions essential key performance indicators for social services (Brown, 1996; Clark, 1995) were identified. This was accomplished by i) performing an exhaustive review of proposed performance dimensions for social services in the literature; ii) identifying current measurement initiatives; iii) reviewing the dimensions being measured, and iv) confirming the relevance of the metrics selected for each one of the performance dimensions under study. The relevance was checked by social service organization directors. Steps i, ii, and iii were mostly reviewed in Chapter 2. The conclusions obtained from this process are herein provided.

In Section 3.2 two different DEA formulations are developed to compute the metrics identified by the model in Section 3.1. This section answers research questions R4 and R5:

R4: How can one integrate the relevant performance dimensions and their respective measures in an overall measure of performance?

R5: What modifications to the existing DEA models will compute the performance measures and generate relevant “performance frontiers”?

To address these questions the relationships among the performance dimensions and measures were investigated and the formulation of the modified DEA model was achieved by:

a. The mathematical definitions of efficiency, effectiveness and service quality were provided. Similar mathematical definitions reported in the performance measurement literature were evaluated. These concepts are defined in Section 2.1.

b. The nature of the relationships between efficiency, effectiveness and service quality were evaluated by carrying out a systemic analysis of performance in
social services. A four-stage modeling approach stem from this process that measures overall performance in these organizations.

c. Appropriate operational models that account for an agency’s social service performance were formulated. These models account for performance at each stage as previously identified.

d. The DEA formulations that result in estimating overall agency performance considering technical efficiency, effectiveness and service quality follow the guidelines provided by Metzger (1994, p. 73). Other DEA modeling issues raised for DEA formulations were considered as well (e.g. Banker & Morey, 1986; Burgess & Wilson, 1996; Girod, 1996; Lothgren & Tambour, 1996; Ruggiero, 1998, 2000). These include:

♦ The unit of analysis for DEA is the DMU, which can be the organization, one department or one project as long as there is consistency in the data. The researcher can also choose between analyzing one production unit over several time periods or several similar production units over one time period.

♦ One needs to analyze the production system. This can be done by considering the internal mechanisms of the production “black box” where inputs are introduced and transformed into outputs. The analysis of nonprofit production required the amplification of this concept so as to introduce outcomes and service quality into the analysis. (Section 3.1)

♦ The relationship between the output and the related inputs should make economic sense. Output is usually good or conforming output.

♦ The DEA assumptions must hold. Verification of the DEA assumptions described in Chapter 2.0 (Section 2.5.1.2) was also performed using analytical and quantitative methods. In the analytical side, the conceptual model was treated as a black box and the relationship between inputs and outputs was established and verified with non-profit managers. Outputs
selected were demonstrated to be the product of inputs. This relationship was closed and nor free neither infinite production is feasible. Also, statistical and quantitative tests were applied to prove the positive correlation between inputs and outputs. Convexity was assumed.

Having other similar modifications to the original DEA formulation as guidelines, an adaptation necessary for this research was deemed viable. Based on the operationalized definitions of the performance dimensions, the DEA objective function and constrains were formulated.

3.1 Development of a Conceptual Model to Measure Performance in Social Services

Several authors (e.g. Kaplan, 2001; Letts et al., 1999; Moore, 1995; Sola and Prior, 2001) have described the unique characteristics of nonprofit organizations in the delivery of social services. Compared to for-profit operations, the main objective of social services is not to make profit but to have an impact on society (the so called outcome). To be able to do so, these organizations need to be active fundraising agents either in the form of donations, income generation activities or grants, and neet to make a wise use of resources. Furthermore, nonprofit organizations dedicated to the delivery of services need to use this income to develop organizational capacity. Organizational capacity enables them to properly operate and deliver services to accomplish their mission. This capacity can be seen as of having different forms: service delivery or programmatic capacity, growth capacity, adaptive capacity (Letts et. al., 1999). I use the notion of organizational (or service delivery and programmatic) capacity as defined in section 2.2.4. Organizational capacity encompasses the capacities for program delivery. Staff is recruited according to programmatic skills and experience. This type of capacity provides the organization the ability to accomplish relevant and sought after outcomes. Likewise, capacity for program expansion, or the "capacity to grow" requires more resources and capabilities to respond to issues the organization once handled more informally. Metrics to analyze organizational performance under the umbrella of growth are related to organization, management and program expertise to respond to future
challenges or added needs in the community. This model does not evaluate this aspect of capacity and these variables are not considered.

Finally, the delivery of such services to the community needs to have an impact or generate a positive change. This is the ultimate goal of social service organizations. The extent to which they do so represents their level of effectiveness.

Consequently, there are at least four processes occurring inside nonprofit organizations with a social charter: (i) Nonprofit organizations do fundraising or generate income either by approaching local, state or Federal government agencies for grants and contracts or by doing fundraising in the community they serve to obtain the resources necessary to perform their mission. (ii) They create organizational capacity through the acquisition of human and physical resources, training and research. (iii) They utilize the created capabilities to deliver their services, and (iv) they deliver their services and as a consequence they generate an impact or change in the service recipients. While not all these concepts are new, no author has attempted to combine them in an Input-Output framework. The sections that follow attempt to do just this. These four processes form the foundation of a performance evaluation model for nonprofit agencies that incorporates outcomes, service quality and technical efficiency as valid measures of performance. The resulting conceptual model will facilitate the DEA model formulation of Section 3.2.

3.1.1 The Financial Health Node

Financial objectives of non-profits differ considerably from their for-profit counterparts. Recent research regarding appropriate performance measures for nonprofit organizations (Kaplan, 2001; Letts, 1999) suggests that the main financial objectives of the sector are to increase the amount of income (donations, grants, contracts) so that their funds or revenues increase while their fundraising and managerial expenses decrease or are kept constant. It is well known, however, that several environmental and non-controllable factors affect the amount of income organizations are able to get as well as the level of fundraising expenses such as, the wealth of the community, the average age of the population, race and cultural aspects, etc. Regardless of these influences, social
service organizations utilize resources to do fundraising (inputs) to produce or generate revenue (donations, grants, contracts) on the output side. The following Deming diagram in Figure 3-1 depicts this Financial Health node.

Figure 3-1. Deming Diagram for the Financial Node

3.1.2 The Capacity Creation Node

The concept of capacity creation has been explored in the literature as fundamental for service delivery organizations in the nonprofit world (Letts, 1999; Moore, 1995, Sola and Prior, 2001). Nonprofit organizations utilize the donations they receive to create organizational capacity. They do this by hiring specialized and support staff, acquiring equipment, expending on training, etc. The capacity to grow and the capacity to deliver are concepts being explored in the literature as fundamental to evaluate nonprofit performance.

Inputs for the capacity creation node are the resources acquired through fund raising. The outputs of this node are the so-called "Capabilities" of the organization
such as, employees and the degree of training of those employees, the development of programs and services, equipment purchased etc. Figure 3-2 depicts the capacity creation node.

![Deming Diagram for the Capacity Creation Node](image)

Figure 3-2. Deming Diagram for the Capacity Creation Node

### 3.1.3 The Service Delivery Node

The center of the production function for social service organizations is the service delivery node. These organizations utilize their capabilities to create value in the form of programs and services to the community. These services in turn can be defined by their characteristics (timeliness, empathy, convenience, etc.) that are service quality attributes. In the literature, service quality characteristics have been considered as outputs of the service delivery process (e.g. Lothgren and Tambour, 1996; Soteriou & Zenios, 1999; Thanassoulis et al., 1995). On one hand, organizations can produce a specific amount of services, on the other, these services can be with high quality, mediocre quality or poor quality according to a series of criteria discussed in Chapter 2. Figure 3-3 depicts this node.
The non-controllable characteristics of the population served shape the way services are delivered as well as client's perceptions of the service and are included in the model.

### 3.1.4 The Customer Effectiveness Node

The customer node represents the ultimate objective of social service organizations (i.e. to produce a change in the people they serve). Clients receive services with their perceived quality and in turn experience a degree of satisfaction with the service plus a change in knowledge, attitudes, behavior, quality of life, etc. In short, the outputs of the customer node are both the outcomes achieved, and the level of customer satisfaction. As in previous nodes, client's characteristics that are of non-discretionary nature have great influence on the results of this node. Figure 3-4 depicts this node.
Finally, all four nodes can be combined to represent service delivery technology for nonprofit organizations. Figure 3-5 depicts the stage-by-stage relationships of the four nodes.

Figure 3-5. Conceptual Model of the Production Technology of Nonprofit Organizations.
The variables depicted for each node are either classified as controllable (the achievement is discretionary to the organization's management), and non-controllable or non-discretionary, (out of the control of the organization). The variables are classified as follows:

*Fundraising (Income generating) and Development Resources*

\[X^{fr} = \text{Controllable fundraising and managerial expenses}\]

*Income from the Community*

\[Z^{e} = \text{Non-controllable environmental characteristics of the community/population served (such as wealth of the community)}\]

\[X^{d} = \text{Controllable resources donated/provided to the organization by the community/government.}\]

*Capacity Created*

\[X^{p} = \text{Controllable resources created/acquired by the organization to deliver the services}\]

\[X^{c} = \text{Qualitative characteristics of the resources created/acquired by the organization to deliver the services.}\]

*Outputs of the Service Delivery Process*

\[Y^{p} = \text{Services delivered.}\]

\[Y^{q} = \text{Qualitative characteristics of those services.}\]

*Outcomes*

\[Y^{o} = \text{Outcomes or the degree of achievement of the services objectives.}\]

*Environment*

\[Z^{e} = \text{Single factor of environmental harshness that affects the operating environment of the DMUs grouping them according to this level.}\]

\[X^{e} = \text{Environmental factors (inputs) that affect nonprofit production and that are out of the control of the organization}\]
3.2 Development of a DEA Formulation to Measure Performance in Non-profit Organizations

The conceptual model developed in Section 3.1 depicts different nodes of production, which are sequentially related. This clearly calls for the evaluation of performance in a stage-by-stage approach. Several authors have applied stage-by-stage concepts in DEA (e.g. Ruggiero, 1996; Soteriou and Zenios, 1999; Athassanopoulos, 1999; Blank & Valdmanis, 2001).

Athanassopoulos (1999) presented a two-stage approach for evaluating bank branches performance in light of a service quality frontier and a technical efficiency (or efficiency-X frontier). Soteriou and Zenios (1999) expanded on Athanasopoulos’ work to include operations, quality and profitability as indicators of strategic benchmarking and efficiency benchmarking. They did so by developing three models: (i) an operational efficiency model, (ii) a service quality efficiency model, and (iii) a profitability efficiency model. All authors, (Athanassopoulos, Soteriou and Zenios) approached the development of their two-stage DEA formulations in the form of a "cascade" of models. The cascade of models, according to them, can be used in different ways. One approach is to run each model separately and then correlate the results. The second approach used by these authors utilizes the projected virtual outputs of one model as the input to the second model. With this approach they determine the peak profitability (since the profitability frontier is the second stage) of the bank branches under study after they have been "re-engineered" to deliver superior quality. I have used a stage-by-stage approach in the same fashion as Athanassopoulos, Soteriou and Zenios's models. I have proposed two different formulations: one that carries the projected values or performance targets that DEA achieves for each node as inputs to the following node. The other formulation uses the real values of the variables at each stage, without utilizing projections or targets at all. Also, as stated before, it is very important to consider influential environmental characteristics at each stage of the proposed model. The next section deals with this inclusion.
3.2.1 Incorporation of the Environment in Nonprofit Production: Proposed New Formulation for all Models

As stated in sections 2.2.5.1, public and nonprofit production technology is greatly impacted by the environment and non-controllable characteristics of the customers or clients. Thus, any DEA formulation for a social service agency ought to account for the impact of a harsher environment on the DMU's production. As stated before in Chapter 2, Ruggiero (1996, 1998) proved that the most commonly used DEA formulation to deal with non-controllable external environmental factors—presented by Banker and Morey in 1986—would underestimate the level of technical efficiency. This is true if, for the composite reference group determined by the DEA formulation \( (X^*, Y^*) \notin T(Z_o) \) where \( X^* \) and \( Y^* \) are the optimal level for inputs and outputs of the DMU being evaluated, \( T \) is the production possibility set of DMU being evaluated, and \( Z_o \) the level of a factor representing the external environment for the DMU being evaluated. This happens because in the Banker and Morey model, the restrictions are placed on the level of \( Z_o \) for the composite reference group but not for the individual DMUs in the reference set. Thus, an optimal solution for DMU may include in its reference set DMU where \( Z_j > Z_o \). Consequently, without further restricting the reference set, a solution \( (X^*_o, Y^*_o) \) may not be feasible in a practical sense, since it would be recommending a target for X and Y that was calculated using DMUs operating in a different, more favorable environment.

For example, let's say a field office operates in a very wealthy community subsequently being able to raise a considerable amount of money. Another field office operates in a very disadvantaged community, raising far fewer dollars. A third field office is in an intermediate situation. Therefore, when using Banker and Morey's formulation, the first and last field offices might be deemed efficient and therefore, might be part of the reference set for the second one. However, the targets offered by this DEA solution for the second field office in a disadvantaged environment might not be feasible at all, given the different environment at which these units operate.

To account for the environmental impact on production, i.e. DMUs with a more favorable environment are able to produce more with less inputs, Ruggiero (1996) extends the DEA production postulates explained in section 2.1.5.1, to include the effect
of the environment. This author introduces an additional postulate in order to assert the relationship between the environment and the production frontier.

\[ T(Z_1) \subseteq T(Z_2) \text{ for all } Z_1 \leq Z_2 \]

This postulate states that the composite reference set for each DMU must be comprised of only those DMUs in the sample, which present an equal or harsher environment. \( Z_1 \) is a harsher environment than \( Z_2 \). For that, the variable \( Z_j \) that represents some influential environmental factor should be calculated in a way that a lesser value represents a harsher environment.

This is done to acknowledge that the optimal solution might not be always feasible if the peer groups used are not operating at the same environment. Thus, Ruggiero modifies the DEA model deleting from the formulation the equality regarding environmental variables as non-controllable inputs that was included by Banker and Morey (1986). Instead he includes a modification in the algorithm that creates different frontiers for each DMU under analysis. This modification allows the inclusion in the benchmark set of only those DMUs with the same or lesser level of environmental harshness.

This multiple-frontier approach calculates different frontiers for each DMU. For each DMU under consideration, the algorithm considers the DMUs in the sample at the same or lesser level of environmental harshness only. This approach works well whenever two characteristics of the environmental variables are met:

- First, the researcher must identify and define one single determining variable that best accounts for the impact of the environment on production. One can do this making use of statistical techniques (correlation and regression) or also using the decision-maker’s knowledge as to what factors impact production the most. This variable can be either continuous or categorical. Even though Ruggiero's formulation is only for continuous variables, over the course of the present research it has been shown that the algorithm works well for categorical variables as well (see chapter 5.

154
Second, it is possible to distinguish or separate DMUs in several subgroups according to different ranges of the environmental variable. The level of this variable must be in a form that is inversely proportional to the concept of "favorable production environment." Hence, the variable represents the level of "environmental harshness" and the lesser the level, the less conducive to production the environment it is e.g., market size, level of education, aggregated household income, etc.

To solve the DEA linear program the algorithm will select only DMUs operating at the same or lesser level of harshness than the DMU under analysis.

Thus, whenever more than one independent environmental variable impacts production—which is a common phenomenon in social service production—and whenever the differences among the level of the environmental variables are subtle, the Ruggiero's algorithm cannot be meaningfully applied.

The Banker and Morey formulation, on the other hand, assures that the weighted average of the reference set is equal to the value of the environmental variable for the DMU under evaluation. This formulation does not assure that DMUs in the composite set are selected according to the level of the environmental variable of the DMU under evaluation. As Ruggiero pointed out, the DMUs in the peer group can come from a more favorable environment, making the comparisons invalid.

So, as part of this research, a mixed model that includes both Ruggiero's and Banker and Morey's formulations is being proposed to deal with environmental non-controllable variables. In those cases, when two or more variables are important, but it is possible to identify one that would partition the DMU set according to the Ruggiero's model, one could think that the Banker and Model formulation could be used in parallel to include those other important environmental variables that not necessarily bind a viable reference set to DMUs with exactly the same or lesser level. The purpose is to include the effect of the second environmental variable in the selection of the peers.
This second environmental variable, X*, need not to be of the type that divides the sample of DMUs into categories or groups to limit comparisons. It also does not need to be an ordinal variable varying from lower to higher level of environmental harshness. This second environmental variable can be of any type, such as variables whose impact on production is known but that do not present a clear grouping of the DMUs into categories. It is not necessary to know the extent of its impact on production, neither the strength of this relationship. It can be that the effect is not as clear, but nevertheless the researcher deems the variable as useful to be included in the analysis. The purpose is ensure that the composite reference set will be computed considering this variable as well but its insertion does not limit the comparison of all DMUs against all DMUs in the sample. It is important to note that this variable does not need not to be correlated to the Ruggiero variable. A good example of a variable that could be included in the model as non-controllable environmental input could be level of educational attainment in the community. While organizations can consider educational attainment as an important factor in service delivery, —the higher the educational attainment the easier service delivery is expected—a lower educational attainment in the community is not a definite factor that can classify the environment as “harder”. So, it is nice to include it but it doesn’t mean that the analysis should be restricted to DMUs under the same or worst level of educational attainment only.

In the sections that follow two DEA formulations that interpret two possible approaches to the nonprofit production problem have been developed. Both can be considered stage-by-stage models and have been developed following Athanasopoulos (1999) and Soteriou and Zenios (1999) approach to incorporate DEA formulations for each one of the production nodes described in section 3.1. This new formulation does consider the impact of the environment through a mixed model of BM and Ruggiero when applicable.

The first one considers the real level of each variable at each stage. This formulation uses the outputs of one node as the inputs of the following. The second approach carries the optimal projection to the frontier of the variables in the first stage to the second, the second to the third, and so on, thus mimicking Athanasopoulos (1999)
and Soteriou and Zenios (1999) approaches. I present these models in the sections that follow.

### 3.2.2 A Four-Stage Approach for Non-profit Performance Evaluation: Using the Observed Values of the Input/Output Variables

The first model developed will carry the outputs from the first stage to the second stage, from the second to the third, and from the third to the fourth, as they are observed. This formulation calculates four different frontiers and then in order to give a final score, for ranking purposes, calculates an index number that is the product of the inverse of the efficiency scores of all stages.

The formulations at each stage are output increasing, as nonprofit organizations are more concerned with performance of units producing more outputs for the same level of inputs. The output maximization model sets goals for each stage that can be achieved given the resources utilized by the DMU (Soteriou and Zenios, 1999). All DEA formulations in this chapter assume variable returns to scale to account for non-linear relationships between inputs and outputs, if they exist (Soteriou and Zenios, 1999) and belong to the family of non-radial efficiency models. Non-radial models represent a more realistic approach to capture the behavior of certain performance variables. In the case of service quality, for example, in the real world the percent projection to the efficient frontier should be different than percent projection of the number of services delivered. Radial models assume that all variables are projected at the same angle and proportion. Thus, possible recommendations for improvements are more realistic if the model allows different variables to be projected in a different fashion. So, the optimal solution for these models yields improvement rates for individual inputs and outputs according to their representation in the objective function (i.e. according to their individual projection rates). Targets will be estimated for all service-quality dimensions given different rates of projection to the frontier as in Athanassopoulos (1996 and 1999). Environmental variables are included in two fashions, as in Ruggiero (1998) and Banker and Morey (1986).
3.2.2.1. Stage I. The Financial Frontier -DEA Model for the Financial Node

Let us consider a set of $j=1\ldots n$ nonprofit organizations that will be referred to as DMUs. In the first stage, this set of DMUs uses a set of discretionary inputs, i.e., controllable fund raising, managerial and/or outreach expenses, $x_{ij}^{fr}$ that represent the input $i$ of DMU $j$ in $F \in \mathbb{R}_+$. $F$ is the set of all fundraising/managerial and outreach cost variables and $f$ is the number of such variables. DMU is impacted by continuous non-discretionary environmental factors $z_{mj}$, where the environmental input $m$ of DMU $j$ in $E \in \mathbb{R}_+$. $E$ is the environmental factors set and $a$ is the number of such variables. These affect the production of output quantities which are the monetary and in-kind donations $x_{kj}^d$ the donation $k$ of DMU $j$ in $D \in \mathbb{R}_+$ where $D$ is set of donations and in-kind contributions and $b$ represents number of such variables.

Let us also consider the set of $w_k$ of user specified constants — decision-maker's weights— reflecting the decision maker's preferences over the types of donations received. The decision-maker based on experience and the requirements sets the values of the weights $w$ in the objective function. The values of these represent the relative importance of one type of donation with respect to the other types of donations. The weights are obtained from the decision-maker, mostly asking him/her about the relative importance that each one of the variables in the objective function has.
Stage 1:

Max $\sum w_k \phi_{k_i}$  \hspace{1cm} (3.1)

s.t.

$\sum_{j=1}^{n} \eta_j x_{j(t-1)}^{fr} \leq x_{j(t)}^{fr} \hspace{1cm} i \in F \hspace{1cm} (3.2)$  fundraising, managerial & outreach costs

$\sum_{j=1}^{n} \eta_j x_{j(t-1)}^{ce} = x_{j(t-1)}^{ce} \hspace{1cm} m \in E \hspace{1cm} (3.3)$  Non – controllable environmental factors

$\sum_{j=1}^{n} \eta_j x_{j(t-1)}^{do} \geq \phi_{j(t)} x_{j(t-1)}^{do} \hspace{1cm} k \in D \hspace{1cm} (3.4)$  Donations & in – kind contributions

$\sum_{j=1}^{n} \eta_j = 1 \hspace{1cm} (3.5)$

$\eta_j \geq 0 \forall j \hspace{1cm} z_{j(t-1)}^{e} \leq z_{j(t-1)}^{e} \hspace{1cm} (3.6a)$  Non – controllable environmental factor

$\eta_j = 0 \forall j \hspace{1cm} z_{j(t-1)}^{e} > z_{j(t-1)}^{e} \hspace{1cm} (3.6b)$

Equation 3.2 treats expenses incurred for fundraising events, income generating activities managerial expenses incurred to prepare grant proposals and get contracts as well as campaigns as discretionary inputs. I have incorporated a variation to the normal treatment of inputs in DEA since this formulation utilizes the expenses incurred at time t-1. This is to recognize that the money collected during year t-1 usually pays for the services provided in year t.

Equation 3.3 accounts for non-controllable environmental factors that affect the organization in its ability to generate income or do fundraising following Banker and Morey (1986). Factors such as the average age of community members, their race, etc. are factors influencing fundraising. Equation 3.3 is an equality to account for the non-discretionary nature of such factors, meaning that the weighted sum of the environmental factors under analysis for all DMUs in the sample, need to be equal to the one of the DMU in question.. Equation 3.4 treats donations and in kind contributions in time (t-1) as discretionary outputs. Since this is a non-radial model, it allows each type of donation to be projected to the frontier not necessarily in the same proportion as the other donations. This formulation will compute a score $\phi_k$ for each type of donation/contribution obtained by the DMU being evaluated.
Equation 3.5 accounts for variable returns to scale (Banker et al., 1984). Equations 3.6a and 3.6b are based in the Ruggiero's model to deal with environmental variables that allows for the partition of the DMUs set into groups of the same or worse level of environmental harshness. Only those DMUs with the same or harsher level of environment than the DMU being evaluated are considered in the analysis. These equations account for the one —continuous or categorical— environmental variable that affects the fundraising or income generating activities (e.g. the wealth of the community where the fund raising is performed) the most. Equation 3.6a assures that the peer group set is chosen from all those DMUs operating at the same or lower level of environmental harshness. Equation 3.6b assures that if a DMU in the sample operates at a more favorable environment, it will not be included in the DEA computations.

The objective function in 3.1 maximizes the value of the sum of $\varphi_{kjo}$ and incorporates the concept of weights or preferences given by the DMU’s decision-makers. These incorporate their notion of which types of donations or outputs are more important in estimating the performance targets. In that $w_k$ are the weights for output k given by the decision-maker.

### 3.2.2.2. Stage II: The Capabilities Frontier - DEA Model for the Capacity Creation Node

For the second stage (the capacity creation node) the outputs of the first stage are used as resources to acquire capabilities that are the products of the second node. In this second stage, the set of DMUs uses a set of discretionary inputs, i.e., the controllable income (funds obtained from donations, grants or contracts) from the community, corporations or government, $x_{kj}^d$ which is the donation k of DMU j in $D \in \mathbb{R}_+^b$. DMU j is affected by non-discretionary environmental factors, $x_{mj}^e$, the environmental input m of DMU j in $E \in \mathbb{R}_+^a$ and $z_{ej}$ the level of environmental harshness in which DMU j is operating. The discretionary inputs and environmental variables yield the capabilities of the organization, i.e., the staff, technology, programs and processes which are represented by $x_{rj}^o$ the output or capability r of DMU j in $P \in \mathbb{R}_+^e$ where P is the set of capabilities and e is the number of such variables. The attributes or qualitative characteristics of those capabilities are represented by $x_{sj}^c$ where s is the capability.
attribute being considered. These variables represent the degree of training and/or other attributes of the human resources, and of other created capabilities in \( C \in \mathbb{R}^g \) where \( C \) is the set of all capability attributes and \( g \) represents the number of such variables. All variables are realized at time \( t \) except for \( x_{kj}^d \) which are realized at time \( t-1 \).

**Stage II:**

\[
\text{Max } \sum w_j \beta_j + w_s \gamma_s \quad (3.7)
\]

s.t.

\[
\sum_{j=1}^{n} \alpha_j x_{mj}^c = x_{mj}^e, \quad m \in E \quad (3.8) \quad \text{Non–controllable environmental variable}
\]

\[
\sum_{j=1}^{n} \alpha_j x_{kj}^d = x_{kj}^d, \quad k \in D \quad (3.9) \quad \text{Donations from the community}
\]

\[
\sum_{j=1}^{n} \alpha_j x_{rj}^b \geq \beta_r x_{rj}^c, \quad r \in P \quad (3.10) \quad \text{Controllable created capabilities}
\]

\[
\sum_{j=1}^{n} \alpha_j x_{sj}^e \geq \gamma_s x_{sj}^e, \quad s \in C \quad (3.11) \quad \text{Controllable attributes of capabilities}
\]

\[
\sum_{j=1}^{n} \alpha_j = 1 \quad (3.12)
\]

\[
\alpha_j \geq 0 \quad \forall j \ni z_j^c \leq z_{oj}^c, \quad (3.13a) \quad \text{Factor of Environmental Harshness}
\]

\[
\alpha_j = 0 \quad \forall j \ni z_j^c > z_{oj}^c, \quad (3.13b)
\]

Equation 3.8 accounts for non-controllable environmental factors that affect the organization's ability to create capacity — according to the Banker and Morey formulation. Factors such as cultural characteristics of the organization, of their personnel, the available technology, the size of the market, etc. are factors that affect capacity creation. As with Eq. 3.3, Eq. 3.8 is also an equality, meaning that the weighted sum of the environmental factor under analysis for all DMUs in the sample, need to be equal to the level of environmental factor of the DMU that is being evaluated.

Equation 3.9 treats monetary contributions and in-kind donations at time \( (t-1) \) as resources to create organizational capacity. Because it is not the objective of nonprofit organizations to reduce the amount of public contributions or donations, this
equation is in the form of equality. The equality prevents rewarding DMUs that use less donations but instead, forces them to be compared to a peer group composed by DMUs utilizing the same amount of these resources. According to Athanasopoulos (1999) the equality gives donations a weakly disposable character. This means that it is not possible to indefinitely reduce the inputs and still have monetary revenue (donations, grants or contracts).

Equation 3.10 treats the created capabilities at time= t, as discretionary (controllable) outputs. This is a non-radial model that allows each type of capability to be projected to the frontier not necessarily in the same proportion as the other capabilities. As such, this formulation will obtain a score \( \beta_r \) for each type of capability created by the DMU being evaluated. Likewise, equation 3.11 treats the attributes of the created capabilities at time= t, as discretionary outputs. This is also of non-radial nature and allows each type of attribute to be projected to the frontier in a non-radial way. As such, this formulation will obtain a score \( \gamma_s \) for each one of the s attributes created by the DMU being evaluated.

Equations 3.13a and 3.13b account for the environmental variable that the researcher wants to use to group the DMUs (the Ruggiero variable) and account for the one —continuous or categorical— environmental variable that affects capacity creation (e.g. level of professionalism found in the community, size of the operation, etc.) the most. Equation 3.13a assures that the peer group set is chosen from all those DMUs operating at the same or lower level of environmental harshness. Equation 3.13b assures that if a DMU in the sample operates at a more favorable environment, it will not be included in the calculations.

The objective function in 3.7 maximizes the value of the sum of \( \beta_r \) and \( \gamma_s \), weighted by the preference given by the decision-makers. So, \( w_r \) is the weight for capability \( r \) and \( w_s \) is the weight for attribute \( s \) given by the decision-maker.

Equation 3.12 accounts for variable returns to scale (Banker et al., 1984).
3.2.2.3.  **Stage III: The Service Delivery Frontier -DEA Model for the Services Node**  

For the third stage (the service delivery node) the outputs of the second stage are used as inputs of the service delivery process. In this second stage, the set of DMUs uses this set of discretionary inputs (controllable organizational capabilities): $x_{rj}^p$ the capability $r$ of DMUj that belongs to the set $P \in \mathcal{R}_{+}^e$ where $e$ represents the number of such variables, and attributes or qualitative characteristics of those capabilities, $x_{sj}^c$ that belongs to the set $C \in \mathcal{R}_{+}^g$ where $g$ represents the number of such variables. The non-discretionary environmental factor $x_{mj}^e$ is the environmental non-controllable input $m$ of DMUj and it belongs to the set $E \in \mathcal{R}_{+}^d$ and $z_j^e$ represents the level of harshness of the operating environment of DMUj. The DMU yields services or products, e.g., the number of services, the number of courses taught, the number of people trained, etc. These output services are represented by $y_{tj}^p$ which is the service $t$ of DMUj in $S \in \mathcal{R}_{+}^v$ where $S$ denotes the set of all the services and $v$ represents the number of such service variables. The attributes or service quality characteristics of the outputs measured as either perceived or operational quality, are denoted by $y_{qj}^a$ and belong to the set $Q \in \mathcal{R}_{+}^u$ where $Q$ represents the set of quality attributes and $u$ denotes the number of such variables.
Stage III:

\[
\text{Max } \sum (w_r \rho + w_q \pi_q) \quad (3.14)
\]

\[\text{s.t.} \]

\[
\sum_{j=1}^{n} \sigma_j x^e_{mj} = x^e_{mj}, \quad m \in E \quad (3.15) \quad \text{Non-controllable environmental variable}
\]

\[
\sum_{j=1}^{n} \sigma_j x^p_{rj} \leq x^p_{rj}, \quad r \in P \quad (3.16) \quad \text{controllable capabilities}
\]

\[
\sum_{j=1}^{n} \sigma_j x^c_{sj} = x^c_{sj}, \quad s \in C \quad (3.17) \quad \text{attributes of capabilities}
\]

\[
\sum_{j=1}^{n} \sigma_j y^o_{qj} \geq \rho_j y^o_{qj}, \quad t \in S \quad (3.18) \quad \text{Services delivered}
\]

\[
\sum_{j=1}^{n} \sigma_j y^q_{qj} \geq \pi_q y^q_{qj}, \quad q \in Q \quad (3.19) \quad \text{Attributes of Services delivered}
\]

\[
\sum_{j=1}^{n} \sigma_j = 1 \quad (3.20)
\]

\[
\sigma_j \geq 0 \quad \forall j \ni z^e_{j} \leq z^e_{j,0}, \quad (3.21) \quad \text{Factor of Environmental Harshness}
\]

\[
\sigma_j = 0 \quad \forall j \ni z^e_{j} > z^e_{j,0}, \quad (3.22)
\]

Equation 3.15 has the same function as before, accounts for the non-controllable environmental inputs that affect the organization ability to deliver high quality services in expected numbers, such as the percentage minority or at risk populations in the community. As in 3.8, it is also an equality. Equation 3.16 accounts for the capabilities as discretionary inputs at time = t. Equation 3.17 keeps the level of quality of the inputs constant (attributes of the capabilities) thus, assuming weak disposability of attributes.

Equation 3.18 accounts for each discretionary output (services) in a non-radial fashion. This allows each type of service to be projected to the frontier in ways other than radial, and not necessarily in the same proportion as the others. As such, this formulation will obtain a score \( \rho_k \) for each type of service provided by the DMU being evaluated. Likewise, equation 3.19 treats the service quality attributes as discretionary outputs. As such, this formulation will obtain a score \( \pi_q \) for each one of the q dimensions of quality for the DMU being evaluated.
Equations 3.21 and 3.22 are the equations that incorporate the Ruggiero's approach to deal with environmental non-controllable factors and account for the one — continuous or categorical— environmental variable $z^e_j$ that affects the most service delivery (e.g., population density of the area of influence of the organization, wealth, etc.). Equation 3.21 assures that the peer group set is chosen from all those DMUs operating at the same or lower level of environmental harshness. Equation 3.22 assures that if a DMU in the sample operates at a more favorable environment, it will not be included in the computation of the efficiency scores.

The objective function in 3.15 maximizes the value of the sum of $\rho_t$ and $\pi_q$ weighted by the preference given by the decision-makers. So, $w_t$ is the weight for service $t$ and $w_q$ is the weight for service quality attribute $q$ given by the decision-maker. Equation 3.20 accounts for variable returns to scale (Banker et al., 1984).

3.2.2.4. **Stage IV: The Effectiveness Frontier - DEA Model for the Customer Node**

In the last stage (the customer node) the service outputs from the third stage are used as inputs. In this last stage, the set of DMUs uses this set of discretionary inputs (services and service quality attributes), i.e., $y^o_{tj}$ the input $t$ of DMU $j$ that belongs to the set $S \in \mathbb{R}_+^v$ and the attributes or qualitative characteristics of those services, i.e., $y^q_{sj}$ that belong to the set $Q \in \mathbb{R}_+^u$. The non-discretionary environmental factors, i.e., $x^e_{mj}$ is the environmental input $m$ of DMU $j$ that belongs to the set $E \in \mathbb{R}_+^a$ and that in conjunction with the inputs yield outcomes experienced by the customer or client receiving the service (change in knowledge, attitudes, behaviors, quality of life, etc.). These outcomes are $y^o_{hj}$ which represents the outcome $h$ of DMU$j$ and belongs to the set $O \in \mathbb{R}_+^o$ where $O$ is the set of all outcomes and $o$ is the number of such variables. $z^e_j$ again is the level of environmental harshness of the operating environment (which is out of the control of the decision-maker).
Equation 3.24 follows the same logic as in previous nodes for the treatment of non-controllable environmental characteristics. Equation 3.25 accounts for the services as discretionary inputs to the process. Equation 3.26 keeps the level of service quality constant. Both equations are equalities which assume weak disposability of inputs and make no preconceived assumptions as to the effect of either services or service quality on the achievement of outcomes (Athassanopoulos, 1999).

Equation 3.27 accounts for each discretionary outcome in a non-radial fashion. As such, this formulation will obtain a score $\vartheta_h$ for each type of outcome created by the DMU being evaluated.

The objective function in 3.23 maximizes the value of the sum of $\vartheta_h$ weighted by preference given by the decision-makers. So, $w_h$ is the weight for outcome $h$ given by the decision-maker.
Equations 3.29a and 3.29b deal with the environmental variable according to Ruggiero's approach to partition the DMU set according to the level of environmental harshness and account for the one continuous or categorical environmental variable that influence the extent of the impact that the services have on the client or community (e.g. poverty level; % minority in the population, cultural characteristics, etc.). Equation 3.29a assures that the peer group set is chosen from all those DMUs operating at the same or lower level of environmental harshness $Z_j$. Equation 3.29b assures that if a DMU in the sample operates at a more favorable environment, it will not be included in the analysis.

Because the four formulations presented above do not use the resulting score from the previous node as input to the subsequent one, then the final performance score in Stage IV cannot be considered as being the overall score. If ranking of the DMUs is a need that the researcher has, then this formulation is not appropriate. However, in some applications ranking of the DMUs might not be necessary, since the most important products of DEA analysis are the performance targets or recommendations for improvement, the peer set for benchmarking, the weights for the input/output set etc.

The next section presents an alternative model that carries the projections to the subsequent node.

### 3.2.3 A Four-Stage Approach for Non-profit Performance Evaluation Using the Optimal Projections

The second model developed will carry the optimal projections of the outputs of the first stage as inputs in the second stage, the second to the third, the third to the fourth as in Athanassopoulos (1999) and Soteriou and Zenios (1999). This formulation accounts for the relation between two frontiers /each defined by each successive stage. It provides the best possible picture of the DMU being evaluated, in the sense that it considers the best possible level of the output variable in the preceding stage as an input to the next stage. The final score of the last stage (Stage IV) can be considered the actual performance score of the DMU being evaluated. This assumes management has taken action to make the corrective improvements recommended in the
previous stages, so that it is efficient with respect to the financial (fundraising or income generating), capacity creation, and service delivery level. The score in the Stage IV yields the overall effectiveness score.

The formulations at each stage are also output increasing as in the previous model, assume variable returns to scale, and are non-radial. Figure 3-6 illustrates the difference between using the actual values and using the projections of the DMU (the values of the virtual efficient DMU) as input for the subsequent stage. Figure 3-6 is an adaptation from Athanassopoulos (1999). The treatment of non-controllable environmental variables is the same as in the models proposed in the previous section.

The left part of Figure 3-6 depicts the assessment of income generation efficiency and the right part represents the assessment of capacity creation efficiency, given the amount of donations received. The level of expenses in fundraising, grant management and outreach directly affects the number of donations.

The income generation (fundraising) efficiency of DMU_k is assessed by projecting it to the financial frontier (line AB). The efficient position K' shows the level that K would have had to achieve with respect to fund raising, had it fully utilized its inputs to collect the maximum possible amount of donations. This efficient level K' is used as an input into the assessment of the capacity creation efficiency of DMU_k as illustrated on the right part of Figure 3-6. The capacity creation efficiency is estimated by projecting K' to K_2 on the line segment EF. The figure shows that in the case when one uses the real values of variables associated with DMU_k instead of using the projections of the variables the performance targets would have been underestimated. According to Athanassopoulos, the performance targets when considering observed values instead of projecting values from the previous stage are generally smaller than the corresponding targets obtained when using the projections (i.e. the maximum level of income generation in the example). That is, K'K_2 ≥ KK_1 in Figure 3-6.
As Athanassopoulos (1999) states, the rationale behind the pre-estimation of maximum levels of an output variable after controlling for the inputs, is that this process would eliminate any inconsistencies in the data of the variable being maximized, as well as biased answers obtained through questionnaires administered to obtain service quality perceptions and so forth.

In other words, if for a given level of inputs (e.g. managerial expenses) one can take the optimal projected value of the output "donations" and input it in the
capacity creation node, hence one is controlling for donations. This in the sense that all DMUs will be deemed as operating at their maximum potential given the level of inputs they have. Therefore, the capacity efficiency of node 2 will be strictly in the terms of the amount of capabilities generated, all things being equal (all DMUs at maximum level). Any inconsistencies or biased introduced in the donations variable, will be eliminated by this method.

3.2.3.1. Stage I. The Financial Frontier - DEA Model for the Financial Node

The formulation for the first stage is exactly the same as in the previous model, as follows. The first stage yields the efficient level of donations and in-kind contributions for each DMU.

Stage 1:
Max \( \sum w_i \varphi_{b_{ij}} \)  \( \quad (3.31) \)

s.t.
\[
\sum_{j=1}^{n} \eta_j x_{b_{ij}(t-1)}^{fr} \leq x_{b_{ij}(t-1)}^{fr} \quad i \in F \quad (3.32) \quad \text{fundraising \& outreach cost}
\]
\[
\sum_{j=1}^{n} \eta_j x_{m_{ij}(t-1)}^{e} = x_{m_{ij}(t-1)}^{e} \quad m \in E \quad (3.33) \quad \text{Non- controllable environmental factors}
\]
\[
\sum_{j=1}^{n} \eta_j x_{k_{ij}(t-1)}^{d} \geq \varphi_k x_{k_{ij}(t-1)}^{d} \quad k \in D \quad (3.34) \quad \text{Donations \& in- kind contributions}
\]
\[
\sum_{j=1}^{n} \eta_j = 1 \quad (3.5)
\]

\( \eta_j \geq 0 \forall j \ni \varphi_{m_{ij}} \leq \varphi_{m_{ij(o)}} \), \( \quad (3.36a) \quad \text{Factor of Environmental Harshness} \)
\( \eta_j = 0 \forall j \ni \varphi_{m_{ij}} > \varphi_{m_{ij(o)}} \), \( \quad (3.36b) \)
3.2.3.2. Stage II: The Capabilities Frontier - DEA Model for the Capacity Creation Node utilizing the Efficient Level of Income Generation (fundraising) from Stage I.

In the second stage of this model, the targets for donations (projections) recommended by DEA for DMUs not in the best-practice frontier, $\varphi_{i jo} x_{i jo}$, are used as discretionary inputs (Eq. 3-26). $\varphi_{i jo} x_{i jo}$ is the efficient level of donations and in-kind contributions obtained from Stage I and these levels are treated as inputs in Stage II. This implies that for higher levels of donations there will be always higher levels of acquired capabilities. This stage controls for environmental variables in 3.38.

Stage II:

Max $\sum w_k \beta_{rjo} + w_k \gamma_{sjo}$ \hspace{0.5cm} (3.37)

s.t.

$\sum_{j=1}^{n} \alpha_j x_{mj}^c = x_{mj_s}$ \hspace{0.5cm} $m \in E$ \hspace{0.5cm} (3.38) \hspace{0.5cm} Non-controllable environmental variable

$\sum_{j=1}^{n} \alpha_j \varphi_{sj} x_{b(i-1)}^d \leq \varphi_{sj} x_{b(i-1).o} \hspace{0.5cm} k \in D \hspace{0.5cm} (3.39)$ \hspace{0.5cm} Projection of Max. Level of Donations

$\sum_{j=1}^{n} \alpha_j x_{rj}^p \geq \beta_{rjo} x_{rj.o} \hspace{0.5cm} r \in P \hspace{0.5cm} (3.40)$ \hspace{0.5cm} Controllable created capabilities

$\sum_{j=1}^{n} \alpha_j x_{sj}^c \geq \gamma_{sjo} x_{sj.o} \hspace{0.5cm} s \in C \hspace{0.5cm} (3.41)$ \hspace{0.5cm} Controllable attributes of capabilities

$\sum_{j=1}^{n} \alpha_j = 1 \hspace{0.5cm} (3.36)$

$\alpha_j \geq 0 \forall j \ni z_{mj}^e \leq z_{mj_o}^e$, \hspace{0.5cm} (3.42a) \hspace{0.5cm} Factor of Environmental Harshness

$\alpha_j = 0 \forall j \ni z_{mj}^e > z_{mj_o}^e$, \hspace{0.5cm} (3.42b)
3.2.3.3. Stage III: The Service Delivery Frontier - DEA Model for the Services Node with the Projection of the Maximum Level of Capabilities from Stage II

In the third stage of this model, $\beta_{rjo}x^p_{rjo}$ and $\gamma_{sjo}x^c_{sjo}$ are used as discretionary inputs, which are the projections to the frontier obtained in the previous stage. $\beta_{rjo}x^p_{rjo}$ is the efficient level of created capabilities obtained in Stage II which is treated as an input in Stage III. This implies that for higher levels of donations there will be always higher levels of acquired capabilities. $\gamma_{sjo}x^c_{sjo}$ is the efficient level of attributes of those capabilities obtained in Stage II which are used as inputs in Stage III.

Qualitative attributes or characteristics of those capabilities (such as level of training of staff and volunteers) are assumed to have a weak disposable behavior (Eq. 3-41). So, a reduction in qualitative characteristics does not always imply a reduction on services and services characteristics. Hence, this formulation calculates the efficient level of service and service quality given the maximum attainable level of capabilities and capabilities attributes for that DMU.

\textit{Stage III}:

\[
\text{Max} \sum (w_3, \rho_v + w_4, \pi_{qjo}) \quad (3.43)
\]

1. \[
\sum_{j=1}^{n} \sigma_j x^c_{mj} = x^c_{mjo} \quad m \in E \quad (3.44) \text{ Non-controllable environmental variable}
\]

2. \[
\sum_{j=1}^{n} \sigma_j \beta_{rjo} x^p_{rjo} \leq \beta_{rjo} x^p_{rjo} \quad r \in P \quad (3.45) \text{ Projection of controllable capabilities}
\]

3. \[
\sum_{j=1}^{n} \sigma_j \gamma_{sjo} x^c_{sjo} = \gamma_{sjo} x^c_{sjo} \quad s \in C \quad (3.46) \text{ Projection of attributes of capabilities}
\]

4. \[
\sum_{j=1}^{n} \sigma_j y^p_{vq} \geq \rho_{qjo} y^p_{vq} \quad v \in S \quad (3.47) \text{ Services delivered}
\]

5. \[
\sum_{j=1}^{n} \sigma_j y^q_{qo} \geq \pi_{qjo} y^q_{qo} \quad q \in Q \quad (3.48) \text{ Attributes of Services delivered}
\]

6. \[
\sum_{j=1}^{n} \sigma_j = 1 \quad (3.49)
\]

(a) \[
\sigma_j \geq 0 \quad \forall j \quad z^c_{mj} \leq z^c_{mj}, \quad (3.50a) \text{ Factor of Environmental Harshness}
\]

(b) \[
\sigma_j = 0 \quad \forall j \quad z^c_{mj} > z^c_{mj}, \quad (3.50b)
\]
3.2.3.4. Stage IV: The Effectiveness Frontier -DEA Model for the Customer Node using the Maximum Possible Level of Service and Service Quality from Stage III

In the fourth stage of this model, $\theta_{tjo} y^p_{tjo}$ and $\lambda_{qjo} y^q_{qjo}$ are used as discretionary inputs. $\theta_{tjo} y^p_{tjo}$ is the efficient level of service obtained in Stage III and for which weak disposability is assumed (Eq. 3.53). $\lambda_{qjo} y^q_{qjo}$ is the efficient level of service quality (attributes of the services) obtained in Stage III and for which weak disposability is assumed (Eq. 3.54). This implies that for higher numbers of service or higher levels of service quality, there is no guarantee that there will be higher levels of outcomes. This embodies the lack of unequivocal evidence in the literature regarding the direction and sign of the causal relationship among those variables in social services, if any.

Stage IV:

Max $\sum w_{5h} \hat{\delta}_{bjo}$  \hfill (3.51)

s.t.

$\sum_{j=1}^{n} \phi_j x_{mj}^e = x_{mj_o}^e \quad m \in E$  \hfill (3.52) Non – controllable environmental variable

$\sum_{j=1}^{n} \phi_j \rho_{vj} y^p_{vj} = \rho_{vj_o} y^p_{vj_o} \quad t \in S$  \hfill (3.53) Projection of Services delivered

$\sum_{j=1}^{n} \phi_j \pi_{qj} y^q_{qj} = \pi_{qj_o} y^q_{qj_o} \quad q \in Q$  \hfill (3.54) Projection of Service Quality (Attributes)

$\sum_{j=1}^{n} \phi_j y_{hj}^o \geq \hat{\delta}_{h} y_{hj_o}^o \quad h \in O$  \hfill (3.55) Outcomes

$\sum_{j=1}^{n} \phi_j = 1$  \hfill (3.56)

$\phi_j \geq 0 \quad \forall j \quad z^e_{mj} \leq z^e_{mj_o}$, \hfill (3.57a) Factor of Environmental Harshness

$\phi_j = 0 \quad \forall j \quad z^e_{mj} > z^e_{mj_o}$, \hfill (3.57b)

The previous sections outlined two alternative DEA models to approach the problem of evaluating efficiency, effectiveness, and service quality in the provision of social services in the nonprofit sector. The results and insights expected from the
application of both are very different. The first one provides an evaluation of the current real situation of the organization under study, and recommends targets for those inefficient performance indicators. The second projections-based model provides an evaluation assuming the maximum level possible of performance given a constant amount of inputs to the first node.

Chapter 4 presents the development of the data collection instruments, sampling and data analysis procedures used, as well as issues on variable selection in DEA. Chapter 5 introduces an illustrative application to a reduced sample of DMUs at the a social services organization that participated in this study and Chapter 6 presents the discussion and conclusions for future research.
CHAPTER 4 Research Methodology: Data Collection and Variable Selection

Data collection is a crucial step in the development of performance measurement systems and for the success of any DEA analysis.

The purpose of this chapter is to describe the data collection methodology developed for this study. This methodology was necessary to implement the models developed in Chapter 3 in a “real-life” application. Hence, Chapter 4 includes a discussion regarding the development of appropriate data collection instruments, data sampling, and data analysis procedures. A brief description of the case study research site is also included as well as of the two different units of analysis to which the models were applied. A description of all variables selected is included as well. This chapter addresses several of the research objectives described in Chapter 1.

4.1 The Case Study Research Site

4.1.1 The Agency Under Study: A Social Service Organization

The agency under study is a humanitarian organization guided by volunteers. With a large network of field offices. National headquarters are located in the Washington D.C. area. The Agency has five service divisions, four of which operate under the field office administration in Falls Church (Armed Forces Services; Disaster Services\(^4\); Health Services; and International Services). Each field office is required to provide a number of services, besides other optional services that are generally adapted to their community.

\(^4\) As Of March 2002, Disaster Services became an independent division but still operating at the field office level.
4.1.1.1. Disaster Services

The authority of the agency to provide emergency relief after disasters was formalized in 1905, when chartered by Congress. Disaster relief focuses on meeting people's immediate emergency disaster-caused needs. When a disaster threatens or strikes, the agency provides food, health, shelter and mental health services to address basic human needs. In addition to these services, the core of the agency's disaster relief is the assistance given to individuals and families affected by disaster to enable them to resume their normal daily activities independently.

The Agency also feeds emergency workers, handles inquiries from concerned family members outside the disaster area, provides blood and blood products to disaster victims, and helps those affected by disaster to access other available resources. This is done through their offices located nationwide. Also, its response is done through a system of trained disaster workers and volunteers, the Disaster Services Volunteers system (DSV) that are on call to be mobilized in case of disaster in less than 24 hours to any part of the country where the Agency is responding to an emergency. They are a clear measure of capacity to respond.

Unit of Analysis

The unit of analysis will be in one case, Disaster Services departments located in field offices. Specifically, disaster services delivered at division level will be considered for the first run of analysis. The second illustrative example is done with data from all lines of service at the field office level, and therefore, the unit of analysis will be the total operation of the field office. Data regarding financial measures as well as process measures (inputs and outputs) are regularly reported to National Headquarters. Service quality and effectiveness data was collected on a pilot basis from the field. All measures have been treated and interpreted with the authorization of the agency.

4.2 Variable Selection: Input and Output Considerations for DEA

This section discusses some fundamentals of variable selection for DEA. This is a crucial stage in all DEA implementations. While the creation of the
formulations and the development of the code to execute these formulations are important, the modeling effort can be squandered if the variable selection is not done correctly. DEA is extremely sensitive to the variables chosen. The following is a list of guidelines found in the literature (Banker et al, 1989; Chalos and Cherian, 1995; Epstein & Henderson, 1989; Lewin et al., 1982) and also taken from our own experience with the present research study. Whenever possible bibliographic references have been provided as to the author(s) reporting specific guidelines. All guidelines were applied during the course of this research and the insights obtained will be further discussed in Chapter 6:

1. Outputs must be the result of—or a product of—input consumption. The model needs to make sense in a microeconomic and production sense.

2. An empirical precedent needs to exist for variable inclusion in the model (Epstein & Henderson, 1989). For example, the researcher needs to substantiate the inclusion of aggregated household income or community wealth in the model as a variable that affects fund-raising ability with the latest findings in nonprofit literature.

3. Banker (1989) states that it is highly recommended that the researchers rely on expert judgment within the organization under analysis. In this way, one can clarify which variables have a major impact on performance and therefore should be included in the DEA model.

4. With respect to the number of inputs and outputs that are included in the DEA model, there is no ideal quantity. DEA is, however, better able to identify inefficient performance when the number of periods or DMUs is greater than the number of inputs and outputs combined. DEA can handle as many separate inputs and outputs as desired but the discriminating power of the approach diminishes considerably with each variable added. In the worst case scenario, all DMUs will appear to be efficient if the number of variables restricts the solution space so much that each DMU appears to be "unique", and consequently, efficient.

5. There may be a time lag associated with specific input variables. Some inputs used now may not effect the production of outputs until a later time period.
6. Ideally, outputs and inputs should be measured in tangible/physical units. This prevents distortions due to inflation and provides a physical interpretation of each variable. Other measures such as indexes, ratios, percentages, are also admissible.

7. The relationship between the inputs and outputs should also stay consistent. Major changes in the production process can have an influence on performance that is not measured in the model —i.e. the production function/process should not change dramatically during the study. This stability allows the input/output correspondence set to remain constant.

8. To insure reliable results, Lewin et al. (1982) suggested that the analyst use a step-wise regression approach to identify potential input/output variables. Incorporating only those variables that are statistically significant helps in determining those inputs most closely associated with performance. During the final stage of the variable selection process variables should be tested for statistical significance. Inputs should be regressed against outputs and outcomes. I used SPSS to run regressions on all selected inputs, outputs, and outcomes. While it is true that statistical significance does not by itself demonstrate the necessity of including a variable in the analysis, it is also true that including and hence evaluating non-significant variables may lead to taking action on inputs that have little benefit to the DMUs being analyzed. Others support this approach (Chalos and Cherian, 1995) in order to minimize the specification error. There is also a step-wise mathematical programming approach that computes the impact that each additional variable (input/output) has on the efficiency score (see Ruiz and Sirvent, 2001).

The following guidelines should be applied once the variable set has been selected:

9. The data used in the analysis should be checked for consistent reporting. DMUs need to report the data on a consistent basis.

10. The data must be checked for unusual conditions (vacations, strikes, natural disasters, wars, special grants, gifts etc.) that affect operations. These special conditions might
affect the analysis by shifting the frontier to an unrealistic level, due to DMUs exhibiting unusual behavior.

11. The data should have no zero elements unless zero resource consumption or zero production is part of the production possibility set.

12. The data should have no missing values. DEA cannot handle missing values. In the course of this study DMUs that did not report complete data sets were deleted and could be included in the analysis once all variables are reported.

13. Data should be checked for influential observations (outliers and leverage points). Influential observations can represent true production occurrences or measurement error. In the case of measurement error the data should be corrected or eliminated. (Seaver and Triantis (1995), Seaver, Triantis and Reeves (1999)).

All these guidelines outlined above were followed during variable selection and depuration of the variable set. The Agency has access to a large data set from most field offices, which includes over 100 different types of service delivery and financial data. Following the guidelines above assured the performance measurement model to be a precise and comprehensive representation of field office operations, without redundant variables.

### 4.3 Variable Selection for the Present Study

Of all available variables regularly collected and available for inputs and outputs, a sub-set was selected following the steps described in the previous section. Table 4.1 shows the type of data (performance indicators) and data collection instruments needed for this research.

#### 4.3.1 Data collection for inputs and outputs

Data for inputs and outputs was obtained from the Field Office Administration and Evaluation division at the agency's national headquarters. Field offices report these measures on a regular basis as part of an information system called the Field Operations Information System, FOIS. Table 4.2 presents the available inputs and outputs for Disaster services in FOIS.
The variables currently being collected can be classified into the following categories:

- Financial indicators: indicators of financial health that are unique to non-profits (such as per household revenue due to fund raising efforts; days of net assets etc.).
- Human Resources: volunteer and paid staff data at different levels of aggregation (total number of volunteers, registered volunteers, youth volunteers, etc.)
- Service Delivery data: Inputs and outputs at different levels of aggregation (Number of disaster specialist staff, total trained people, numbers of trained people by course type, etc.)
- Demographic data pertaining to each individual field office, like population served, % minority, aggregated household income, etc.

A service delivery profile of each field office includes 61 variables. If one adds the service quality dimensions, customer satisfaction and outcome data for different lines of service, the number of variables can easily jump to 90.
<table>
<thead>
<tr>
<th>Performance Dimension</th>
<th>Performance Indicators Required</th>
<th>Data Collection Instrument</th>
<th>Source of the Data (Administrative Unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service Quality</td>
<td>♦ Customer Satisfaction ♦ Service Quality</td>
<td>Customer Survey Customer Survey</td>
<td>♦ Administration and Evaluation Division of the Agency, located at National Headquarters, in Falls Church, VA. ♦ Data requested directly from selected field offices in the nation.</td>
</tr>
<tr>
<td>Technical efficiency</td>
<td>♦ Outputs ♦ Inputs</td>
<td>Internal Reports of the Organization</td>
<td>♦ Administration and Evaluation Division of the Agency, located at National Headquarters, in Falls Church, VA.</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>♦ Outcomes Obtained</td>
<td>Customer Survey</td>
<td>♦ Data requested directly from selected field offices in the nation through Administration and Evaluation Division of the Agency, located at National Headquarters, in Falls Church, VA.</td>
</tr>
<tr>
<td>Demographic and Environmental Data</td>
<td>♦ Environmental Indexes</td>
<td>Public and Internal Sources</td>
<td>♦ Data is available at the Administration and Evaluation Division from different sources, including census bureau and marketing firms that have provided the data to the agency.</td>
</tr>
</tbody>
</table>
Table 4.2. Inputs and Outputs Used for the First Data Set regarding Disaster Services collected by FOCIS

<table>
<thead>
<tr>
<th>Type of Variable</th>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Node I: Financial (fundraising or income generating) Node</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Input variables</strong></td>
<td>FNDRSNEXP</td>
<td>Fundraising, income generating and managerial expenses incurred by the field office for fundraising activities, grant and contract management</td>
</tr>
<tr>
<td><strong>Output variables</strong></td>
<td>FRAISREV99</td>
<td>Total revenues that can be attributed to the field office's fund raising efforts including: monetary contributions and special events for FY 99, revenues from legacies and bequests for FY 99, contributions to the Disaster fund, in kind contributions.</td>
</tr>
<tr>
<td><strong>Node II: Capacity Creation Node</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Input Variables</strong></td>
<td>FRAISREV99</td>
<td></td>
</tr>
<tr>
<td><strong>Output Variables</strong></td>
<td>DSV_ENROL</td>
<td>Number of DSV members ready to be deployed in case of disaster.</td>
</tr>
<tr>
<td></td>
<td>DSVOLS_REG</td>
<td>Number of Disaster registered adult volunteers (18 and older). These are volunteers who have demonstrated an ongoing commitment to the Agency and whose name and address are on file at the field office.</td>
</tr>
<tr>
<td></td>
<td>DSTOTSTAFF</td>
<td>The number of total staff assigned to Disaster activities. Depends on the actual number of hours worked per week, including job related training, seminars, workshops, travel time during working hours, but excluding paid leave such as holidays, annual leave, and sick leave.</td>
</tr>
<tr>
<td></td>
<td>DISASEXP</td>
<td>Expenses for services to families or individuals affected by local disasters.</td>
</tr>
<tr>
<td></td>
<td>DAYS_NET_ASSETS</td>
<td>Number of days a field office can operate at its current level of expense making use of its current assets.</td>
</tr>
<tr>
<td><strong>Node III: Service Delivery Node</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Input Variables</strong></td>
<td>DSV_ENROL</td>
<td>Same as above</td>
</tr>
<tr>
<td></td>
<td>DSVOLS_REG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DSTOTSTAFF</td>
<td></td>
</tr>
<tr>
<td><strong>Output Variables</strong></td>
<td>Disaster_Fam_assist</td>
<td>Number of families who suffered a local disaster (generally fire, flood, etc.) who received disaster assistance.</td>
</tr>
<tr>
<td><strong>Service</strong></td>
<td>Timeliness</td>
<td>% of times that an agency’s worker was on scene in less</td>
</tr>
<tr>
<td>Quality Index</td>
<td>Service quality index calculated as a composite of service quality items in survey.</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------------------------------------------------------</td>
<td></td>
</tr>
</tbody>
</table>

**Node III: Customer Node**

**Input Variables**
- Lack of timeliness
- Lack of Service Quality

(1-Timeliness) = % of times the agency was not on scene in less than 2 hours.

Index of Poor quality (inverse of Service Quality Index) calculated from the service quality results.

**Outcome Variables**
- Recovery Needs

Proportion of clients who stated that the service helped them recovery back to their pre-disaster status
Proportion of clients who stated that the service covered their most immediate needs

**Other variables**
- CITY
- STATE
- REGION
- POP99
- Percent Minority
- 2000 Households

City Name
State abbreviation
Region Name
The total population in the field office’s jurisdiction
Percentage of total population that are minorities including: Black, Hispanic, Asian or Pacific Islander, and American Indian or Alaskan Native based on estimates from CACI Marketing Systems.
Total number of households in the field office's jurisdiction based on estimates from CACI Marketing Systems.

Initial DEA runs to calculate technical efficiency of inputs and outputs only with the help of available commercial software⁵ proved that the number of variables limited the degrees of freedom to such an extent that no meaningful results could be obtained. Furthermore, as several variables are the aggregation of others, some of the inputs and outputs were duplicated.

All these limitations indicated a need for refinement of the data set with the help of statistical techniques for variable data reduction (principal component analysis)

---

⁵ DEA Solver
• Principal component analysis was used to refine the data set identifying groups of variables measuring the same factor. In that way, perhaps only one variable of each factor could be used to measure performance.

• Regression analysis (regressing variables to the most important outputs to see which are the best predictors of output performance.)

• Correlation analysis (for DEA analysis it is not advisable that input variables be strongly correlated among themselves but need to be correlated to the output variables).

The reduced set of variables chosen by intuition which are also statistically-significant for subsequent input into the DEA model must satisfy the following conditions (Chalos & Cherian, 1995):

1. An empirical precedent existed for variable inclusion in the model (Epstein & Henderson, 1989). The research work exposed in Chapter 3 supported the inclusion of each variable.

2. Agency's officials at all levels considered the variables to be appropriate measures of efficiency and effectiveness. They also validated the assumption that the use of the inputs selected for each node in fact produced or generated the outputs and outcomes expected from each node. The need to ensure that the measures used in the DEA model reflect the relevant perspectives of the organization has been emphasized in the literature (Rouse et al., 1997). The validity of DEA results is often supported by testimony of their usefulness by managers within the organization.

3. A statistically significant relationship existed between the chosen inputs and outputs. All inputs and outputs were checked using regression analysis to test the significance of including them in the model. Furthermore, inputs were not correlated among themselves.

4. The minimum measurement had an amount of zero.
5. Few missing values existed. When they existed, for illustrative purposes those DMUs were excluded from the analysis.

6. A non-arbitrary data collection source and methodology existed for all variables. The FOIS system instructs field offices in how to report these variables, and performs quality assurance of the numbers reported.

The same guidelines plus extended brainstorming sessions with agency’s officials were applied to select the variables for the total field office operations assessment, i.e. to validate the second data set.

4.4 Development of Data collection Instruments and Data Collection

The model(s) developed in Chapter 3 require data with respect to all the performance dimensions chosen for this analysis.

For service quality and effectiveness data, there are two potential sources of data. One is data collected with respect to operational measures of customer satisfaction, such as the percentage of mistakes in an emergency communication or the percentage of times the service was available immediately to the client. Effectiveness data can be collected in a proxy manner from public records. For example, a program targeted at providing training and employment to social-security dependent single mothers could conceivable claim as long-term outcomes the reduction in social security dependants in the county where the program is implemented. But these types of linkages are difficult to prove and data for these outcomes are not always available. Hence, the most common data collection method is through written surveys.

Survey questionnaires have been applied in the past to elicit outcome-related data (Schalock, 1995; Ruggiero, 1995). Although hard data from public and internal sources has also been used as outcome indicators, such as felony and misdemeanor statistics and budget allocations (Todd & Ramanathan, 1994), the use of public data has limitations. It is difficult to link such indices to the services actually provided. Even if public data is available, the problems targeted by some programs are so specific that only in rare instances it will be possible to find published indicators that
measure what we want to assess. Furthermore, the objective of certain programs might be to provide services for clients on an ongoing fashion, which in and of itself benefits the community. In this situation, the provision of infrastructure is the service, thus, outcome data here is more related to customer satisfaction. In these cases, measurement tools such as surveys and client interviews should be designed to collect appropriate data (Triantis and Medina-Borja, 1996). The best instruments are those that can show a change in knowledge, skills, attitudes, behavior, etc.

### 4.4.1 Survey Instrument for Service Quality and Outcome Data

Ten survey instruments were developed to collect service quality and outcome data from agency’s clients across four lines of service: Disaster Services, Health and Community Services (6 questionnaires), Armed Forces and International Services (2 questionnaires). The constructs of the survey will follow those proposed by Parasuraman et al. (1988) for the SERVQUAL tool but adapted to fit the Agency’s operating environment. For example, the appearance of the facilities is not important since most services are provided over the phone, on the disaster scene or through fax or e-mail. In that respect, the authors of SERVQUAL assessed the importance of service quality dimensions in four service sectors (Berry et al., 1988). They found that the construct “tangibles” was consistently ranked last with a mean importance of around 7 on a 10-point scale, while all other four dimensions were ranked in the 9 range. In chapter 2, I presented a discussion about the SERVQUAL scale comparing it to other frameworks developed to measure service quality and customer satisfaction.

While SERVQUAL measures perceptions minus expectations of customers, the use of instruments measuring perceptions only has been widely supported in recent years. One of the developers of SERVQUAL (Parasuraman, 1995) has concluded that a perceptions-only approach to measuring quality is even more acceptable from a predictive validity point of view. De Ruyter et al. (1997) also found that the perception of the service performance is the most important indicator of service quality over expectations or other measures.
I used a scale for service quality performance that measures timeliness, empathy and other constructs comparable to those on SERVQUAL and one item with respect to overall customer satisfaction. All items will measure perceptions and not expectations. I will not use, however, a replica of the 22 items in SERVQUAL since the actual questions need to be adapted to the language and culture particular to the agency. While I recognize both constructs —service quality and customer satisfaction— as different, one of the premises of this research is that they are complementary and that service quality causes customer satisfaction.

4.4.1.1. Selection of Scales

There is a lot of discussion in academia about how many points and what kind of labels (anchors) comprises the most effective measurement tool (e.g. Bruner et al, 2001; Devlin et al, 1993, Schmalensee, 1994). There are dozens of ways a scale can be labeled, and researchers do not always agree in what a good scale is. Each scale has variations, some more reliable than others (Devlin et al., 1993; NCS Pearson, 2002). One of the statistical tests on scales that can be performed to assure that the respondent understands the measurement intent of the scale, or the “survey performance” defined as stability, repeatability and accuracy (Devlin et al, 1993; Grapentine, 1994; NCS Pearson, 2002) is the reliability coefficient Alpha Cronbach (1951) that measures internal consistency. Another potential test to check whether the instrument measures what is supposed to measure is to perform factor analysis over all the items in a survey (Devlin et al, 1993; Grapentine, 1994). The objective is to check whether the grouping of the items based on the trends in responses actually agree with the grouping of the intended constructs the researcher wants to measure. According to NCS Pearson, a survey cannot be more valid than it is reliable. The following is a list of issues that one needs to address in building survey scales:

**Even vs Odd.** Several researchers believe that even numbered scales better discriminate between satisfied or unsatisfied customers, positive and negative reactions or perceptions because there is not a neutral option (NCS Pearson, 2002).
the other hand, others believe that without a midpoint option respondents often choose a positive response, therefore skewing the data. For this study, the service quality and effectiveness constructs all had even numbers scales (six points).

**Number of Points.** Common scales in survey research go from 3 to 10 points (Bass et al, 1974; Cohen, 1983). There is a lot of controversy also in this issue (Devlin et al, 1993; Grapentine, 1994; Schmalensee, 1994; NCS Pearson, 2002). Although seven to ten point scales may gather more discriminating information, these scales also face the problem of whether respondents are even capable of discriminating differences among points or whether they will take their time to actually even consider all points presented to them and make an informed decision when answering. If respondents cannot discriminate between an 8 and a 9 in a 10-point scale, then the value of using such a scale is questionable. At the end, for reporting purposes many researchers and companies collapse the results into 3 or 5 points scales. On the other hand, three and four point scales have little discriminative value. A two-point scale is used more for binary yes or no questions items not for measuring degrees of perception of some attribute from customers.

Chang (1994) argues that respondent knowledge of the subject matter is important to consider. He suggests that a lack of respondent knowledge may lead to an abuse of the endpoints of longer scales resulting in lower reliability than with the shorter scales. Consequently, it is worth considering how familiar your potential respondents are with the issues you are addressing in your survey. Chang also suggests considering the similarity (or lack) of the respondents' frames of reference for the issues addressed in the survey. Specifically, he argues using more response options may introduce error when a respondent group has very different frames of reference. If the points are not labeled, this scenario provides more opportunity for the respondents to introduce their own unique frames of reference.

For this research study four points were considered not to have discriminatory power. Furthermore, this study faced the constraint of not wanting clients to be undecided, so another requirement was to avoid a middle point, therefore a five point scale was discarded. Hence, a 6 point scale was selected to assure not middle point
and enough discriminatory power. NCS Pearson (2002) research suggests that there is no best number of response options for all applications.

**Labels.** Once the number of points in the scale is defined, the researcher needs to determine the labels, if any. Sometimes only the end points of the scale are labeled with words, which are called “anchor” points. The logic behind this method is that it provides a “nominal” scale with equal intervals between each point in the scale. It avoids respondents making more subjective decisions about the differences between word labels for each scale point (Fink, 1998, Hayes, 1992). However, there are as many researchers in favor of this approach as there are against (NCS Pearson, 2002). Researchers against the anchors at the extreme have two major concerns: first, some respondents may not figure out what the numbers between extremes really mean, unless words are attached to them. Second, they argue that by labeling each scale point all respondents attach the same word to each numerical value. This helps to avoid response misinterpretation of scale definitions and furthermore, it allows reports from the surveys to be written in more concrete terms. This last statement was also one of the drivers of some evaluation units at the agency to suggest labeling each point in the scale.

Bass et al (1974) conducted research in this area and provided a table summarizing a variety of response anchors to use for a number of different scenarios. Two examples of their research, which have shown to discriminate well, are:

- Never, occasionally, fairly many times, very often, always
- None, some, very often, quite a bit, extremely often, all

Of course, the issue of equally spaced points comes into play and some could argue that the labels used in Bass’ study do not correspond to equal spacing. The real issue is to accommodate the labels to what the target respondent can identify and understand. In the case of organizational research, the labels need also to fit the company’s culture. Additional response anchors are available in the literature and can be used to fit a variety of situations. For example, a group of organizations that survey regularly and share their survey results with other members of the consortium, have
several response anchors for the scales they use. Some of these are reproduced below in table 4-3.
Table 4-3. Sample of Labels Used in Organizational Research (Bruner et al, 2001; NCS Pearson, 2002)

<table>
<thead>
<tr>
<th>Far too much</th>
<th>Much higher</th>
<th>One of the best</th>
<th>Very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>Too much</td>
<td>Higher</td>
<td>Above average</td>
<td>Good</td>
</tr>
<tr>
<td>About right</td>
<td>About the same</td>
<td>Average</td>
<td>Fair</td>
</tr>
<tr>
<td>Too little</td>
<td>Slightly lower</td>
<td>Below average</td>
<td>Poor</td>
</tr>
<tr>
<td>Far too little</td>
<td>Much lower</td>
<td>One of the worst</td>
<td>Very poor</td>
</tr>
</tbody>
</table>

As one can see, there is no one right scale to use for all occasions nor the anchors need to be labeled representing equally spaced concepts, even though that is the theoretical assumption (e.g. Flynn and Pearcy, 2001; Schmalensee, 1994). Also, to evaluate whether the difference between “below average” and “one of the worst” is equal to “average” and “below average” is usually a subjective judgement. In fact, Bruner et al (2001) presented a total of 941 scales being used at this moment.

At the agency under study, all surveys suffered several interactions before and after the pilot test, including refinement of the labels from other evaluation units/experts in the office. Also, the labels used tried to mimic whenever was appropriate those used for customer surveys done in the past within the organization. In that way, each group of labels for each survey varied by department as shown in Table 4.4 of this document. Some had the end point or anchors only, and some had labels for each point in the scale. The choice depended a lot of the results of the refinement process with the program evaluation units at each department. Different agency's researchers argued in pro or against one or the other. Whenever a service had clients or customers with very low levels of education, the tendency was to label all points in the scale. The reliability analysis results shown in table 4-5 shows that the issue of labeling affected very little the reliability of the instruments.
4.4.2 Data collection for Outcome Indicators

As stated before, the same customer survey used to collect service quality and customer satisfaction data can be and was used to collect outcome-related data. The procedure to determine the right scales was basically the one described in section 4.4.1 above. The only difference is that the labeling of the scales needed to reflect the aimed change on the client and degree of outcome achievement that the program is having on the client. In that sense, labels such as “Describes me totally” or ‘Describes me a little” where used for outcomes that required the intention of the client to perform an action. In order to demonstrate that the desired outcome is related to the program, the agency has to design and apply a specific measurement instrument that will track the changes in the actual clients who are receiving program’s services. Agency’s field office directors identified the appropriate outcomes for agency’s services. Triantis and Medina-Borja identified these outcomes during several outcome-based objectives development workshops facilitated from 1996 to 1998. The outcome-based objectives development process is described in Appendix A.

The development of valid constructs for client outcomes is a crucial step in the proposed methodology. To do this, observations are classified according to a standardized procedure to indicate the presence, absence, or degree of the given conceptual attribute. Measuring client outcomes requires rules for making inferences about knowledge, skills, attributes, intentions or motives; or systematic procedures for classifying observed behavior or reports of behavior (Magura and Moses, 1986). In many instances, multiple items are needed to operationalize adequately each construct (Steward, Hays and Ware, 1992). This was the case for some of the surveys developed during the time of this study, as is the case for the different types of Health & Safety training. For these cases, one survey item would have not been sufficient by itself to operationalize the sought outcome. For example capturing the behavior of an individual that uses his/her skills during an emergency situation, such as in the presence of a heart attack victim.
4.4.2.1. The Design of the Instrument

A number of data gathering tools to capture data (service quality / customer satisfaction and outcome indicators) that were subsequently used by the DEA model were developed as follows:

a. The variables that the questionnaire measures were identified and selected. In the case of outcome indicators, this has been done after following the outcome-based development process with experienced social workers, social science and program evaluation officials, etc. The outcomes obtained by these practitioners during a series of workshops facilitated by this researcher and her advisor will be selected as the constructs for the questionnaire. After the initial pre-selection, evaluation officials at the Agency were consulted as to the validity of the variables from a programmatic point of view.

b. Theoretical recommendations (e.g. Alreck and Settle, 1995; Fink, 1995; Sudman and Bradburn, 1982) with respect to survey construction were examined and applied regarding the planning and design phase of survey construction —i.e. construct identification, composing questions, creating item scales. The overall purpose is to build the final questionnaire with the respondents' in mind, for their ease and highest comprehension. Other similar data collection instruments built to elicit data regarding similar problems inside and outside the organization were also examined and the historical results (response rate, problems faced, etc.) were taken into consideration. The following decisions were made:

- The number of items to operationalize each construct.
- The grammatical construction of the items. Looking for the best way to get the respondent to understand the question and identify him-herself with the question.
- The size of the scale and the value of each point in the scale. There were two very different inclinations at the agency. One that pushed for a 10-point Likert scale advocating that a wider scale gives better spread of the variance. The second one wanted a 4-point Likert scale to accomplish two things. First, to make it simpler and clearer for the customer to understand the scale and second, to guarantee a higher
percentage of "Excellent" scores from respondents who otherwise, will tend to give a "Very Good" score. This researcher chose a compromise by using a 6 point, (no middle point) scale for all service quality and effectiveness outcomes. The advantages of this decision are that many respondents who were satisfied with the service but not "delighted" with it had the opportunity to rate the service as a "Very good" instead of good. In doing so, the DEA model will recommend the field office to thrive to achieve a higher percentage of "excellent" ratings. This is important because there is a tendency inside some organizations of increasing the proportion of excellent ratings obtained from survey research by giving the respondent fewer "good" choices. Conversely, a 6-point scale grants opportunity for improvement at the field office side. The main disadvantage of using a wider scale is that for some clients/respondents differentiating among 6 different rating options may be difficult. Some respondents, for example, may not distinguish between a rating 2 vs. 3 or between a 4 vs. 5. So, a 6-point scale for some respondents may not be meaningful. This problem can be overcome by labeling each of the points in the scale very well.

♦ Formatting issues such as the font size, the distribution of the items in the page, the use of one or two sided-pages, the use of color and the information and instructions provided to the respondents in the instrument were widely discussed with several evaluation professionals and officers in the field.

c. Pilot testing of the instrument was conducted to refine the tool through an actual pilot application. The survey was applied to clients that received a service from the agency during 2001. Over 50 field offices all over the country were asked to participate in the pilot. Less than 20 accepted and distributed the questionnaires to their clients. At least 40 responses per field office were needed to perform the reliability tests. At least 100 responses were received for each type of service at the time the reliability analysis was performed.

d. Using SPSS, the reliability coefficient Alpha Cronbach (1951) was calculated for each construct in each of the eight surveys. Correlation coefficients among items were also obtained. These two statistics measure the strength of the correlation of items pertaining to the same construct. The analysis of the data collected
served as a test of reliability and validity of the instrument. Table 4.4 shows the variables, constructs and items in the Disaster Services questionnaire (Appendix D). All other 10 surveys were developed in the same way.

e. As an additional statistical aid, factor analysis was conducted on the results obtained from each survey. Table 4.5 shows the results of the reliability test and a Factor analysis conducted in all items. For Disaster Services, for example, three factors were identified. These 3 factors grouped the question items by variance. In that way, items under the same factor can be interpreted as measuring a similar concept. Factor I can be interpreted as the Service Quality construct. Factor II encompasses the item on Information and referrals to other agencies and Factor III can be seen as representing the outcomes of the service. This analysis turned out to be useful to re-affirm the results obtained from the reliability analysis. A total of 5500 responses were received, processed and analyzed across surveys.

f. Some of the items included in the instrument are not necessarily part of the same construct. Some items are included to gather important programmatic information that is not necessarily related to the constructs service quality and effectiveness. For management decision making all items will be included in the survey reports created. However, only items within the same construct and with the same scale were considered for the Alpha Cronbach reliability analysis.

g. After the pilot test, modifications to the survey were performed particularly in the wording of certain items to increase readability. Certain items were eliminated to increase reliability.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Construct</th>
<th>Item</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service Quality</td>
<td>Reliability</td>
<td>7. I would contact the Agency in the future for disaster recovery assistance</td>
<td>5-point Likert scale Definitely would to Definitely Would not</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.1 Ease of contacting Agency</td>
<td>6-point Likert scale Poor to Excellent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.4 Timeliness of assistance</td>
<td>Ordinal – 5 points 2 hours or less to more than 48 hours</td>
</tr>
<tr>
<td>Timeliness</td>
<td></td>
<td>1. When did the Agency first get in touch with you after your disaster?</td>
<td>6-point Likert scale Poor to Excellent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.4 Timeliness of assistance</td>
<td>6-point Likert scale Poor to Excellent</td>
</tr>
<tr>
<td>Responsiveness</td>
<td></td>
<td>8.1 Helpfulness of case worker</td>
<td>6 point Likert scale Poor to Excellent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. Where did you FIRST get Agency help?</td>
<td>On scene By phone At Agency office</td>
</tr>
<tr>
<td>Assurance</td>
<td></td>
<td>8.3 Agency worker's Knowledge</td>
<td>6-point Likert scale Poor to Excellent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.6 Usefulness of names and phone numbers of other groups that can help</td>
<td>6-point Likert scale Poor to Excellent</td>
</tr>
<tr>
<td>Empathy</td>
<td></td>
<td>2. The caseworker was at every moment kind and caring</td>
<td>6-point Likert scale Poor to Excellent</td>
</tr>
<tr>
<td>Customer Satisfaction</td>
<td>Item on expectations of service</td>
<td>1. Was the help you got from Agency?</td>
<td>4-point Likert scale: More than expected What you expected Less than you expected Did not know what to expect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.9 Overall satisfaction with Agency Disaster help</td>
<td>6-point Likert scale Poor to Excellent</td>
</tr>
<tr>
<td>Effectiveness (Immediate outcomes)</td>
<td>Helping Disaster victims recover and get back to their normal pre-disaster status</td>
<td>8.7 Helpfulness in beginning your recovery back to your normal life</td>
<td>6-point Likert scale Poor to Excellent</td>
</tr>
<tr>
<td></td>
<td>Meeting immediate needs</td>
<td>8.8 Effectiveness in meeting your most immediate needs</td>
<td>6-point Likert scale Poor to Excellent</td>
</tr>
<tr>
<td>Qualitative questions</td>
<td>Quality related</td>
<td>10. Anything else you want to tell us about the services you received from the agency?</td>
<td>Open ended</td>
</tr>
</tbody>
</table>

*Item numbering comes from the numbering in the actual questionnaire*
Table 4.5 Results of Internal Consistency and Reliability Analysis Of Service Quality And Effectiveness Surveys

<table>
<thead>
<tr>
<th>Survey/Construct</th>
<th>No. of valid responses in pilot</th>
<th>Items</th>
<th>Current Alpha Cronbach reliability coefficient</th>
<th>Items deleted</th>
<th>New alpha coefficient after changes</th>
<th>Factor after principal component analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Disaster Services</td>
<td>180</td>
<td>8.1, 8.2, 8.3 8.4, 8.5, 8.6</td>
<td>0.74</td>
<td>8.1</td>
<td>0.80</td>
<td>I</td>
</tr>
<tr>
<td>Service Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empathy, Reliability,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcomes:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Help in recovery back to normal life</td>
<td>8.7</td>
<td>No items deleted</td>
<td></td>
<td>One item construct</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Fulfilling victims most immediate needs</td>
<td>8.8</td>
<td>No items deleted</td>
<td></td>
<td>One item construct</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>2. First Aid/ CPR and AED training – Health &amp; Safety Services</td>
<td>1142</td>
<td>7.1, 7.2, 7.3, 74, 7.5, 7.6, 7.7</td>
<td>0.61</td>
<td>7.6 &amp; 7.7</td>
<td>0.88</td>
<td>I</td>
</tr>
<tr>
<td>Service Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confident, prepared, willing and comfortable to act during an emergency</td>
<td>8.1, 8.2, 8.3, 8.4</td>
<td>0.90</td>
<td>No items deleted</td>
<td>III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. AFES</td>
<td>272</td>
<td>7.1, 7.2, 7.3, 7.4, 7.5</td>
<td>0.82</td>
<td>No items deleted</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>Service Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduce family stress and Increase effectiveness dealing with crises</td>
<td>8.1, 8.2</td>
<td>0.77</td>
<td>No items deleted</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reunite family members</td>
<td>8.3</td>
<td>No items deleted</td>
<td></td>
<td>III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. International Tracing Services</td>
<td>62</td>
<td>3.1, 3.2, 3.3, 3.4, 3.5, 3.6</td>
<td>0.74</td>
<td>3.7</td>
<td>0.81</td>
<td>I</td>
</tr>
<tr>
<td>Service Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service Quality</td>
<td>4.1, 4.2, 4.3, 4.4, 4.5</td>
<td>0.71</td>
<td>4.5</td>
<td>0.89</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>------------------------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase knowledge of agency’s international activities, Geneva conventions, IHL and Increase compassion for people impacted by international crises.</td>
<td>5.1, 5.2, 5.3, 5.4</td>
<td>0.59</td>
<td>5.3 5.4</td>
<td>0.85</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>Increase knowledge of IHL and increased compassion</td>
<td>5.3, 5.4</td>
<td>0.51</td>
<td>No items deleted</td>
<td>I and III</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**HIV/AIDs education** 68

<table>
<thead>
<tr>
<th>Service Quality</th>
<th>4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7</th>
<th>0.70</th>
<th>No items deleted</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>4.6</td>
<td>0.71</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Materials</td>
<td>4.4</td>
<td>0.72</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Outcomes:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase in knowledge, change in attitude, change in intentions, willingness to share acquired information, acceptance of vulnerability</td>
<td>5.2, 5.3, 5.4, 5.5, 5.6, 5.7</td>
<td>0.69</td>
<td>5.5</td>
<td>0.72</td>
</tr>
</tbody>
</table>
Table 4.5. Sources of Data for the Service Quality and Effectiveness Dimensions: Disaster Services Questionnaire

<table>
<thead>
<tr>
<th>Sample</th>
<th>Service Quality Survey</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>643 clients selected randomly representing 15 field offices</td>
<td>Field offices were asked to volunteer for the pilot study. Survey was sent to at least 100 randomly selected clients per field office. Response rate per field office varies from 5 to 65 clients. All 15 field offices were selected for DEA analysis as a demonstration only.</td>
</tr>
</tbody>
</table>

| Items | 22 items | Demographic items are not included in this count |

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Service Quality (Likert scale)</th>
<th>Customer satisfaction (Likert scale)</th>
<th>Outcome indicators (Likert scale)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Reliability</td>
<td>• This researcher tailored the questionnaire.</td>
<td>• Constructs used for service quality are supported by the literature and were repeatedly submitted for review of Agency evaluation officials.</td>
</tr>
<tr>
<td></td>
<td>• Responsiveness</td>
<td></td>
<td>• Constructs represent what the organization sees as important regarding its service quality performance.</td>
</tr>
<tr>
<td></td>
<td>• Assurance</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Empathy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Problems with the instrument | 3 field offices have less than 10 responses for the Disaster Services survey. | Data collection depended on individual field offices mailing the surveys to clients (time & resource constraint). |
|                            | • Specific sampling guidelines were given to each field office | |

199
4.4.2.2. Sampling Guidelines

Field offices where provided with clear specifications as of the sample size and sampling procedures. A copy of the booklet created with that purpose is in Appendix B.

Field offices with less than 300 clients in any line of service were directed to survey all its clients (census). Field offices with more than 300 clients for any determined type of service were instructed to draw a random sample. An excerpt from the guidelines provided follow.

"Random Sample Procedure"

Field offices that serve more than 300 customers with a given service: You may decide to mail surveys to a sample of customers (rather than to all customers). If so, you will need to draw a random sample. This is the standard procedure for drawing a random sample:

1. Find out how many customers you have in total.

2. Assign a different number to each case, starting 1,2, etc. Generate random numbers with a random number program, (you can use Excel, Lotus, Access, or simply a random numbers table provided at the end of most basic statistics text books). In MS Excel you can use the function =INT(RAND()*TOTAL) being TOTAL the number of cases you have. To generate a table of random numbers, copy the function to a range of cells in the spreadsheet and use it to select your cases. Sort as many random numbers as you need. Be advised that each time you open the spreadsheet, the random numbers will be recalculated. So, if you want to keep the numbers for your records, then copy the table into a Word document. For example, if you wish to mail questionnaires to 400 customers out of a total of 2,000, the program should draw 400 random numbers (without replacement of those numbers already selected) and sort these numbers in order from lowest to highest. Copy =INT(RAND()*2000) to 400 cells and select those cases.

3. Count through the customers using the sorted list of random numbers to identify the ones that will be included in the sample.

4. Transfer customer names and addresses to the envelopes."

Field offices with a population over 50,000 were directed to mail surveys to 200-300 people for each service delivery survey. Large field offices should aim to have at least 100 usable responses for each type of survey, although more would be a better number for each fiscal year. These numbers translate into at least 800 clients surveyed across all lines of service, a number that represents a good sample from which inferences about service quality and effectiveness at the field office level can be made.
Field offices serving smaller populations (under 50,000) could use smaller samples and should aim for a minimum of 50 completed surveys of each type.

### 4.4.3 Service quality index

An overall service quality index that considers all service quality dimensions in the questionnaire was calculated. Regarding the definition and operationalization of service quality and customer satisfaction, research has suggested that service quality and customer satisfaction are distinct constructs (Hurley & Estelami, 1998; Oliver, 1997; Taylor & Baker, 1994) and that there is a causal relationship between the two (Cronin & Taylor, 1992; Gotlieb, Grewal & Brown, 1994; Spreng and Mackoy, 1996). Unfortunately, as Hurley and Estelami (1998) point out, an examination of the measurement work in industry suggests that this distinction is often ignored in practice.

Two relative important areas where practice has lagged research are: (1) defining and operationalizing service quality and (2) determining the best indexes to describe service quality and customer satisfaction data (Hurley & Estelami, 1998).

De Ruyter et al. (1997) presents an integrative model of service quality and customer satisfaction. These authors examine whether service satisfaction is a super ordinate concept to quality or vice versa. Super ordinate means that one concept antecedes or encompasses the other. They determined that there is a conceptual overlap as well as distinctions between these two customer judgments and that service quality is an antecedent of service satisfaction. Furthermore, they conclude that these two constructs can be integrated and that usually, one can be a proxy of the other.

The formulation of the appropriate index weights was attempted as part of this research based on agency’s officials’ perceptions of importance of each one of the items. The weights were deemed to be equal for all inputs and output variables. The conclusion was that it was better to assume equal importance of all aspects of the operation.
4.4.3.1. Index Calculations

According to Hurley and Estelami (1998), customer satisfaction and service quality measures obtained through customer surveys invariably have skewed distributions (Peterson & Wilson, 1992) either to the excellent or poor side. As such, researchers have questioned the appropriateness of the generalized approach of using the mean rating to summarize the data. The calculated mean of all responses is generally used as a service quality index for that service.

These researchers performed two independent studies in which they examined the validity of the various indexes that can be used to summarize consumers’ service quality ratings (e.g. mean, median, mode, kurtosis, skewness, top/bottom tail percentiles). In both studies, both the mean and top-box percentiles were found to be the best indicators of service quality. Furthermore, in one of these studies these authors performed a simulation varying such factors as the underlying distribution of customer ratings and the strength of the relationships between customer ratings and business performance measures.

In general, the field office's service quality index will be calculated as follows:

\[ SQ_j = \frac{\sum_{i=1}^{N} I_i}{N} \]

The following Table 4.6 illustrates how an index for service quality would be calculated for the Disaster Services survey.
Table 4.6 Example of the Calculation of Service Quality Index

<table>
<thead>
<tr>
<th>Construct</th>
<th>Item</th>
<th>Answer</th>
<th>Numeric value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responsiveness</td>
<td>Ease of contacting the Agency</td>
<td>Excellent</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Helpfulness of Agency Worker</td>
<td>Very Good</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Convenience of Assistance</td>
<td>Good</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Where did you FIRST get Agency help?</td>
<td>On scene</td>
<td>3</td>
</tr>
<tr>
<td>Reliability</td>
<td>Timeliness of assistance</td>
<td>Very Good</td>
<td>5</td>
</tr>
<tr>
<td>Timeliness</td>
<td>When did the Agency first get in touch with you?</td>
<td>3 to 4 hours</td>
<td>5</td>
</tr>
<tr>
<td>Assurance</td>
<td>Agency worker's knowledge</td>
<td>Excellent</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Usefulness of names and phone numbers of other groups that can help</td>
<td>Poor</td>
<td>1</td>
</tr>
<tr>
<td>Customer Satisfaction</td>
<td>Overall satisfaction with Agency disaster help</td>
<td>Satisfied</td>
<td>4</td>
</tr>
<tr>
<td>Service quality index assigned by individual respondent</td>
<td>= Sum of all numeric values of the respondent’s answers</td>
<td>In this case =</td>
<td>39</td>
</tr>
<tr>
<td>Total service quality index for the field office</td>
<td>The mean of all the individual indices collected for that field office</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Chapter 5 presents the application of the DEA models to two data sets. First, as an illustration to the Disaster Services line of service at the field office level and second to the whole field office operation encompassing all four lines of service. For this purpose the data set will include 32 agency’s field offices.
CHAPTER 5 Application Results and Discussion

This chapter reports on the application of the two DEA models devised in Chapter 3. Each of these models — Model I and Model II — is characterized by the following:

- Model I uses the real values from the output of node 1 as input to node 2 and subsequently the inputs from node 2 as inputs to node 3, etc.

- Model II uses the output targets or projections to the frontier of inefficient units from node 1 as inputs for node 2, and so forth.

The objective of this chapter is to test, validate, and illustrate the use of the formulations with real life data. Two data sets were compiled to illustrate the behavior of each of these models:

The first data set came from a sample of field offices that participated in a pilot test for the survey instruments. The main objective of this pilot was to test the surveys, and collect the service quality and outcome data to run the DEA formulations in nodes 3 and 4 to validate its results and usefulness with key decision-makers.

For the first data set, — Disaster Services— one of the lines of services or divisions of the agency was chosen for the application of the survey instrument. Six hundred and forty three disaster victims served by 15 field offices completed questionnaires for assessing service quality and effectiveness. In addition, operational and environmental data related to the operations of the disaster services departments in those 15 field offices was gathered. The service quality and effectiveness indicators collected for each one of the 15 field offices was matched with their respective operational data prior to running both DEA models (I and II).

For the second data set, input and output indicators representing the complete field office operation for a group of 32 field offices were obtained. These 32
field offices were classified by the agency based on the fact that these field offices were serving populations ranging from 750 thousand to 1.25 million inhabitants. Data for fund raising and financial indicators, capacity creation and service delivery in all four lines of service (Disaster, Armed Forces Emergency Services, Health & Safety and Community Services and International Services) was also reviewed. Service quality and effectiveness indicators were randomly generated for illustrative purposes, since most field offices in that population range did not participate in the pilot and have not yet conducted their customer survey process. This data set was only used with Model I.

Sections 5.1 and 5.2 present a discussion for Models I and II regarding the results obtained for the first data set. Section 5.3 presents the application of model I to the total field office operations (second data set). Section 5.4 presents a discussion related to the aggregated value upon the inclusion of environmental variables in the models and section 5.5 presents the conclusions about model behavior and performance. Finally, section 5.6 presents a draft of potential DEA reports that would translate to decision-makers the results for each field unit into action driven recommendations for performance improvement.

5.1 Application of Model I in Disaster Services: Using Observed Values of the Variables in Each Node

The formulation for the four-stage model using observed values of the variables in each node was applied to a sample of 15 field offices that piloted the Disaster Services service quality and effectiveness surveys.

Since this is a limited sample, this application will only be used for the purpose of illustrating the mechanism and potential results with real data. Field offices in the sample come from all regions of the United States and represent different levels of service delivery. It is worth noting that in a more systematic and rigorous application, field offices would be evaluated in distinctively different peer groups considering other more comparable field offices and not necessarily together as it is suggested in the present illustration.
DEA works best when the number of DMUs is significantly larger than the number of variables. Therefore, the discriminatory power of the model in some nodes was greatly reduced. Also, the service quality index was calculated for each DMU as described in Chapter 4.

The following sections present the results for each one of the four nodes. The formulations are used with the data in its current form (volume) and also standardized as ratios per household (by dividing the variables by population factors or number of households). Results using both types of data are shown and comparisons drawn.

Standardized data was used because one of the objectives of nonprofit agencies, when implementing performance measurement systems, is to establish performance standards applicable to all DMUs (field offices or programs). Because of the human/social nature of the services provided, it makes sense in this field of application to have measures that represent the penetration of their services and the amount of resources obtained from the population served. Hence the use of ratios “per household.”

The overall characteristics of the 15 DMUs in the sample are shown in Table 5.1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5132</td>
<td>0.558</td>
<td>742,355</td>
<td>14,722,932</td>
<td>102,837</td>
</tr>
<tr>
<td>5180</td>
<td>0.655</td>
<td>5,737,588</td>
<td>121,962,619</td>
<td>1,698,041</td>
</tr>
<tr>
<td>5264</td>
<td>0.395</td>
<td>2,810,321</td>
<td>70,424,969</td>
<td>1,541,012</td>
</tr>
<tr>
<td>5503</td>
<td>0.455</td>
<td>4,475,547</td>
<td>129,404,072</td>
<td>2,404,178</td>
</tr>
<tr>
<td>10060</td>
<td>0.688</td>
<td>2,278,320</td>
<td>41,682,798</td>
<td>1,135,074</td>
</tr>
<tr>
<td>12024</td>
<td>0.101</td>
<td>1,251,752</td>
<td>10,247,825</td>
<td>256,538</td>
</tr>
<tr>
<td>15080</td>
<td>0.083</td>
<td>463,430</td>
<td>10,621,037</td>
<td>258,027</td>
</tr>
<tr>
<td>22172</td>
<td>0.120</td>
<td>342,220</td>
<td>7,202,271</td>
<td>177,780</td>
</tr>
<tr>
<td>22322</td>
<td>0.296</td>
<td>4,081,169</td>
<td>97,514,301</td>
<td>2,055,642</td>
</tr>
<tr>
<td>25150</td>
<td>0.171</td>
<td>1,448,301</td>
<td>34,804,630</td>
<td>786,095</td>
</tr>
<tr>
<td>29042</td>
<td>0.014</td>
<td>133,082</td>
<td>2,269,208</td>
<td>70,152</td>
</tr>
<tr>
<td>35072</td>
<td>0.154</td>
<td>1,540,943</td>
<td>34,996,065</td>
<td>778,808</td>
</tr>
<tr>
<td>35168</td>
<td>0.295</td>
<td>1,454,014</td>
<td>33,562,930</td>
<td>710,111</td>
</tr>
<tr>
<td>38360</td>
<td>0.291</td>
<td>3,211,087</td>
<td>80,765,331</td>
<td>1,727,033</td>
</tr>
<tr>
<td>45010</td>
<td>0.023</td>
<td>99,491</td>
<td>1,843,303</td>
<td>52,226</td>
</tr>
</tbody>
</table>
5.1.1 Results of Node 1: Financial Frontier

In Node 1, the total fund raising effort was expressed through the variables fund raising expenses as input and revenue as output. The environmental variables considered here were the aggregated household income —as the indicator of environmental harshness or “Ruggiero variable” (Ruggiero, 1998) — since it is considered that the level of wealth in the community is a determinant of the fund raising ability of nonprofits. The lesser the wealth in the community, the harsher the environment for this node. Another environmental factor that is considered influential is the percentage of minority in the community. Some cultures and/or ethnic groups have a stronger voluntarism and giving tradition than others. In general, Hispanic and Asian groups among other minorities might not be as inclined to give to national social organizations as other groups. Since this last assumption is based only in anecdotal experience, I decided to include minority as the “Banker and Morey variable” to assure that the weighted average of the peers equals the percentage minority in the community. With this last step I ensure that peers are comparable also at this level.

5.1.1.1. Node 1 with Volume Data

Table 5.2 below presents the results obtained for this node when volume data (data in its original form without standardization) was used. Because it is an output increasing model the best-practice score is equal to the number of outputs in the variable set. This example has one output hence best-performing units will have a score of 1 and under-performing units will have a score greater than 1.

Only 2 DMUs turned out to be inefficient, mainly because of the number of constraints and equalities in the formulation and the reduced sample size. Both inefficient DMUs are large field offices with high levels of wealth present in their communities and potentially more income generation opportunities. They also have a minority percentage of 16% and 30% respectively.

Regarding the adequacy of the peers selected by the model, for field office 35072 peers include field offices 25150 and 29042. Field office 25150 has a very much comparable community composition with almost the same level of aggregated household
income and minority population. Moreover, field office 25150 has spent approximately half of its funding to fundraise almost twice as much as field office 35072. Field office 29042 is a DMU with an almost negligible minority population but also with a community wealth seventeen times lower than that for field office 35072, therefore, operates in a much disadvantaged environment. Nonetheless, it has fundraised approximately one quarter of what 35072 has been able to raise while spending less than 5% of what field office 35072 has spent in its fundraising effort. Thus, to become efficient DMU 35072 needs to spend $440,446.7 less, and raise 73% more, i.e. 1,806,490.23.

Similarly, the other DMU that is inefficient — field office 22322 should raise 162% beyond its present amount for its level of expenses. Its peers are all DMUs operating in less wealthy environments that are proportionally doing better in their fundraising effort. The level of minority in their populations varies from 60% to 2%, but their weighted averages are equal to the level of the inefficient DMU. However, this is under the assumption that in theory a virtual DMU can exist with the same levels of minority and wealth while producing a higher level of output.

5.1.1.2. Node 1 with Standardized Data

Input and output variables were standardized classifying them by number of households in the community. This standardization makes sense from the managerial point of view because the agency is trying to come up with standards of performance for field offices that can be applied independent of their size or populations. However, the DEA application to this type of data requires that the DMUs be grouped in similar categories of field offices presenting the similar level of operations. Failing to do so will make the comparisons and peer groups less likely to be accepted by the field office managers. This is because DEA will select peers whose weighted averages of inputs and outputs equal the virtual DMU, and when data is standardized, the notion of big DMUs having a reference set (peer group) closer to their scale of operations is lost. Therefore, peers for a field office can come from any place in the sample, independent of their volume of operations. Table 5.3 presents the results for Node 1 for these standardized variables.
The number of inefficient DMUs increased from two to four. The two DMUs that were deemed inefficient using volume data continued to be inefficient here but their inefficiencies increased to over 200%. Other two DMUs that were deemed efficient (best-performers) in the previous case are now inefficient. One of them only by 9% and the other one by 30%. DMU 12024, for example, has 10% of minority and 10 million aggregated household income. It needs both, to reduce its expenses and to increase its fundraising. DMU 12024 peers are one DMU with much less wealth in its community (2.3 million dollars of aggregated household income). This result is consistent with the premises of our model, that DMUs can be compared to those in the same or worst operating conditions. This DMU also has a 1.4% minority. The second peer has 12% minority and 7.2 million dollars of aggregated household income, which is very comparable to the DMU under analysis. In general, the number of inefficient DMUs increases and the efficiency scores are also reduced when standardized data were used as opposed to volume data.

Standardizing the data before implementing the model makes a lot of sense in cases where there is wide variety of dispersion in the operation of the various DMUs as it is in this case, where field offices in the sample have a population varying between less than a 100,000 and more than a 5.5 million. By way of standardization, one is able eliminate the wide variation of the scale of operations and is able to compare these field offices meaningfully, especially in a case like this with a limited number of DMUs.

In the next pages, the tables on which the above discussions are based will be presented.
### Table 5.2 Results with Volume Data for Node 1

<table>
<thead>
<tr>
<th>DMU</th>
<th>Performance Score: Phi</th>
<th>Peers</th>
<th>Slack Fund Raising Exp</th>
<th>Fund Raising expense Actual</th>
<th>Fund Raising exp Target (actual – slack)</th>
<th>Fund raising revenue (Donations)</th>
<th>Fund raising revenue (Donations) Projected</th>
</tr>
</thead>
<tbody>
<tr>
<td>45010</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$66,901</td>
<td>$183,513.00</td>
<td>$183,513.00</td>
<td></td>
</tr>
<tr>
<td>29042</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$34,097</td>
<td>$270,419.00</td>
<td>$270,419.00</td>
<td></td>
</tr>
<tr>
<td>22172</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$134,954</td>
<td>$507,013.00</td>
<td>$507,013.00</td>
<td></td>
</tr>
<tr>
<td>12024</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$348,803</td>
<td>$713,950.00</td>
<td>$713,950.00</td>
<td></td>
</tr>
<tr>
<td>15080</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$123,067</td>
<td>$755,368.00</td>
<td>$755,368.00</td>
<td></td>
</tr>
<tr>
<td>5132</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$186,613</td>
<td>$513,425.00</td>
<td>$513,425.00</td>
<td></td>
</tr>
<tr>
<td>35168</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$508,955</td>
<td>$2,285,015.00</td>
<td>$2,285,015.00</td>
<td></td>
</tr>
<tr>
<td>25150</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$406,969</td>
<td>$1,993,832.00</td>
<td>$1,993,832.00</td>
<td></td>
</tr>
<tr>
<td>35072</td>
<td>1.726605</td>
<td>25150, 29042</td>
<td>$440,446.70</td>
<td>$806,883</td>
<td>$366,436.00</td>
<td>$1,046,267.00</td>
<td>$1,806,490.23</td>
</tr>
<tr>
<td>10060</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$398,730</td>
<td>$3,910,004.00</td>
<td>$3,910,004.00</td>
<td></td>
</tr>
<tr>
<td>5264</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$1,022,436</td>
<td>$4,463,398.00</td>
<td>$4,463,398.00</td>
<td></td>
</tr>
<tr>
<td>38360</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$1,168,798</td>
<td>$7,170,390.00</td>
<td>$7,170,390.00</td>
<td></td>
</tr>
<tr>
<td>22322</td>
<td>2.617358</td>
<td>29042, 10060, 38360</td>
<td>0</td>
<td>$1,049,960</td>
<td>$2,494,295.00</td>
<td>$2,494,295.00</td>
<td>$6,528,464.11</td>
</tr>
<tr>
<td>5180</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$2,039,211</td>
<td>$7,812,053.00</td>
<td>$7,812,053.00</td>
<td></td>
</tr>
<tr>
<td>5503</td>
<td>1</td>
<td></td>
<td>0</td>
<td>$1,379,087</td>
<td>$10,613,322.00</td>
<td>$10,613,322.00</td>
<td></td>
</tr>
</tbody>
</table>
# Table 5.3 Results with Standardized Data for Node 1

<table>
<thead>
<tr>
<th>DMU</th>
<th>Performance Score: Phi</th>
<th>Peers</th>
<th>Slack Fund Raising Exp</th>
<th>Fund Raising expense Actual</th>
<th>Fund Raising exp Target (actual – slack)</th>
<th>Fund raising revenue (Donations)</th>
<th>Fund raising revenue (Donations) Projected</th>
</tr>
</thead>
<tbody>
<tr>
<td>45010</td>
<td>1.00</td>
<td></td>
<td>0</td>
<td>$1.28</td>
<td>$1.28</td>
<td>$3.51</td>
<td>$3.51</td>
</tr>
<tr>
<td>29042</td>
<td>1.00</td>
<td></td>
<td>0</td>
<td>$0.49</td>
<td>$0.49</td>
<td>$3.85</td>
<td>$3.85</td>
</tr>
<tr>
<td>22172</td>
<td>1.00</td>
<td></td>
<td>0</td>
<td>$0.76</td>
<td>$0.76</td>
<td>$2.85</td>
<td>$2.85</td>
</tr>
<tr>
<td><strong>12024</strong></td>
<td><strong>1.09</strong></td>
<td>29042, 22172</td>
<td>$0.65</td>
<td>$1.36</td>
<td>$0.71</td>
<td>$2.78</td>
<td><strong>3.03</strong></td>
</tr>
<tr>
<td>15080</td>
<td>1.00</td>
<td></td>
<td>0</td>
<td>$0.48</td>
<td>$0.48</td>
<td>$2.93</td>
<td><strong>2.93</strong></td>
</tr>
<tr>
<td>5132</td>
<td>1.00</td>
<td></td>
<td>0</td>
<td>$1.81</td>
<td>$1.81</td>
<td>$4.99</td>
<td><strong>4.99</strong></td>
</tr>
<tr>
<td>35168</td>
<td>1.00</td>
<td></td>
<td>0</td>
<td>$0.72</td>
<td>$0.72</td>
<td>$3.22</td>
<td><strong>3.22</strong></td>
</tr>
<tr>
<td>25150</td>
<td>1.00</td>
<td></td>
<td>0</td>
<td>$0.52</td>
<td>$0.52</td>
<td>$2.54</td>
<td><strong>2.54</strong></td>
</tr>
<tr>
<td><strong>35072</strong></td>
<td><strong>3.09</strong></td>
<td>29042, 5132</td>
<td>$0.21</td>
<td>$1.04</td>
<td>$0.83</td>
<td>$1.34</td>
<td><strong>4.15</strong></td>
</tr>
<tr>
<td>10060</td>
<td>1.00</td>
<td></td>
<td>0</td>
<td>$0.35</td>
<td>$0.35</td>
<td>$3.44</td>
<td><strong>3.44</strong></td>
</tr>
<tr>
<td><strong>5264</strong></td>
<td><strong>1.34</strong></td>
<td>29042, 5132, 10060</td>
<td>0</td>
<td>$0.66</td>
<td>$0.66</td>
<td>$2.90</td>
<td><strong>3.88</strong></td>
</tr>
<tr>
<td>38360</td>
<td>1.00</td>
<td></td>
<td>0</td>
<td>$0.68</td>
<td>$0.68</td>
<td>$4.15</td>
<td><strong>4.15</strong></td>
</tr>
<tr>
<td><strong>22322</strong></td>
<td><strong>3.16</strong></td>
<td>29042, 10060, 38360</td>
<td>0</td>
<td><strong>$0.51</strong></td>
<td><strong>$0.51</strong></td>
<td>$1.21</td>
<td><strong>3.84</strong></td>
</tr>
<tr>
<td>5180</td>
<td>1.00</td>
<td></td>
<td>0</td>
<td>$1.20</td>
<td>$1.20</td>
<td>$4.60</td>
<td><strong>4.60</strong></td>
</tr>
<tr>
<td>5503</td>
<td>1.00</td>
<td></td>
<td>0</td>
<td>$0.57</td>
<td>$0.57</td>
<td>$4.41</td>
<td><strong>4.41</strong></td>
</tr>
</tbody>
</table>
5.1.2 Results of Node 2: Capacity Building Frontier

In node 2, the capacity building effort for the Disaster Services line at each field office was expressed by having the donations from node 1 as inputs to node 2 and the capabilities created being the employees, volunteers and the DSV personnel available as its output. "days of net assets" was the variable chosen as a proxy of financial capacity. Population was the environmental variable chosen as the variable for environmental harshness “Ruggiero variable” assuming that the smaller the community more difficult is it to acquire resources. Minority was again the “Banker and Morey variable” selected assuming it affected capacity building. Again, the selection of environmental variables in some of the nodes was done exclusively for the sake of illustrating the mechanisms used by the model to achieve the results and recommendations provided. It is probable that other external variables for which there are not available data at this time would be more appropriate for a more rigorous implementation, such variables may have included for example the educational attainment in the community's workforce.

Because this node has four outputs, the performance score is 4 for DMUs considered best-performers and greater than 4 for DMUs under-performing.

5.1.2.1. Results of Node 2 with Volume Data

Table 5.4 shows the performance scores and peers with volume data for this node. Two DMUs were deemed under-performing. Projections or targets for these units are shown in Table 5.5.
### Table 5.4 Performance Score and Peers for Node 2 with Volume Data

<table>
<thead>
<tr>
<th>DMU</th>
<th>Performance Score: Phi</th>
<th>Peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>45010</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>29042</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>22172</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>15080</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5132</td>
<td>4.75</td>
<td>45010, 22172</td>
</tr>
<tr>
<td>12024</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>25150</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>35168</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>35072</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>10060</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5264</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>38360</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>22322</td>
<td>8.22</td>
<td>35072, 10060, 5264</td>
</tr>
<tr>
<td>5503</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5180</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

The peers for DMU 5132 both operate in smaller communities, which represent in this case harsher environments. However, the Banker and Morey environmental variable (% minority in this case) selects a composite reference set with fewer minorities for both peers (45010 with 2% and 22172 with 12%). In the case of the peers for 22322, which is a field office serving over 4.4 million people (of which 30% are minority), all three of the selected peers operate in smaller communities but relatively comparable in size to 22322. The smaller peer office (35072) serves a community of 1.5 million and the largest (5264) a community of 2.8 million. Regarding minority the composite reference set shows one community with an small minority (15%) and the two other peers serving a more diverse community (39.5 and 69%). In this case, the peer group definitely corresponds to what was expected.
Table 5.5 Projections and Actual Values for Inefficient DMUs in Node 2

<table>
<thead>
<tr>
<th>DMU</th>
<th>Phi 1</th>
<th>Phi 2</th>
<th>Phi 3</th>
<th>Phi 4</th>
<th>5132</th>
<th>22322</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual DSV Enrollees</td>
<td>1.354112</td>
<td>1.069394</td>
<td>1.18308</td>
<td>1.14027</td>
<td>49</td>
<td>311</td>
</tr>
<tr>
<td>Target DSV Enrollees</td>
<td>66.35149</td>
<td>3.997728</td>
<td>2.189243</td>
<td>1.037005</td>
<td>80</td>
<td>292</td>
</tr>
<tr>
<td>Actual Disaster registered volunteers</td>
<td>332.5814</td>
<td>332.5814</td>
<td>1167.337</td>
<td>1167.337</td>
<td>80</td>
<td>292</td>
</tr>
<tr>
<td>Target Disaster registered volunteers</td>
<td>3.54924</td>
<td>15.3247</td>
<td>3.54924</td>
<td>15.3247</td>
<td>80</td>
<td>292</td>
</tr>
<tr>
<td>Actual DAYS OF NET ASSETS</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>825</td>
<td>825</td>
</tr>
<tr>
<td>Target DAYS OF NET ASSETS</td>
<td>912.2157</td>
<td>912.2157</td>
<td>855.5288</td>
<td>855.5288</td>
<td>825</td>
<td>825</td>
</tr>
</tbody>
</table>

Table 5.5 shows the different projections "Phi" for each type of variable. Being a non-radial model, it gives different factors to reflect the increase for each of the variables. One needs to remember that DEA assumes substitution, so that it is then possible that an inefficient DMU is efficient in one dimension or variable and very inefficient in another. This is the case of 22322 for the number of DSV enrollees, which is an efficient dimension, while the number of disaster volunteers is almost 300% inefficient.

5.1.2.2. **Results for Node 2 with Standardized Data**

I standardized the data by dividing the discretionary input and output variables by households (ten thousand households (hh) in order to have higher order values for better intelligibility). Table 5.6 shows these results. Six DMUs were deemed inefficient using these data as compared to two DMUs in the volume data case. One of the DMUs deemed inefficient in the previous application turned out to be efficient here. The other one remained inefficient. Most DMUs operating in large communities were deemed as non-performing, which is a notion validated by decision-makers at the agency. Large field offices collect more money but at the same time it is more difficult for them to build capacity and service delivery for adequately tending to their larger communities. While volume data may not account for this problem, the use of standardize data will. Capacity in this line of service is not only a measure of volume but also
of preparedness to deal with the unexpected. Indicators per household are more likely to capture inefficiencies in this respect than are volume data not only because of the scale of the data but because it will make more difficult in some cases to find peers. By standardizing, the DMUs are also brought closer to one another in the solution space.

<table>
<thead>
<tr>
<th>DMU</th>
<th>Overall Performance Score: Phi</th>
<th>Peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>45010</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>29042</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>22172</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>15080</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5132</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>12024</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>25150</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>35168</td>
<td>13.5964325</td>
<td>45010, 22172</td>
</tr>
<tr>
<td>35072</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>10060</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5264</td>
<td>19.0845777</td>
<td>45010, 22172, 25150</td>
</tr>
<tr>
<td>38360</td>
<td>25.5357921</td>
<td>45010, 22172, 5132</td>
</tr>
<tr>
<td>22322</td>
<td>24.4555949</td>
<td>45010, 22172, 25150</td>
</tr>
<tr>
<td>5503</td>
<td>11.7808903</td>
<td>45010, 22172, 25150</td>
</tr>
<tr>
<td>5180</td>
<td>11.1198599</td>
<td>45010, 22172</td>
</tr>
</tbody>
</table>

In table 5.6, all peers of the inefficient DMUs operate in smaller communities as expected. However, as Ruggiero and others have argued, the percentage minority variable inside the model (following the Banker and Morey formulation) does not guarantee that the reference set will be at the same or worst level and only that the virtual DMU will be the convex combination of this variable. As such, peers have a variety of minority ratios, from lower than the one in the DMU being evaluated to comparable, to higher, as is the case of DMUs 45010, 22172 and 25150 which are the peers for DMU 5264. All three have less population, so harsher environment for capacity creation is assumed. They present 2.3% and 17.1% minority, as compared to approximately 40% of DMU 5264.
Field Office 5264 is so inefficient as compared to 45010, for example, because this DMU has 9 DSV enrollees per 10 thousand households, while 5264 has only 1 DSV per 10 thousand households. The later is evident given that it has 200 times more population from where to get its enrollees. Also, regarding total registered volunteers, 5264 has only 6 per 10 thousand households, when it should be having 9 as compared to 45010, which has 10 per 10 thousand households. A similar situation occurs with paid staff, where 5264 presents 0.04 employees per 10 thousand households while 45010 presents 0.38. The days of net assets dimension is efficient for both DMUs 5264 and 45010 in this example.

Table 5.7 below shows the projections for the inefficient DMUs.

Let's analyze for example, DMU 35168, which has a population of 1.5 million of which 30% are minorities. Its peers are 45010 and 22172. Both are smaller communities with 2% and 12% minority respectively. However, 45010 has 9 DSV enrollees per 10K hh while 35168 has only 4.5; 10 disaster volunteers while 45010 has 5 and should have 27.5 according to the level of donations this office receives. 45010 has 0.38 paid staff per 10K hh and 251 days of net assets compare to 471 while based on the level of inputs 35168 should have 1263 days of net assets. A similar situation occurs for all other inefficient DMUs. In fact, DMUs 45010 and 22172 dominate the other observations as they are peers for all the under-performing field offices.
### Table 5.7. Projections for Inefficient DMUs Node 2

<table>
<thead>
<tr>
<th>DMU</th>
<th>Actual DSV Enrollees</th>
<th>Target DSV Enrollees per 10K HH</th>
<th>Actual Disaster Registered Volunteers per 10K HH</th>
<th>Target Disaster Registered Volunteers per 10KHH</th>
<th>Actual Disaster Total Paid Staff per 10KHH</th>
<th>Target Disaster Total Paid Staff per 10KHH</th>
<th>Actual DAYS OF NET ASSETS</th>
<th>Target DAYS OF NET ASSETS</th>
</tr>
</thead>
<tbody>
<tr>
<td>45010</td>
<td>9</td>
<td>9.3823</td>
<td>10</td>
<td>9.573776</td>
<td>0.38</td>
<td>0.38</td>
<td>251</td>
<td>251</td>
</tr>
<tr>
<td>29042</td>
<td>2</td>
<td>1.568024</td>
<td>18</td>
<td>18.10355</td>
<td>0.07</td>
<td>0.07</td>
<td>160</td>
<td>160</td>
</tr>
<tr>
<td>22172</td>
<td>4</td>
<td>4.331196</td>
<td>28</td>
<td>28.46214</td>
<td>0.25</td>
<td>0.25</td>
<td>1318</td>
<td>1318</td>
</tr>
<tr>
<td>15080</td>
<td>1</td>
<td>1.162669</td>
<td>5</td>
<td>5.076988</td>
<td>0.09</td>
<td>0.09</td>
<td>1094</td>
<td>1094</td>
</tr>
<tr>
<td>5132</td>
<td>5</td>
<td>4.764822</td>
<td>30</td>
<td>30.24203</td>
<td>0.29</td>
<td>0.29</td>
<td>800</td>
<td>800</td>
</tr>
<tr>
<td>12024</td>
<td>2</td>
<td>1.949029</td>
<td>24</td>
<td>24.2849</td>
<td>0.31</td>
<td>0.31</td>
<td>244</td>
<td>244</td>
</tr>
<tr>
<td>25150</td>
<td>2</td>
<td>1.513812</td>
<td>5</td>
<td>4.808579</td>
<td>0.24</td>
<td>0.24</td>
<td>790</td>
<td>790</td>
</tr>
<tr>
<td>35168</td>
<td>2</td>
<td>4.590747</td>
<td>5</td>
<td>27.49156</td>
<td>0.11</td>
<td>0.26</td>
<td>471</td>
<td>1263.2</td>
</tr>
<tr>
<td>35072</td>
<td>1</td>
<td>1.001531</td>
<td>14</td>
<td>13.73894</td>
<td>0.21</td>
<td>0.21</td>
<td>935</td>
<td>935</td>
</tr>
<tr>
<td>10060</td>
<td>1</td>
<td>0.81052</td>
<td>17</td>
<td>16.80067</td>
<td>0.10</td>
<td>0.10</td>
<td>297</td>
<td>297</td>
</tr>
<tr>
<td>5264</td>
<td>1</td>
<td>6.202447</td>
<td>6</td>
<td>8.718831</td>
<td>0.04</td>
<td>0.32</td>
<td>504</td>
<td>504</td>
</tr>
<tr>
<td>38360</td>
<td>1</td>
<td>8.650513</td>
<td>4</td>
<td>12.5622</td>
<td>0.09</td>
<td>0.37</td>
<td>374</td>
<td>374</td>
</tr>
<tr>
<td>22322</td>
<td>0</td>
<td>3.232148</td>
<td>1</td>
<td>10.4977</td>
<td>0.03</td>
<td>0.26</td>
<td>825</td>
<td>825</td>
</tr>
<tr>
<td>5503</td>
<td>2</td>
<td>5.504737</td>
<td>3</td>
<td>10.06811</td>
<td>0.07</td>
<td>0.31</td>
<td>610</td>
<td>610</td>
</tr>
<tr>
<td>5180</td>
<td>2</td>
<td>6.94745</td>
<td>7</td>
<td>18.67878</td>
<td>0.18</td>
<td>0.32</td>
<td>376</td>
<td>765.3</td>
</tr>
</tbody>
</table>
The problem encountered with this model is that the selection of the DMUs to enter the frontier according to Ruggiero is very rigid. In that case, DMUs in the harshest environments cannot be compared to any other in a better environment, even if the difference is just slightly better. This problem worsens with a sample of DMUs so reduced as the ones available for this study. See for example DMU 29042, which has a much reduced level of Days of Net Assets and nonetheless turned out to be efficient because its peer group is composed of only one DMU.

In cases like this node, where one has a relatively large number of input and output variables the analysis is restricted in terms of the limited number of DMUs at hand. Thus, a constraint like the one based on the Ruggiero variable allows a lot of the DMUs with harsher environments to be efficient by default even without a chance of being compared to others.

However, if no Ruggiero harshness variable is used, then DEA will select peers across the entire sample, and DMUs operating in much better environments will be compared to units operating in disadvantage, hence invalidating the notion of “comparable” peers.

Analytically, this node is exactly the same as node 1 since it has the same formulation structure. However, conceptually speaking, monetary resources are on the input side being used to create capacity in the form of human resources and financial stability on the output side. There is little in the nonprofit literature addressing the evaluation of this transformation of monetary resources into operational capacity.

In the next section, I look at evaluating the transformation of the capacity created in this node to deliver services.

### 5.1.3 Results for Node 3: Service Delivery

In the service delivery node, the capabilities created in Node 2 become the input expressed as the resources needed to produce services for the node’s output. The variable number of families assisted during disasters was selected for this demonstration as output. Other outputs of this node include service quality characteristics such as an index of perceived service quality
and an operational quality measure characterized by the percentage of times the field office was 
on the disaster scene in less than two hours.

It has been suggested that the size of the population is an important factor in 
service delivery. This is the Ruggiero factor of environmental harshness included in the model. 
However, this factor is included in a different form as opposed to node 2. In node 2, larger places 
have better opportunities to acquire capacity in number and quality such as better-trained 
personnel. In node 3, the larger the place, the more difficult is it to deliver services in a timely and 
consistent manner. This is because for service delivery, the larger the unit, the more difficult 
service delivery becomes.

The performance score for this node is 3 if the DMU is a high-performing unit and 
greater than 3 if the DMU is under-performing.

5.1.3.1. Results with Volume Data for Node 3

Table 5.8 presents the performance score for this node. All the DMUs in this node 
are efficient. The discriminatory power of DEA is greatly reduced because of the presence of six 
variables (3 inputs and 3 outputs) in this node. No peers are determined since all DMUs are 
deemed efficient. Table 5.9 presents the desirable level of inputs or resources. This drawback can 
be easily observed by looking up the case of field office 5132. This field office has zero % 
Timeliness and still came out to be efficient. This is further analyzed below by further studying 
the performance of this field office.
The reason for the overestimation of performance in this case is probably due to the fact that field office 5132 has the least number of DSV Enrollees (49) among the field offices to which it was allowed to be compared with. The only other field office which also had the least number of DSV Enrollees (49) was field office 45010. However, field office 45010 is in a much better environment with respect to the Ruggiero variable (Population) and hence could not be in its reference set for comparison.

When the standardized data results were analyzed (see Table 5.13), it was found that this field office 5132 turned out to be inefficient with field offices 35072, 12024 as its peers. Here, by using standardized data, the model was able to detect inefficiencies in all the three output variables, namely, %Timeliness, Service Quality Index and Disaster Families assisted. In the case of the second and the third output variables, the inefficiency was captured by the proportional increase required. However, this was not possible in the case of Timeliness, as the data had a zero value. But the ability of the model to determine output shortfalls or deficits also by means of slacks came in handy when a Timeliness value of 55.13 was recommended for this field office. The above illustration helps to explain that the model performs well even in the presence of zero values in the data. This is also true for in the other nodes because of similar formulation structures.
Table 5.9 Levels of Inputs for Node 3

<table>
<thead>
<tr>
<th>DMU</th>
<th>DSV Enrollees</th>
<th>Target DSV ENROLLES</th>
<th>Disaster Registered Volunteers</th>
<th>Target Disaster Registered Volunteers</th>
<th>Disaster Total Paid Staff</th>
<th>Target Disaster Total Paid Staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>5180</td>
<td>262</td>
<td>262</td>
<td>1122</td>
<td>1122</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>5503</td>
<td>410</td>
<td>410</td>
<td>758</td>
<td>758</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>22322</td>
<td>80</td>
<td>80</td>
<td>292</td>
<td>292</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>38360</td>
<td>87</td>
<td>87</td>
<td>636</td>
<td>636</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5264</td>
<td>115</td>
<td>115</td>
<td>891</td>
<td>891</td>
<td>6.04</td>
<td>6.04</td>
</tr>
<tr>
<td>10060</td>
<td>92</td>
<td>92</td>
<td>1907</td>
<td>1907</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>35072</td>
<td>78</td>
<td>78</td>
<td>1070</td>
<td>1070</td>
<td>16.01</td>
<td>16.01</td>
</tr>
<tr>
<td>35168</td>
<td>107</td>
<td>107</td>
<td>351</td>
<td>351</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>25150</td>
<td>119</td>
<td>119</td>
<td>378</td>
<td>378</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>12024</td>
<td>50</td>
<td>50</td>
<td>623</td>
<td>623</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>5132</td>
<td>49</td>
<td>49</td>
<td>311</td>
<td>311</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>15080</td>
<td>30</td>
<td>30</td>
<td>131</td>
<td>131</td>
<td>2.35</td>
<td>2.35</td>
</tr>
<tr>
<td>22172</td>
<td>77</td>
<td>77</td>
<td>506</td>
<td>506</td>
<td>4.5</td>
<td>4.5</td>
</tr>
<tr>
<td>29042</td>
<td>11</td>
<td>11</td>
<td>127</td>
<td>127</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>45010</td>
<td>49</td>
<td>49</td>
<td>50</td>
<td>50</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 5.10 Output Targets for Node 3 Volume Data

<table>
<thead>
<tr>
<th>DMU</th>
<th>Timeliness</th>
<th>Target Timeliness</th>
<th>SQ_Index</th>
<th>Target SQ_Index</th>
<th>Disaster Families Assisted</th>
<th>Target Disaster Families Assisted</th>
</tr>
</thead>
<tbody>
<tr>
<td>5180</td>
<td>40</td>
<td>40</td>
<td>29.74</td>
<td>29.74</td>
<td>574</td>
<td>574</td>
</tr>
<tr>
<td>5503</td>
<td>54</td>
<td>54</td>
<td>30.48</td>
<td>30.48</td>
<td>855</td>
<td>855</td>
</tr>
<tr>
<td>22322</td>
<td>58.00</td>
<td>58.00</td>
<td>31.09</td>
<td>31.09</td>
<td>1676</td>
<td>1676</td>
</tr>
<tr>
<td>38360</td>
<td>40.00</td>
<td>40.00</td>
<td>30.16</td>
<td>30.16</td>
<td>1078</td>
<td>1078</td>
</tr>
<tr>
<td>5264</td>
<td>55.00</td>
<td>55.00</td>
<td>31.85</td>
<td>31.85</td>
<td>248</td>
<td>248</td>
</tr>
<tr>
<td>10060</td>
<td>59.00</td>
<td>59.00</td>
<td>31.68</td>
<td>31.68</td>
<td>246</td>
<td>246</td>
</tr>
<tr>
<td>35072</td>
<td>60</td>
<td>60</td>
<td>31.24</td>
<td>31.24</td>
<td>815</td>
<td>815</td>
</tr>
<tr>
<td>35168</td>
<td>65.00</td>
<td>65.00</td>
<td>31.78</td>
<td>31.78</td>
<td>369</td>
<td>369</td>
</tr>
<tr>
<td>25150</td>
<td>60.00</td>
<td>60.00</td>
<td>31.88</td>
<td>31.88</td>
<td>564</td>
<td>564</td>
</tr>
<tr>
<td>12024</td>
<td>35</td>
<td>35</td>
<td>20.8</td>
<td>20.8</td>
<td>162</td>
<td>162</td>
</tr>
<tr>
<td>5132</td>
<td>0.00</td>
<td>0.00</td>
<td>25.25</td>
<td>25.25</td>
<td>79</td>
<td>79</td>
</tr>
<tr>
<td>15080</td>
<td>70</td>
<td>70</td>
<td>32.77</td>
<td>32.77</td>
<td>144</td>
<td>144</td>
</tr>
<tr>
<td>22172</td>
<td>75.00</td>
<td>75.00</td>
<td>33.00</td>
<td>33.00</td>
<td>101</td>
<td>101</td>
</tr>
<tr>
<td>29042</td>
<td>40.00</td>
<td>40.00</td>
<td>24.88</td>
<td>24.88</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>45010</td>
<td>35</td>
<td>35</td>
<td>25.67</td>
<td>25.67</td>
<td>17</td>
<td>17</td>
</tr>
</tbody>
</table>
5.1.3.2. *Results with Standardized Data: Node 3*

The numbers of families assisted as well as the capabilities were standardized by dividing the raw data by the number of households in the community served by the DMU.

Table 5.11 presents the performance scores as well as the peers. There is only one inefficient DMU in node III. This is of course due to the reduced discriminatory power of the model with so few DMUs and so many inputs and outputs. This DMU is using more resources per household to deliver fewer services with lesser quality.

<table>
<thead>
<tr>
<th>DMU</th>
<th>Phi</th>
<th>Peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>5180</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>5503</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>22322</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>38360</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>5264</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>10060</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>35072</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>35168</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>25150</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>12024</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>5132</td>
<td>3.413758</td>
<td>35072, 12024</td>
</tr>
<tr>
<td>15080</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>22172</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>29042</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>45010</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.12 presents the desired level of inputs. No slacks were computed for this model so the level of inputs remained constant.
**Table 5.12 Desired Level of Inputs for Node 3 with Standardized Data**

<table>
<thead>
<tr>
<th>DMU</th>
<th>DSV per 10KHH</th>
<th>Disaster Registered Volunteers per 10KHH</th>
<th>Disaster Total Paid Staff per 10KHH</th>
</tr>
</thead>
<tbody>
<tr>
<td>5180</td>
<td>2</td>
<td>7</td>
<td>0.18</td>
</tr>
<tr>
<td>5503</td>
<td>2</td>
<td>3</td>
<td>0.07</td>
</tr>
<tr>
<td>22322</td>
<td>0</td>
<td>1</td>
<td>0.03</td>
</tr>
<tr>
<td>38360</td>
<td>1</td>
<td>4</td>
<td>0.09</td>
</tr>
<tr>
<td>5264</td>
<td>1</td>
<td>6</td>
<td>0.04</td>
</tr>
<tr>
<td>10060</td>
<td>1</td>
<td>17</td>
<td>0.10</td>
</tr>
<tr>
<td>35072</td>
<td>1</td>
<td>14</td>
<td>0.21</td>
</tr>
<tr>
<td>35168</td>
<td>2</td>
<td>5</td>
<td>0.11</td>
</tr>
<tr>
<td>25150</td>
<td>2</td>
<td>5</td>
<td>0.24</td>
</tr>
<tr>
<td>12024</td>
<td>2</td>
<td>24</td>
<td>0.31</td>
</tr>
<tr>
<td><strong>5132</strong></td>
<td><strong>5</strong></td>
<td><strong>30</strong></td>
<td><strong>0.29</strong></td>
</tr>
<tr>
<td>15080</td>
<td>1</td>
<td>5</td>
<td>0.09</td>
</tr>
<tr>
<td>22172</td>
<td>4</td>
<td>28</td>
<td>0.25</td>
</tr>
<tr>
<td>29042</td>
<td>2</td>
<td>18</td>
<td>0.07</td>
</tr>
<tr>
<td>45010</td>
<td>9</td>
<td>10</td>
<td>0.38</td>
</tr>
</tbody>
</table>

The peer group for DMU 5132 (56% minority, population 742K) are 35072 which is in a more disadvantageous environment for service delivery, with double the population to serve but less minority (15%) and 12024 which also operates in a harsher environment with 1.2 million population but presents much less minority (1%). So the significance of the peer group regarding environmental aspects is not clear.
<table>
<thead>
<tr>
<th>DMU</th>
<th>Timeliness</th>
<th>Target Timeliness</th>
<th>SQ_Index</th>
<th>Target SQ_Index</th>
<th>Disaster _Families assisted per 10K HH</th>
<th>Target Disaster _Families assisted per 10K HH</th>
</tr>
</thead>
<tbody>
<tr>
<td>5180</td>
<td>40</td>
<td>40</td>
<td>29.74</td>
<td>29.74</td>
<td>3.380366</td>
<td>3.380366</td>
</tr>
<tr>
<td>5503</td>
<td>54</td>
<td>54</td>
<td>30.48</td>
<td>30.48</td>
<td>3.556309</td>
<td>3.556309</td>
</tr>
<tr>
<td>22322</td>
<td>58</td>
<td>58</td>
<td>31.09</td>
<td>31.09</td>
<td>8.153171</td>
<td>8.153171</td>
</tr>
<tr>
<td>38360</td>
<td>40</td>
<td>40</td>
<td>30.16</td>
<td>30.16</td>
<td>6.241919</td>
<td>6.241919</td>
</tr>
<tr>
<td>5264</td>
<td>55</td>
<td>55</td>
<td>31.85</td>
<td>31.85</td>
<td>1.609332</td>
<td>1.609332</td>
</tr>
<tr>
<td>10060</td>
<td>59</td>
<td>59</td>
<td>31.68</td>
<td>31.68</td>
<td>2.16726</td>
<td>2.16726</td>
</tr>
<tr>
<td>35072</td>
<td>60</td>
<td>60</td>
<td>31.24</td>
<td>31.24</td>
<td>10.46471</td>
<td>10.46471</td>
</tr>
<tr>
<td>35168</td>
<td>65</td>
<td>65</td>
<td>31.78</td>
<td>31.78</td>
<td>5.196371</td>
<td>5.196371</td>
</tr>
<tr>
<td>25150</td>
<td>60</td>
<td>60</td>
<td>31.88</td>
<td>31.88</td>
<td>7.174705</td>
<td>7.174705</td>
</tr>
<tr>
<td>12024</td>
<td>35</td>
<td>35</td>
<td>20.8</td>
<td>20.8</td>
<td>6.314854</td>
<td>6.314854</td>
</tr>
<tr>
<td>5132</td>
<td>0</td>
<td>55.13</td>
<td>25.25</td>
<td>29.20714</td>
<td>7.68206</td>
<td>9.656656</td>
</tr>
<tr>
<td>15080</td>
<td>70</td>
<td>70</td>
<td>32.77</td>
<td>32.77</td>
<td>5.580811</td>
<td>5.580811</td>
</tr>
<tr>
<td>22172</td>
<td>75</td>
<td>75</td>
<td>33</td>
<td>33</td>
<td>5.681179</td>
<td>5.681179</td>
</tr>
<tr>
<td>29042</td>
<td>40</td>
<td>40</td>
<td>24.88</td>
<td>24.88</td>
<td>8.837952</td>
<td>8.837952</td>
</tr>
<tr>
<td>45010</td>
<td>35</td>
<td>35</td>
<td>25.67</td>
<td>25.67</td>
<td>3.255084</td>
<td>3.255084</td>
</tr>
</tbody>
</table>

Field office 5132 utilizes more resources than its peers to deliver far fewer services than DMU 35072, which performs above most regarding all output dimensions. DMU 12024, conversely, spent less in two input dimensions and produces more output in one dimension. The prescribed targets are an increase in timeliness from 0% to 55.13 % of the cases; increase the service quality index from 25.25 to 29.2 and increase the number of disaster families assisted from 7.68 per 10K households to 9.65 families per 10K households.

### 5.1.4 Node 4: Customer Effectiveness

The effectiveness node has been expressed with the service quality indicators as inputs and the desired outcomes as outputs. In this node, the number of families assisted was not included since outcomes were measured on individual clients who received only one service. Hence the number of services should not affect the outcome result. The premise is that higher
service quality on an individual will produce better or higher outcomes. The level of minority in
the population is a proxy of cultural environmental variables that can be considered as
fundamental for positive outcome generation. In general, minority populations have less
education, higher unemployment rate and more social problems than Caucasian populations.
Therefore, the percentage minority in the population served would inversely affects the impact
potential of social services.

One characteristic of DEA is that it will tend to reward DMUs that either produce
more outputs with a fixed level of inputs or that produce a fixed level of outputs with less inputs.
In the formulation for this node the DEA technique would find a reference set for which the
weighted average of the inputs would be equal to the level of inputs of the DMU under
evaluation. This means that for some cases, DMUs showing fewer inputs will become part of the
reference set of a given DMU. This is not a problem except for this node, where it is implied that
DMUs with less service quality could become part of the reference set of a DMU with higher
quality.

To explore how this issue would affect the validity of the results obtained by DEA
Node 4 is run with two types of inputs. The first run, with normal service quality levels and then a
second run with service quality indexes transformed to their inverse —meaning lack of service
quality.

Because in this node all variables are percentages or indices, there is no variation
for a standardized data model. In both approaches, the data would be the same.

5.1.4.1. Results for Node 4 with Service Quality Indexes

Table 5.14 presents the performance scores for node 4. Two field offices are
under-performing. DMU 35072 and DMU 29042. DMU 29042 operates in a very positive
environment, 1% minority, in a very small community. DMU 35072 operates in an environment
with 15% minority, but in a much larger community.

The peers for DMU 29042 are 10060, 12024, 45010 are all DMUs operating in
higher minority concentration communities. However, 45010 and 12024 are very much
comparable to 29042, since their minority ratio is 2% and 10% respectively. DMU 10060, however, operates in a much more difficult environment, with 65.5% minority and a much larger community. The peer group, therefore, is meaningful for this DMU.

Table 5.14. Performance Scores for Node 4 with Service Quality as inputs

<table>
<thead>
<tr>
<th>ECODE</th>
<th>SCORE</th>
<th>Peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>10060</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>5180</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>5132</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>5503</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>5264</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>22322</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>35168</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>38360</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>25150</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>35072</td>
<td>2.28</td>
<td>5503, 5132, 35168</td>
</tr>
<tr>
<td>22172</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>12024</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>15080</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>45010</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>29042</td>
<td>3.39</td>
<td>10060, 12024, 45010</td>
</tr>
</tbody>
</table>

Moreover, DMU 29042 has two peers that have poor service quality. In fact, DMU 12024 and DMU 45010 are the DMUs with lower service quality indicators except for 5132 that presents 0% timeliness.

Table 5.15 below presents the service quality indicators for this node. There were no slacks associated with the input side, so the actual and target inputs are the same.
Table 5.15 Service Quality Inputs for Node 4

<table>
<thead>
<tr>
<th>ECODE</th>
<th>Timeliness</th>
<th>SQ_Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>10060</td>
<td>59.00</td>
<td>31.68</td>
</tr>
<tr>
<td>5180</td>
<td>40.00</td>
<td>29.74</td>
</tr>
<tr>
<td>5132</td>
<td>0.00</td>
<td>25.25</td>
</tr>
<tr>
<td>5503</td>
<td>54.00</td>
<td>30.48</td>
</tr>
<tr>
<td>5264</td>
<td>55.00</td>
<td>31.85</td>
</tr>
<tr>
<td>22322</td>
<td>58.00</td>
<td>31.09</td>
</tr>
<tr>
<td>35168</td>
<td>65.00</td>
<td>31.78</td>
</tr>
<tr>
<td>38360</td>
<td>40.00</td>
<td>30.16</td>
</tr>
<tr>
<td>25150</td>
<td>60.00</td>
<td>31.88</td>
</tr>
<tr>
<td>35072</td>
<td>60.00</td>
<td>31.24</td>
</tr>
<tr>
<td>22172</td>
<td>75.00</td>
<td>33.00</td>
</tr>
<tr>
<td>12024</td>
<td>35.00</td>
<td>20.80</td>
</tr>
<tr>
<td>15080</td>
<td>70.00</td>
<td>32.77</td>
</tr>
<tr>
<td>45010</td>
<td>35.00</td>
<td>25.67</td>
</tr>
<tr>
<td>29042</td>
<td>40.00</td>
<td>24.88</td>
</tr>
</tbody>
</table>

The results above need to be compared in light of the outcome side of the formulation. Table 5.16 presents these results. Some issues that require further attention and discussion are highlighted in gray. DMUs 12024 and 45010 have very low performance in the outcome side. However, because their service quality levels are also low, DEA will consider them as making proportionally efficient use of their respective low level of inputs. However, the inputs for this node happen to be service quality indicators, which do not necessarily need to be reduced. Moreover, DMU 5132 has extremely low effectiveness (around 50% of outcome achievement) but two situations are taking place here. First, because one of its inputs has zero % value, then the relative expected output side is also low; therefore, 50% is not a bad result for this DMU. Furthermore, 5132 operates in a community with high minority/ harsh environment; therefore it can only be compared to two other DMUs that operate in an even harsher environment. Although, those other two DMUs perform much better, 5132 uses less input. Hence it is considered efficient by default.

The use of unorthodox inputs and all of the problems associated with them trigger several assumptions different than the common DEA assumptions. This in turn prompted the use of data transformation on the input side.
Table 5.16 Actual and Target Values of Outcomes in Node 4 with Service Quality as Input

<table>
<thead>
<tr>
<th>ECODE</th>
<th>%_Minority</th>
<th>RECOVERY OUTCOME</th>
<th>IMMEDIATE NEEDS OUTCOME</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
<td>Phi Target</td>
<td>Actual Phi Target</td>
</tr>
<tr>
<td>10060</td>
<td>68.78%</td>
<td>77.40 1.00</td>
<td>77.40 93.50 1.00</td>
</tr>
<tr>
<td>5180</td>
<td>65.51%</td>
<td>60.00 1.00</td>
<td>60.00 70.00 1.00</td>
</tr>
<tr>
<td>5132</td>
<td>55.77%</td>
<td>25.00 1.00</td>
<td>25.00 50.00 1.00</td>
</tr>
<tr>
<td>5503</td>
<td>45.51%</td>
<td>64.30 1.00</td>
<td>64.30 83.30 1.00</td>
</tr>
<tr>
<td>5264</td>
<td>39.50%</td>
<td>66.70 1.00</td>
<td>66.70 88.90 1.00</td>
</tr>
<tr>
<td>22322</td>
<td>29.56%</td>
<td>60.90 1.00</td>
<td>60.90 78.30 1.00</td>
</tr>
<tr>
<td>35168</td>
<td>29.50%</td>
<td>80.40 1.00</td>
<td>80.40 89.10 1.00</td>
</tr>
<tr>
<td>38360</td>
<td>29.15%</td>
<td>58.20 1.00</td>
<td>58.20 74.50 1.00</td>
</tr>
<tr>
<td>25150</td>
<td>17.11%</td>
<td>73.10 1.00</td>
<td>73.10 78.80 1.00</td>
</tr>
<tr>
<td>35072</td>
<td>15.40%</td>
<td>68.60 1.09</td>
<td>74.84 72.50 1.19</td>
</tr>
<tr>
<td>22172</td>
<td>12.04%</td>
<td>50.00 1.00</td>
<td>50.00 75.00 1.00</td>
</tr>
<tr>
<td>12024</td>
<td>10.12%</td>
<td>40.00 1.00</td>
<td>40.00 60.00 1.00</td>
</tr>
<tr>
<td>15080</td>
<td>8.31%</td>
<td>77.30 1.00</td>
<td>77.30 86.40 1.00</td>
</tr>
<tr>
<td>45010</td>
<td>2.33%</td>
<td>66.70 1.00</td>
<td>66.70 66.70 1.00</td>
</tr>
<tr>
<td>29042</td>
<td>1.38%</td>
<td>37.50 1.54</td>
<td>57.73 37.50 1.85</td>
</tr>
</tbody>
</table>
5.1.4.2. Results for Node 4 Using Transformed Inputs to Bad Quality Index

The data on the input side has been transformed as the inverse of quality, i.e. bad quality. This means that DEA will look for peer DMUs whose weighted average equals the level of bad quality of the composed virtual DMU. Lesser bad quality and greater level of outcomes will be the two rules that will guide DEA to select effective units. DMUs with less bad quality (equivalent to more good quality) will be rewarded in the performance calculations. These types of transformations are done in the DEA literature to deal with inputs and/or outputs that are not in the right format (see Pasupathy, 2002 for a discussion on the different approaches to treat undesirable outputs).

Table 5.17 presents the performance scores and peers for inefficient DMUs. Nine DMUs are inefficient and one DMU (10060) is the most repeated peer for most of the inefficient DMUs (dominant DMU). In fact, despite operating in a very harsh environment (minority 69%), DMU 10060 outperforms most DMUs operating in better environments in terms of quality of services and outcome achievement (efficiency). In fact, its timeliness is not the higher nor is its service quality index, even though is in the top 6, but its achievement of the “needs” outcome is the higher (93.5%) and the achievement of its “recovery” outcome is the second highest (77.4%).

When using bad quality as input the behavior of this node is the one less affected by the reduced degrees of freedom, even though there are 2 inputs and 2 outputs. Nine DMUs are under performing and they are at all levels of environmental harshness. The target results in this node are also encouraging. The under performing DMUs have outcome targets percentages that are very realistic and achievable and in line to what agency’s historical data show as the desirable levels of outcome achievement or service quality.
Table 5.17 Performance Scores for Node 4 Using Transformations in the Input Side

<table>
<thead>
<tr>
<th>DMU</th>
<th>Phi</th>
<th>Peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>10060</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>5180</td>
<td>2.625714</td>
<td>10060</td>
</tr>
<tr>
<td>5132</td>
<td>4.966</td>
<td>10060</td>
</tr>
<tr>
<td>5503</td>
<td>2.326181</td>
<td>10060</td>
</tr>
<tr>
<td>5264</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>22322</td>
<td>2.465061</td>
<td>10060</td>
</tr>
<tr>
<td>35168</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>38360</td>
<td>2.58493</td>
<td>10060</td>
</tr>
<tr>
<td>25150</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>35072</td>
<td>2.415109</td>
<td>10060, 35168</td>
</tr>
<tr>
<td>22172</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>12024</td>
<td>3.495</td>
<td>35168</td>
</tr>
<tr>
<td>15080</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>45010</td>
<td>2.562219</td>
<td>10060</td>
</tr>
<tr>
<td>29042</td>
<td>4.557333</td>
<td>10060</td>
</tr>
</tbody>
</table>

Table 5.18 shows the desired input level after the input shortfalls (slacks) have been subtracted. The maximum target in the outcome “covering victim’s most immediate needs” is 93.5 percent for seven DMUs—meaning that the field office needs to increase to that number the percentage of clients who perceived their most immediate needs as covered by the service. Basically this happens because of the dominance of DMU 10060 that has that level of effectiveness for that outcome while operating in a harsher environment.

In the outcome “recovery back to normal life”, the maximum target is 80.4% for DMU 12024 whose only peer (35168) operates in a harsher environment and presents better quality in one dimension and less timeliness, however, its outcomes are a much better levels.

As discussed in Chapter 2 (section 2.2.3) even though the relationship among service quality and effectiveness has not been absolutely determined, there is evidence that higher service quality produces higher outcome achievement and conversely, lower service quality produce lower level of effectiveness. The model prescribes a higher proportionately increase for field office 5132 that has zero timeliness making use of the slack in the flat surface of the frontier.
(from 0 to 59%) and therefore, the percentage increase in outcome achievement recommended to achieve the target is also higher (from 25% to 77.4% or a prescribed increase of 210%).

Among all nodes, the use of variable returns to scale (VRS) makes more sense in this node. This is mainly because modeling in this way enables DMUs to be efficient even if not having a proportionately increased level of outcome achievement. Since the degree by which more service quality will induce more effectiveness is still unclear, VRS provides freedom to the model to deem as best-performing DMUs those that do not present a strict linear relationship between SQ and Outcome achievement but that are still in the northeastern most portion of the solution space.

The next section presents the results for Model II, carrying the output projections of each node as inputs to the subsequent node.
Table 5.18 Projections and Desired Levels of Inputs and Outcomes in Node 4

<table>
<thead>
<tr>
<th>DMU</th>
<th>Lack of Timeliness</th>
<th>Target Lack of Timeliness (Actual – slack)</th>
<th>Lack of Service Quality</th>
<th>Target Lack of Service Quality (actual – slack)</th>
<th>Recovery</th>
<th>Target Recovery</th>
<th>Needs</th>
<th>Target Needs</th>
</tr>
</thead>
<tbody>
<tr>
<td>10060</td>
<td>41.00</td>
<td>41.00</td>
<td>4.32</td>
<td>4.32</td>
<td>77.40</td>
<td>77.4</td>
<td>93.5</td>
<td>93.5</td>
</tr>
<tr>
<td>5180</td>
<td>60.00</td>
<td>41.00</td>
<td>6.26</td>
<td>4.32</td>
<td>60</td>
<td>77.4</td>
<td>70</td>
<td>93.5</td>
</tr>
<tr>
<td>5132</td>
<td>100.00</td>
<td>41.00</td>
<td>10.75</td>
<td>4.32</td>
<td>25.00</td>
<td>77.4</td>
<td>50.00</td>
<td>93.5</td>
</tr>
<tr>
<td>5503</td>
<td>46.00</td>
<td>41.00</td>
<td>5.52</td>
<td>4.32</td>
<td>64.3</td>
<td>77.4</td>
<td>83.3</td>
<td>93.5</td>
</tr>
<tr>
<td>5264</td>
<td>45.00</td>
<td>45.00</td>
<td>4.15</td>
<td>4.15</td>
<td>66.70</td>
<td>66.7</td>
<td>88.9</td>
<td>88.9</td>
</tr>
<tr>
<td>22322</td>
<td>42.00</td>
<td>41.00</td>
<td>4.91</td>
<td>4.32</td>
<td>60.90</td>
<td>77.4</td>
<td>78.30</td>
<td>93.5</td>
</tr>
<tr>
<td>35168</td>
<td>35.00</td>
<td>35.00</td>
<td>4.22</td>
<td>4.22</td>
<td>80.40</td>
<td>80.4</td>
<td>89.10</td>
<td>89.1</td>
</tr>
<tr>
<td>38360</td>
<td>60.00</td>
<td>41.00</td>
<td>5.84</td>
<td>4.32</td>
<td>58.20</td>
<td>77.4</td>
<td>74.50</td>
<td>93.5</td>
</tr>
<tr>
<td>25150</td>
<td>40.00</td>
<td>40.00</td>
<td>4.12</td>
<td>4.12</td>
<td>73.10</td>
<td>73.1</td>
<td>78.80</td>
<td>78.8</td>
</tr>
<tr>
<td>35072</td>
<td>40.00</td>
<td>40.00</td>
<td>4.76</td>
<td>4.30</td>
<td>68.6</td>
<td>77.9</td>
<td>72.5</td>
<td>92.8</td>
</tr>
<tr>
<td>22172</td>
<td>25.00</td>
<td>25.00</td>
<td>3.00</td>
<td>3.00</td>
<td>50.00</td>
<td>50</td>
<td>75.00</td>
<td>75</td>
</tr>
<tr>
<td>12024</td>
<td>65.00</td>
<td>35.00</td>
<td>15.20</td>
<td>4.22</td>
<td>40</td>
<td>80.4</td>
<td>60</td>
<td>89.1</td>
</tr>
<tr>
<td>15080</td>
<td>30.00</td>
<td>30.00</td>
<td>3.23</td>
<td>3.23</td>
<td>77.3</td>
<td>77.3</td>
<td>86.4</td>
<td>86.4</td>
</tr>
<tr>
<td>45010</td>
<td>65.00</td>
<td>41.00</td>
<td>10.33</td>
<td>4.32</td>
<td>66.7</td>
<td>77.4</td>
<td>66.7</td>
<td>93.5</td>
</tr>
<tr>
<td>29042</td>
<td>60.00</td>
<td>41.00</td>
<td>11.12</td>
<td>4.32</td>
<td>37.50</td>
<td>77.4</td>
<td>37.50</td>
<td>93.5</td>
</tr>
</tbody>
</table>
5.2 Application of Model II: Considering the Optimal Level of the Outputs from the Previous Node in Disaster Services

In this section the second formulation (Model II) is applied to the volume data of the same 15 field offices evaluated previously. This formulation utilizes the projected values of the output side of each node as inputs to the next. The logic behind this model is to evaluate what would happen with target resources and outputs if the DMU were in the capacity of implementing the performance improvement recommendations at each node. Model II can be used as a planning tool for forecasting resource needs once the improvements have been achieved. The following sections provide the results at each node. Node 1 is the same as in model II in the sense that it does not use projections from any other node.

5.2.1 Node 2: Capacity Creation Considering the Optimal Level of the Outputs of the Financial Node

Table 5.19 presents the scores and peers for node 2 in Model II.

<table>
<thead>
<tr>
<th>DMU</th>
<th>Phi</th>
<th>Peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>45010</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>29042</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>22172</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>15080</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5132</td>
<td>6.26431</td>
<td>22172, 15080</td>
</tr>
<tr>
<td>12024</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>25150</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>35168</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>35072</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>10060</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5264</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>38360</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>22322</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5503</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5180</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>
For node 2, the projections of donations from node 1 were included as inputs. Let’s recall that node 1 had 2 inefficient DMUs so that their projection was input to the second node. In Model I two DMUs were inefficient in node 2 (5132 and 22322). In model II only one DMU is still inefficient utilizing the optimal level of fund raising (5132). In fact, its inefficiency was increased to 6.26 which is an expected result since the output side of this node is unaltered while at the same time input is required, and therefore, an already inefficient DMUs will become more inefficient. However, DMU 22322 — that presented the lowest performance in Model I— became efficient in Model II. This is because in model I DMU 22322 has DMU 35072 as one of its peers (with a heavy weight of more than 0.87), among others. For Node 2 increased values as inputs for the same outputs are in place for DMUs 35072 and 22322. DMU 5132 will not experience any change, as its peers are neither of those two above. The following changes do occur:

- DMU 35072 has its input increased from 1.05 million to 1.81 million. It remained efficient but shifted its position to another area of the frontier.

- DMU 22322 has its input increased from 2.49 million to 6.53 million. Given that its main peer (35072) also increased its resources and remained efficient, DMU 22322 turns out to be efficient because we are moving a vital pivot point (DMU 35072) and therefore, there is no other peer that is performing better. Hence 22322 turned out to be relatively efficient.

Model I is useful not as a ranking model but else, as a planning tool. For example, it provides DMU 5132 with the expected level of resources (capabilities) this DMU could achieve if the inefficiencies in the fund raising node are corrected. The best take out from this model is the determination of the potential or expected outputs and outcomes after improvement.

Table 5.20 presents the projections for DMU 5132. As it can be observed, the targets for capabilities have been considerably increased. One peer remained the same as prescribed by Model I (DMU 22172) and the other peer changed.
Table 5.20 Projections for DMU 5132 in Node 2

<table>
<thead>
<tr>
<th></th>
<th>5132</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSV Enrollees</td>
<td>49</td>
</tr>
<tr>
<td>Target DSV Enrollees</td>
<td>75.78656</td>
</tr>
<tr>
<td>Disaster registered volunteers</td>
<td>311</td>
</tr>
<tr>
<td>Target Disaster registered volunteers</td>
<td>496.3183</td>
</tr>
<tr>
<td>Disaster Total Paid Staff</td>
<td>3.00</td>
</tr>
<tr>
<td>Target Disaster Total Paid Staff</td>
<td>4.444492</td>
</tr>
<tr>
<td>DAYS OF NET ASSETS</td>
<td>800</td>
</tr>
<tr>
<td>Target DAYS OF NET ASSETS</td>
<td>1312.217</td>
</tr>
</tbody>
</table>

5.2.2 Node 3: Service Delivery Considering the Optimal Level of the Outputs of the Capacity Creation Node

Table 5.21 presents the performance scores for node 3 when 5132 has undergone improvements. 5 DMUs are deemed inefficient here, including 5132.

Table 5.21 Performance scores for Node 3 in Model 2

<table>
<thead>
<tr>
<th>DMU</th>
<th>Phi</th>
<th>Peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>45010</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>29042</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>22172</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>15080</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>5132</td>
<td>4.120607</td>
<td>15080</td>
</tr>
<tr>
<td>12024</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>25150</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>35168</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>35072</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>10060</td>
<td>4.395543</td>
<td>15080, 35072</td>
</tr>
<tr>
<td>5264</td>
<td>3.551047</td>
<td>15080, 35072</td>
</tr>
<tr>
<td>38360</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>22322</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>5503</td>
<td>4.054321</td>
<td>22322</td>
</tr>
<tr>
<td>5180</td>
<td>5.415254</td>
<td>22322</td>
</tr>
</tbody>
</table>
Table 5.22 presents all projections and desired levels of outputs and inputs for node 3. While in Model I all DMUs were deemed as best performing units for this node, in Model II five other DMUs will under perform. This is because the level of input has increased for DMU 5132 while keeping the level of outputs and SQ constant. By shifting the input side of DMU 5132 four other units became under performing, as DEA is a technique that works relative to the DMUs in the sample. The output targets for all those 5 DMUs are pretty reasonable and higher than the ones prescribed in model I, which is expected since the bar has been lifted a little more.

It is important to notice that for DMU 5132, for example, when the projections for node 2 are input in node 3 then DEA computes slacks recommending a reduction of input for the levels of output presented by the DMU, which is something expected. For example, DSV enrollees are reduced from 76 (the projection) to 30. Let’s recall that the original number of DSV enrollees was 49. This means that under current circumstances DMU 5132 is not using its DSV enrollees to their fullest potential to produce quality services. In fact, it is recommended that with 30 DSV enrollees this field office should serve almost the double of families affected by disaster with a target increase from 79 to 144 families or 82%. Also, the service quality index prescribed is 32.8 which represents an increase of 31% from an index of 25 and the % of timely services increased from 0 to 72%.
<table>
<thead>
<tr>
<th>DMU</th>
<th>DSV Enrollees</th>
<th>Target DSV Enrollees</th>
<th>Disaster Registered Volunteers</th>
<th>Target Disaster Registered Volunteers</th>
<th>Disaster Total Paid Staff</th>
<th>Target Disaster Total Paid Staff</th>
<th>Timelines</th>
<th>Target Timelines</th>
<th>SQ_Index</th>
<th>Target SQ_Index</th>
<th>Disaster _Fam_assist</th>
<th>Target Disaster _Fam_assist</th>
</tr>
</thead>
<tbody>
<tr>
<td>45010</td>
<td>49</td>
<td>49</td>
<td>50</td>
<td>50</td>
<td>2</td>
<td>2</td>
<td>35</td>
<td>35</td>
<td>25.67</td>
<td>25.67</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>29042</td>
<td>11</td>
<td>11</td>
<td>127</td>
<td>127</td>
<td>0.5</td>
<td>0.5</td>
<td>40.00</td>
<td>40</td>
<td>24.88</td>
<td>24.88</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>22172</td>
<td>77</td>
<td>77</td>
<td>506</td>
<td>506</td>
<td>4.5</td>
<td>4.5</td>
<td>75.00</td>
<td>75</td>
<td>33.00</td>
<td>33</td>
<td>101</td>
<td>101</td>
</tr>
<tr>
<td>15080</td>
<td>30</td>
<td>30</td>
<td>131</td>
<td>131</td>
<td>2.35</td>
<td>2.35</td>
<td>70</td>
<td>70</td>
<td>32.77</td>
<td>32.77</td>
<td>144</td>
<td>144</td>
</tr>
<tr>
<td>5132</td>
<td>76</td>
<td>30</td>
<td>496</td>
<td>131</td>
<td>4.444</td>
<td>2.35</td>
<td>0.00</td>
<td>72</td>
<td>25.25</td>
<td>32.77</td>
<td>79</td>
<td>144</td>
</tr>
<tr>
<td>12024</td>
<td>50</td>
<td>50</td>
<td>623</td>
<td>623</td>
<td>8</td>
<td>8</td>
<td>35</td>
<td>35</td>
<td>20.8</td>
<td>162</td>
<td>162</td>
<td>162</td>
</tr>
<tr>
<td>25150</td>
<td>119</td>
<td>119</td>
<td>378</td>
<td>378</td>
<td>19</td>
<td>19</td>
<td>60.00</td>
<td>60</td>
<td>31.88</td>
<td>31.88</td>
<td>564</td>
<td>564</td>
</tr>
<tr>
<td>35168</td>
<td>107</td>
<td>107</td>
<td>351</td>
<td>351</td>
<td>8</td>
<td>8</td>
<td>65.00</td>
<td>65</td>
<td>31.78</td>
<td>31.78</td>
<td>369</td>
<td>369</td>
</tr>
<tr>
<td>35072</td>
<td>78</td>
<td>78</td>
<td>1070</td>
<td>1070</td>
<td>16.01</td>
<td>16.01</td>
<td>60</td>
<td>60</td>
<td>31.24</td>
<td>31.24</td>
<td>815</td>
<td>815</td>
</tr>
<tr>
<td>10060</td>
<td>92</td>
<td>60</td>
<td>1907</td>
<td>726</td>
<td>11</td>
<td>11</td>
<td>59.00</td>
<td>63.66764</td>
<td>31.68</td>
<td>31.801</td>
<td>246</td>
<td>568.9012</td>
</tr>
<tr>
<td>5264</td>
<td>115</td>
<td>43</td>
<td>891</td>
<td>385</td>
<td>6.04</td>
<td>6.04</td>
<td>55.00</td>
<td>67.29868</td>
<td>31.85</td>
<td>32.356</td>
<td>248</td>
<td>325.2584</td>
</tr>
<tr>
<td>38360</td>
<td>87</td>
<td>87</td>
<td>636</td>
<td>636</td>
<td>16</td>
<td>16</td>
<td>40.00</td>
<td>40</td>
<td>30.16</td>
<td>30.16</td>
<td>1078</td>
<td>1078</td>
</tr>
<tr>
<td>22322</td>
<td>80</td>
<td>80</td>
<td>292</td>
<td>292</td>
<td>7</td>
<td>7</td>
<td>58.00</td>
<td>58</td>
<td>31.09</td>
<td>31.09</td>
<td>1676</td>
<td>1676</td>
</tr>
<tr>
<td>5503</td>
<td>410</td>
<td>80</td>
<td>758</td>
<td>292</td>
<td>17</td>
<td>7</td>
<td>54</td>
<td>58</td>
<td>30.48</td>
<td>31.09</td>
<td>855</td>
<td>1676</td>
</tr>
<tr>
<td>5180</td>
<td>262</td>
<td>80</td>
<td>1122</td>
<td>292</td>
<td>31</td>
<td>7</td>
<td>40</td>
<td>58</td>
<td>29.74</td>
<td>31.09</td>
<td>574</td>
<td>1676</td>
</tr>
</tbody>
</table>
This example also helps to illustrate the value of allowing each variable in the input/output set to increase or decrease at its own proportion rate to approach the frontier. It would not make sense to recommend the same proportional increase for number of families to the timeliness indicator, for example.

5.2.3 Node 4: Effectiveness Considering the Optimal Level of the Service Quality Indicators of the Service Delivery Node

For node 4 the projections for good quality indicators from node 3 were transformed into bad quality indicators to be used as inputs of node 4.

Table 5.23 presents the performance scores for Node 4 after all the five inefficient DMUs in node 3 have improved their performance at the projected optimal value. When that is the case, the results show 9 DMUs under performing in node 4.

<table>
<thead>
<tr>
<th>DMU</th>
<th>Phi</th>
<th>Peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>10060</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>5180</td>
<td>2.62514</td>
<td>10060</td>
</tr>
<tr>
<td>5132</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>5503</td>
<td>2.32618</td>
<td>10060</td>
</tr>
<tr>
<td>5264</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>22322</td>
<td>2.465061</td>
<td>10060</td>
</tr>
<tr>
<td>35168</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>38360</td>
<td>2.58493</td>
<td>10060</td>
</tr>
<tr>
<td>25150</td>
<td>2.216311</td>
<td>10060, 5264</td>
</tr>
<tr>
<td>35072</td>
<td>2.417935</td>
<td>10060</td>
</tr>
<tr>
<td>22172</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>12024</td>
<td>3.495</td>
<td>35168</td>
</tr>
<tr>
<td>15080</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>45010</td>
<td>2.562219</td>
<td>10060</td>
</tr>
<tr>
<td>29042</td>
<td>4.557333</td>
<td>10060</td>
</tr>
</tbody>
</table>

Let’s recall that in Model I, 9 DMUs were under performing as well in Node 4. However, one DMU that in Model I was deemed efficient became inefficient (25150) and one
DMU that was inefficient became efficient (5132). Once corrective action has been taken in all previous nodes, unit 5132 working in a relatively harsher environment will be deemed effective and unit 25150 working in a median environment (therefore allowed to have more DMUs in its reference set) will become ineffective. The peer set for each unit has also changed and DMU 10060 has become the only peer for most of the units. Thus, the performance scores for most units are the same except for those two DMUs.

Since Model II assumes dependency between nodes, one could say that the final performance score for node 4 represents the rank that the DMU under evaluation would have after undertaking performance improvement recommendations in the previous nodes. If that is the case, DMU 29042 is the one ranked lowest followed by DMU 12024. Of course, this assumption arises several concerns:

- First, it assumes that even after re-engineering all other previous three operations, the outcome achievement level will remain the same. Therefore, the performance score reflects an improvement in the input side with no proportional improvement in the output side. This is of course unrealistic in particular because it is expected that an improvement in quality will indeed carry an improvement in outcomes.

- Model II assumes that all DMUs in the sample are undertaking improvement actions at the same time. Since DEA is a relative performance technique, the results are greatly sensitive to the situation of all DMUs in the sample. Therefore, even the projections for planning are dependant of the situation of all DMUs, which is of course, a big drawback.

Finally, Table 5.24 shows the projections and desired levels for node 4. Inefficient units present either a higher level of bad quality in the input side than their peers or a lower level of outcomes for the same level of quality in the output side. In the case of the one unit that became efficient in Model II while in Model I was inefficient, what occurred was that as this unit was inefficient in node 3, the increased service quality target was input as reduced bad quality in
node 4. Thus a smaller number than the real index and therefore, due to lower input consumption
DMU 5132 became efficient even though it still had low level of outcome production.

In the case of the DMU that became inefficient when it was efficient in model I, what probably happened is that its reference set included a DMU that changed its input side after reengineering in node 3, and with this frontier shift DMU 25150 became inefficient. In fact, the targets for outcomes of DMU 25150 in Model II are higher than in Model I because it is compared to another set of DMUs.

The other DMUs that were reengineered in node 3 remained inefficient in node 4 and DEA computed the same performance score than the one computed in Model I with a lot of slack in the input side. This happens when DMUs are located in the flat portion of the frontier.

An important observation drawn from this table is that in Model II the bad quality index was used as the complement of the optimal projection of good quality in node 3. Hence, the level of inputs is being reduced in Node 4—as compared to node 4 in Model I. Therefore, whenever a DMU has more than one DMU in its peer group, DEA will compute a lower level of outcomes as the prescribed targets for the output side in Model II than the targets computed for those DMUs in Model I. This can be seen in DMU 38360 only, because all others have only one dominant peer (10060).
Table 5.24 Projections And Desired Levels of Outcomes and Inputs for Node 4 with Projected Values

<table>
<thead>
<tr>
<th>DMU</th>
<th>Lack of Timeliness</th>
<th>Target Lack of Timeliness</th>
<th>Lack of SQ_</th>
<th>Target Lack of SQ_</th>
<th>Recovery Needs</th>
<th>Target Needs</th>
</tr>
</thead>
<tbody>
<tr>
<td>10060</td>
<td>36.33</td>
<td>36.33</td>
<td>4.20</td>
<td>4.20</td>
<td>77.40</td>
<td>93.50</td>
</tr>
<tr>
<td>5180</td>
<td>42.00</td>
<td>36.33</td>
<td>4.91</td>
<td>4.20</td>
<td>60</td>
<td>77.4</td>
</tr>
<tr>
<td>5132</td>
<td>28.00</td>
<td>28.00</td>
<td>3.23</td>
<td>3.23</td>
<td>25.00</td>
<td>50.00</td>
</tr>
<tr>
<td>5503</td>
<td>42.00</td>
<td>36.33</td>
<td>4.91</td>
<td>4.20</td>
<td>64.3</td>
<td>77.4</td>
</tr>
<tr>
<td>5264</td>
<td>32.70</td>
<td>32.70</td>
<td>3.64</td>
<td>3.64</td>
<td>66.70</td>
<td>88.90</td>
</tr>
<tr>
<td>22322</td>
<td>42.00</td>
<td>36.33</td>
<td>4.91</td>
<td>4.20</td>
<td>60.90</td>
<td>78.30</td>
</tr>
<tr>
<td>35168</td>
<td>35.00</td>
<td>35.00</td>
<td>4.22</td>
<td>4.22</td>
<td>80.40</td>
<td>89.10</td>
</tr>
<tr>
<td>38360</td>
<td>60.00</td>
<td>36.33</td>
<td>5.84</td>
<td>4.20</td>
<td>58.20</td>
<td>74.50</td>
</tr>
<tr>
<td>25150</td>
<td>40.00</td>
<td>35.82</td>
<td>4.12</td>
<td>4.12</td>
<td>73.10</td>
<td>78.80</td>
</tr>
<tr>
<td>35072</td>
<td>40.00</td>
<td>36.33</td>
<td>4.76</td>
<td>4.20</td>
<td>68.6</td>
<td>72.5</td>
</tr>
<tr>
<td>22172</td>
<td>25.00</td>
<td>25.00</td>
<td>3.00</td>
<td>3.00</td>
<td>50.00</td>
<td>75.00</td>
</tr>
<tr>
<td>12024</td>
<td>65.00</td>
<td>35.00</td>
<td>15.20</td>
<td>4.22</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>15080</td>
<td>30.00</td>
<td>30.00</td>
<td>3.23</td>
<td>3.23</td>
<td>77.3</td>
<td>86.4</td>
</tr>
<tr>
<td>45010</td>
<td>65.00</td>
<td>36.33</td>
<td>10.33</td>
<td>4.20</td>
<td>66.7</td>
<td>66.7</td>
</tr>
<tr>
<td>29042</td>
<td>60.00</td>
<td>36.33</td>
<td>11.12</td>
<td>4.20</td>
<td>37.50</td>
<td>37.50</td>
</tr>
</tbody>
</table>
5.3 Application of Model I in the Total Field Office Operation

Model I was also applied to a different unit of analysis to illustrate the generalizability of the models to different applications. Total field office operation was chosen as the unit of analysis encompassing different lines of service (disaster, international, services to the military, health & safety training and services to the community). A data set of 32 field offices was selected. These 32 field offices are part of a population category of field offices handling medium-size communities.

Variables chosen included the resources and capabilities of each office / DMU used to deliver those services and maintain the operation. The selection of the appropriate variables was done through a structured process as described in Chapter 4. Tables 5.25-5.29 show the selected variables recommended for evaluation of the performance for the total field office operation. It is important to note that senior Agency decision makers evaluated the use of standardized variables in order to make performance comparisons meaningful to the realization of the agency’s work in the respective communities and came to the conclusion that results were more meaningful for the field offices in such a format.
Table 5.25 List of Inputs and Outputs used for the Evaluation of Field Office’s Financial (income generation)

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management &amp; Fundraising Expense per Household (FY-1)</td>
<td>Input - resources Used</td>
<td>Expenses incurred by the field office on management and fundraising campaigns and events during the <em>prior fiscal year</em> (FY-1) as reported in financial forms divided by the number of households in the community.</td>
</tr>
<tr>
<td>Donations per Household</td>
<td>Output – financial return</td>
<td>The sum of the prior year’s donations obtained by the field office through Monetary contributions, special events, disaster relief fund, legacies and bequests and in-kind contribution over year (FY-1) as reported in financial forms divided by the number of households in the community.</td>
</tr>
<tr>
<td>Institutional Funds per Household</td>
<td>Output – financial return</td>
<td>Sum of revenue obtained from the federated United Way campaign and Grants &amp; contracts.</td>
</tr>
<tr>
<td>Variable Name</td>
<td>Type</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>-----------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Donations per Household</td>
<td>Input – Resources Used</td>
<td>The sum of donations obtained by the field office through Monetary contributions, special events, disaster relief fund, legacies and bequests and in-kind contribution over year (FY-1) as reported in financial forms</td>
</tr>
<tr>
<td>Institutional Funds per Household</td>
<td>Input – Resources Used</td>
<td>Sum of revenue obtained from the federated campaign (UW) and Grants &amp; contracts.</td>
</tr>
<tr>
<td>All income not included above</td>
<td>Input – Resources used</td>
<td>The sum of all other income received by the field office in year (FY-1) including products and services</td>
</tr>
<tr>
<td>Number of DSV Enrollees per 1000 households</td>
<td>Output – Organizational Capacity</td>
<td>Number of people registered in the Disaster Services Human Resource System (DSV) divided by number of households in the community.</td>
</tr>
<tr>
<td>Number of Registered Volunteers per 1000 households</td>
<td>Output– Organizational Capacity</td>
<td>Volunteers who have demonstrated an ongoing commitment to the Agency and whose name and address are on file at the field office or station divided by number of households in the community</td>
</tr>
<tr>
<td>Total Paid Staff per 1000 households</td>
<td>Output– Organizational Capacity</td>
<td>The sum of full-time and part-time paid staff (PT staff assumed at 50%)</td>
</tr>
<tr>
<td>Health &amp; Safety Instructors per 1000 households</td>
<td>Output– Organizational Capacity</td>
<td>Number of instructors authorized to teach First Aid, CPR, Swimming, Lifeguard training, and HIV/AIDS courses</td>
</tr>
<tr>
<td>Days of Net Assets</td>
<td>Output- Financial Reserve Capacity</td>
<td>Number of days a field office can continue to operate at current expense levels without additional income (Form 4684)</td>
</tr>
<tr>
<td>Variable Name</td>
<td>Type</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>-----------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Health &amp; Safety reach</td>
<td>Output – Service Level</td>
<td>Percent of population reached by H&amp;S courses calculated as total enrollees divided by population times 100</td>
</tr>
<tr>
<td>Disaster Families assisted per 100 K households</td>
<td>Output – Service Level</td>
<td>Number of families assisted in operations/incidents totally funded by the field office divided by the number of households times 100,000</td>
</tr>
<tr>
<td>Program expenses per household</td>
<td>Output – Service Level</td>
<td>Total program expenditures from Field office Services, Disaster, AFES, Health &amp; Safety, International, Community and Biomedical as a proxy measure of all types of services provided.</td>
</tr>
<tr>
<td>Service quality index for the field office</td>
<td>Output – Service Quality</td>
<td>Proxy measure of service quality obtained from clients’ responses to 9 surveys, weighted by the importance of line of service.</td>
</tr>
</tbody>
</table>
Table 5.28 List of Inputs and Outputs used for the Evaluation of Field Office’s Effectiveness

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad Service Quality index</td>
<td>Input - SQ</td>
<td>Difference between the maximum excellent score of 36 and the score of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Service quality computed for the chapter from clients’ surveys.</td>
</tr>
<tr>
<td>Outcome I for Disaster Services</td>
<td>Output- Client</td>
<td>% of clients who reported they had experienced a favorable outcome: recovery</td>
</tr>
<tr>
<td></td>
<td>Outcome</td>
<td>back to normal life, immediate needs covered.</td>
</tr>
<tr>
<td>Outcome for all H&amp;S courses</td>
<td>Output- Client</td>
<td>% of all enrollees who reported the expected outcome for the course (whatever</td>
</tr>
<tr>
<td></td>
<td>Outcome</td>
<td>that is)</td>
</tr>
</tbody>
</table>

All of the above will be analyzed in groups of field offices associated by the following environmental variables shown in Table 5.29, as follows:

Table 5.29. List of Non-controllable Environmental Variables Used in the Evaluation of Total Field Office’s Operations

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Environmental Factor</td>
<td>Field office population is the total population in the field office's</td>
</tr>
<tr>
<td></td>
<td></td>
<td>jurisdiction based on estimates from CACI Marketing Systems</td>
</tr>
<tr>
<td>Aggregated Household Income</td>
<td>Environmental Factor</td>
<td>Total aggregate household income in the field office jurisdiction from CACI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Marketing Systems</td>
</tr>
<tr>
<td>% Minority</td>
<td>Environmental Factor</td>
<td>Percentage of total population that are minorities including: Black, Hispanic,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asian or Pacific Islander, and American Indian or Alaskan Native based on</td>
</tr>
<tr>
<td></td>
<td></td>
<td>estimates from CACI Marketing Systems</td>
</tr>
<tr>
<td>Population density</td>
<td>Environmental Factor</td>
<td>Field office jurisdiction’s population divided by field office jurisdiction</td>
</tr>
<tr>
<td></td>
<td></td>
<td>area in squared miles</td>
</tr>
</tbody>
</table>
5.3.1 Node 1 (Financial Node) Results for Total Field Office Operations

For this node, acting upon the advise of an agency senior evaluation officer and in coordination with the researcher, fundraising and managerial expenses was chosen as a single input and donations and institutional income as two separate outputs. Aggregated household income was the variable of environmental harshness and minority was chosen as the internal non-controllable input variable as suggested by Banker and Morey (1986).

The results from node 1 are shown in Table 5.30 below. 19 field offices turned out to be under performing. DMU 30116 was the most inefficient field office with a need for increasing its institutional income per household by 21.62 times to $3.27 from $0.15. However, it is performing efficiently in the Donations per household dimension. There are other DMUs like 20064 that are performing inefficiently in both of the output dimensions. It can clearly be seen that the advantage of using a non-radial model helps the decision-maker in determining the varying amounts of improvement that can be made in different dimensions.

Table 5.31 presents the peers for those inefficient units in the fundraising node. As there are two outputs, the performance score for best-practice DMUs is two and greater than two for under performing DMUs.

Let’s take DMU 7038 for illustration purposes. This is a DMU with the following environmental variables: aggregated household income of approximately 25 billion and a percentage of minorities in the population of 19.5%. Its peers are 44084, 17340 and 1019. Table 5.32 below shows that all its peers operated in a less wealthy communities with comparable percentage of minority except for 1019 that had a higher percentage (13%, 16%, and 28%). Therefore, all the prescribed peers for field office 7038 are meaningful. Similarly, the same is seen with DMU 30176 that has close to 33% minority and 32 Billion aggregated household income. Peer DMU 1019 was the only one that had less aggregated household income (harsher environment) and similar minority 28% of population.
Table 5.30 Results in Node 1 for Total Field Office Operations

<table>
<thead>
<tr>
<th>DMU</th>
<th>Score</th>
<th>DONATIONS PHH</th>
<th>INST_INC PHH</th>
<th>MGT FNDRS PHH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Actual Phi1</td>
<td>Target</td>
<td>Actual Phi2</td>
</tr>
<tr>
<td>43092</td>
<td>2.00</td>
<td>1.10 1.00 1.10</td>
<td>1.08 1.00 1.08</td>
<td>0.36 0.00 0.36</td>
</tr>
<tr>
<td>46441</td>
<td>2.00</td>
<td>3.80 1.00 3.80</td>
<td>2.00 1.00 2.00</td>
<td>1.47 0.00 1.47</td>
</tr>
<tr>
<td>36308</td>
<td>2.00</td>
<td>3.28 1.00 3.28</td>
<td>5.85 1.00 5.85</td>
<td>3.21 0.00 3.21</td>
</tr>
<tr>
<td>32004</td>
<td>2.00</td>
<td>3.53 1.00 3.53</td>
<td>1.14 1.00 1.14</td>
<td>1.28 0.00 1.28</td>
</tr>
<tr>
<td>21244</td>
<td>2.69</td>
<td>0.60 1.69 1.02</td>
<td>0.79 1.00 0.79</td>
<td>0.52 0.00 0.52</td>
</tr>
<tr>
<td>46372</td>
<td>2.00</td>
<td>5.53 1.00 5.53</td>
<td>5.19 1.00 5.19</td>
<td>2.80 0.00 2.80</td>
</tr>
<tr>
<td>20064</td>
<td>4.32</td>
<td>1.01 1.67 1.70</td>
<td>0.61 2.64 1.62</td>
<td>0.72 0.00 0.72</td>
</tr>
<tr>
<td>3060</td>
<td>2.00</td>
<td>3.78 1.00 3.78</td>
<td>0.59 1.00 0.59</td>
<td>0.86 0.00 0.86</td>
</tr>
<tr>
<td>38196</td>
<td>2.00</td>
<td>4.49 1.00 4.49</td>
<td>3.65 1.00 3.65</td>
<td>1.98 0.00 1.98</td>
</tr>
<tr>
<td>10061</td>
<td>2.00</td>
<td>2.32 1.00 2.32</td>
<td>0.64 1.00 0.64</td>
<td>0.52 0.00 0.52</td>
</tr>
<tr>
<td>33181</td>
<td>3.22</td>
<td>4.72 1.86 8.76</td>
<td>4.58 1.36 6.23</td>
<td>3.51 0.00 3.51</td>
</tr>
<tr>
<td>32140</td>
<td>3.38</td>
<td>3.41 1.76 6.02</td>
<td>2.98 1.62 4.82</td>
<td>2.63 0.00 2.63</td>
</tr>
<tr>
<td>42296</td>
<td>3.91</td>
<td>2.72 1.89 5.16</td>
<td>1.39 2.01 2.80</td>
<td>1.71 0.00 1.71</td>
</tr>
<tr>
<td>1019</td>
<td>2.00</td>
<td>3.87 1.00 3.87</td>
<td>6.75 1.00 6.75</td>
<td>1.40 0.00 1.40</td>
</tr>
<tr>
<td>10198</td>
<td>2.23</td>
<td>2.51 1.00 2.51</td>
<td>2.12 1.23 2.62</td>
<td>0.75 0.00 0.75</td>
</tr>
<tr>
<td>17340</td>
<td>2.00</td>
<td>3.47 1.00 3.47</td>
<td>9.26 1.00 9.26</td>
<td>2.04 0.00 2.04</td>
</tr>
<tr>
<td>39004</td>
<td>5.07</td>
<td>0.97 1.91 1.84</td>
<td>1.20 3.16 3.80</td>
<td>0.81 0.00 0.81</td>
</tr>
<tr>
<td>44084</td>
<td>2.00</td>
<td>5.51 1.00 5.51</td>
<td>0.87 1.00 0.87</td>
<td>1.18 0.00 1.18</td>
</tr>
<tr>
<td>42264</td>
<td>4.35</td>
<td>1.04 3.35 3.47</td>
<td>4.44 1.00 4.44</td>
<td>1.11 0.00 1.11</td>
</tr>
<tr>
<td>23280</td>
<td>2.65</td>
<td>3.23 1.00 3.23</td>
<td>3.80 1.65 6.27</td>
<td>3.02 1.30 1.71</td>
</tr>
<tr>
<td>7038</td>
<td>5.18</td>
<td>1.81 4.18 7.55</td>
<td>4.20 1.00 4.20</td>
<td>2.06 0.00 2.06</td>
</tr>
<tr>
<td>35086</td>
<td>3.15</td>
<td>3.54 2.15 7.62</td>
<td>5.03 1.00 5.03</td>
<td>2.17 0.00 2.17</td>
</tr>
<tr>
<td>43129</td>
<td>5.87</td>
<td>2.72 1.00 2.73</td>
<td>0.95 4.86 4.61</td>
<td>0.98 0.00 0.98</td>
</tr>
<tr>
<td>30067</td>
<td>2.96</td>
<td>2.78 1.00 2.78</td>
<td>1.49 1.96 2.92</td>
<td>0.84 0.00 0.84</td>
</tr>
<tr>
<td>54004</td>
<td>3.36</td>
<td>2.03 1.65 3.35</td>
<td>3.14 1.70 5.36</td>
<td>1.19 0.00 1.19</td>
</tr>
<tr>
<td>30116</td>
<td>22.62</td>
<td>2.23 1.00 2.23</td>
<td>0.15 21.62 3.27</td>
<td>0.77 0.00 0.77</td>
</tr>
<tr>
<td>28016</td>
<td>3.44</td>
<td>0.68 1.46 0.99</td>
<td>0.76 1.97 1.51</td>
<td>0.35 0.00 0.35</td>
</tr>
<tr>
<td>30140</td>
<td>2.00</td>
<td>3.97 1.00 3.97</td>
<td>1.11 1.00 1.11</td>
<td>0.83 0.00 0.83</td>
</tr>
<tr>
<td>30176</td>
<td>6.44</td>
<td>2.72 1.49 4.04</td>
<td>1.42 4.96 7.05</td>
<td>1.46 0.00 1.46</td>
</tr>
<tr>
<td>14164</td>
<td>2.68</td>
<td>2.89 1.07 3.10</td>
<td>3.84 1.61 6.20</td>
<td>1.32 0.00 1.32</td>
</tr>
<tr>
<td>32442</td>
<td>2.08</td>
<td>4.02 1.00 4.02</td>
<td>3.74 1.08 4.03</td>
<td>1.15 0.00 1.15</td>
</tr>
</tbody>
</table>
Table 5.31 Peers of Under Performing Units in Node 1 for Total Field Office Operations

<table>
<thead>
<tr>
<th>Field office</th>
<th>Peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>21244</td>
<td>32004,36308,46441</td>
</tr>
<tr>
<td>20064</td>
<td>20064,46372,43092</td>
</tr>
<tr>
<td>33181</td>
<td>10061,38196</td>
</tr>
<tr>
<td>32140</td>
<td>10061,38196</td>
</tr>
<tr>
<td>42296</td>
<td>10061,38196</td>
</tr>
<tr>
<td>10198</td>
<td>1019,10061,43092</td>
</tr>
<tr>
<td>39004</td>
<td>17340,1019</td>
</tr>
<tr>
<td>42264</td>
<td>1019,10061,43092</td>
</tr>
<tr>
<td>23280</td>
<td>17340,38196</td>
</tr>
<tr>
<td>7038</td>
<td>44084,17340,1019</td>
</tr>
<tr>
<td>35086</td>
<td>44084,17340,1019</td>
</tr>
<tr>
<td>43129</td>
<td>1019,43092</td>
</tr>
<tr>
<td>30067</td>
<td>1019,10061,43092</td>
</tr>
<tr>
<td>54004</td>
<td>1019,43092</td>
</tr>
<tr>
<td>30116</td>
<td>44084,17340,1019</td>
</tr>
<tr>
<td>28016</td>
<td>1019,43092</td>
</tr>
<tr>
<td>30176</td>
<td>1019</td>
</tr>
<tr>
<td>14164</td>
<td>17340,1019</td>
</tr>
<tr>
<td>32442</td>
<td>30140,1019,43092</td>
</tr>
</tbody>
</table>

All of these results demonstrate the significance of including the environmental harshness variable. Comparisons to models that do not present those variables will be performed later in section 5.4.

Table 5.32 Environmental Variables for the Peers of Field Office 7038

<table>
<thead>
<tr>
<th>DMU CODE</th>
<th>POPULATION</th>
<th>AG HH Income</th>
<th>MINORITY %</th>
</tr>
</thead>
<tbody>
<tr>
<td>7038</td>
<td>983,848</td>
<td>25,149,846</td>
<td>19.47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>With peers:</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>44084</td>
<td>1,030,913</td>
<td>21,645,196</td>
<td>13.36</td>
</tr>
<tr>
<td>17340</td>
<td>1,019,628</td>
<td>21,615,931</td>
<td>15.59</td>
</tr>
<tr>
<td>1019</td>
<td>1,013,636</td>
<td>22,692,714</td>
<td>28.36</td>
</tr>
</tbody>
</table>
5.3.2 Node 2 (Capacity Creation) Results for Total Field Office Operations

For the capacity creation node, the field offices were ranked based on population served as indicated by the Ruggiero variable. Population was ordered from the smallest to the largest. Based on such a ranking, DMU 35106 had the least population and field office 30067 had the highest population. Percentage of minority was also included in the model as a non-controllable environmental input — the Banker & Morey variable.

Inputs for this node included an aggregated measure of all the income resulting from the fundraising effort in addition to the income obtained from products and paid services. The output side or capabilities included all the human resource force of the field office, total paid staff, volunteers, disaster staff (DSV), health & safety instructors and days of net assets as a financial proxy of other assets such as buildings, vehicles and equipment.

The results obtained and the target recommendations for the various field offices are shown in Table 5.33 below. This table lists the field office code, the overall score for this node and the actual and target values for each of the five output variables. All field offices that have a score of 5.0 are efficient in all the five output dimensions. Also, these field offices are the ones that appear as peers for the inefficient field offices.

As can be seen in this case, the first DMU in the list (therefore the one operating in the harsher environment after its respective ranking) is always efficient because of the unavailability of any other field office to be compared with. Further down in the table, each DMU can be compared with all the field offices above it. For instance, field office 33181 is under performing in four of the five output dimensions namely, DSV Enrollees per 10K HH, Registered Volunteers per 1K HH, Total Paid Staff per 100K HH and Days of Net Assets. Nonetheless, DMU 33181 does perform well in the Health & Safety Instructors per 1K HH output dimension.
<table>
<thead>
<tr>
<th>DMU</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>35106</td>
<td>5.0</td>
</tr>
<tr>
<td>8004</td>
<td>5.0</td>
</tr>
<tr>
<td>36308</td>
<td>5.0</td>
</tr>
<tr>
<td>40050</td>
<td>5.0</td>
</tr>
<tr>
<td>46372</td>
<td>5.0</td>
</tr>
<tr>
<td>20064</td>
<td>5.0</td>
</tr>
<tr>
<td>33181</td>
<td>6.5</td>
</tr>
<tr>
<td>32004</td>
<td>5.0</td>
</tr>
<tr>
<td>21244</td>
<td>5.0</td>
</tr>
<tr>
<td>46441</td>
<td>5.0</td>
</tr>
<tr>
<td>42296</td>
<td>6.6</td>
</tr>
<tr>
<td>30140</td>
<td>5.0</td>
</tr>
<tr>
<td>20060</td>
<td>12.8</td>
</tr>
<tr>
<td>5112</td>
<td>5.0</td>
</tr>
<tr>
<td>32442</td>
<td>9.2</td>
</tr>
<tr>
<td>38196</td>
<td>5.0</td>
</tr>
<tr>
<td>43092</td>
<td>5.0</td>
</tr>
<tr>
<td>7038</td>
<td>5.0</td>
</tr>
<tr>
<td>23280</td>
<td>5.0</td>
</tr>
<tr>
<td>1019</td>
<td>15.9</td>
</tr>
<tr>
<td>17340</td>
<td>5.0</td>
</tr>
<tr>
<td>44084</td>
<td>13.9</td>
</tr>
<tr>
<td>39004</td>
<td>5.0</td>
</tr>
<tr>
<td>32140</td>
<td>7.4</td>
</tr>
<tr>
<td>3060</td>
<td>17.3</td>
</tr>
<tr>
<td>10198</td>
<td>5.0</td>
</tr>
<tr>
<td>43129</td>
<td>10.9</td>
</tr>
<tr>
<td>30176</td>
<td>10.0</td>
</tr>
<tr>
<td>10228</td>
<td>5.0</td>
</tr>
<tr>
<td>35086</td>
<td>9.6</td>
</tr>
<tr>
<td>30116</td>
<td>9.6</td>
</tr>
<tr>
<td>42264</td>
<td>9.1</td>
</tr>
<tr>
<td>54004</td>
<td>5.0</td>
</tr>
<tr>
<td>30067</td>
<td>32.1</td>
</tr>
</tbody>
</table>
Table 5.34 Results from Node 2 for Total Field Office Operations

<table>
<thead>
<tr>
<th>DMU</th>
<th>Score</th>
<th>DS_DSV_P10_KHH</th>
<th>VL_REG_P1K_HH</th>
<th>PD_TOTAL_P_100KHH</th>
<th>HS_INST_P1_KHH</th>
<th>DAYASET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Actual</td>
<td>Target</td>
<td>Actual</td>
<td>Target</td>
<td>Actual</td>
</tr>
<tr>
<td>35106</td>
<td>5.0</td>
<td>1.9</td>
<td>1.9</td>
<td>6.9</td>
<td>6.9</td>
<td>30.5</td>
</tr>
<tr>
<td>8004</td>
<td>5.0</td>
<td>0.8</td>
<td>0.8</td>
<td>4.3</td>
<td>4.3</td>
<td>17.4</td>
</tr>
<tr>
<td>36308</td>
<td>5.0</td>
<td>3.3</td>
<td>3.3</td>
<td>4.2</td>
<td>4.2</td>
<td>31.3</td>
</tr>
<tr>
<td>40050</td>
<td>5.0</td>
<td>1.8</td>
<td>1.8</td>
<td>5.1</td>
<td>5.1</td>
<td>15.4</td>
</tr>
<tr>
<td>46372</td>
<td>5.0</td>
<td>2.5</td>
<td>2.5</td>
<td>8.3</td>
<td>8.3</td>
<td>30.3</td>
</tr>
<tr>
<td>20064</td>
<td>5.0</td>
<td>0.6</td>
<td>0.6</td>
<td>7.2</td>
<td>7.2</td>
<td>7.0</td>
</tr>
<tr>
<td>33181</td>
<td>6.5</td>
<td>2.2</td>
<td>2.5</td>
<td>5.9</td>
<td>8.1</td>
<td>21.8</td>
</tr>
<tr>
<td>32004</td>
<td>5.0</td>
<td>2.2</td>
<td>2.2</td>
<td>4.6</td>
<td>4.6</td>
<td>11.5</td>
</tr>
<tr>
<td>21244</td>
<td>5.0</td>
<td>0.7</td>
<td>1.6</td>
<td>1.6</td>
<td>10.3</td>
<td>10.3</td>
</tr>
<tr>
<td>46441</td>
<td>5.0</td>
<td>1.7</td>
<td>1.7</td>
<td>12.5</td>
<td>12.5</td>
<td>19.6</td>
</tr>
<tr>
<td>42296</td>
<td>6.6</td>
<td>1.8</td>
<td>2.0</td>
<td>7.2</td>
<td>8.3</td>
<td>14.9</td>
</tr>
<tr>
<td>30140</td>
<td>5.0</td>
<td>0.6</td>
<td>0.6</td>
<td>5.9</td>
<td>5.9</td>
<td>21.7</td>
</tr>
<tr>
<td>20060</td>
<td>12.8</td>
<td>0.7</td>
<td>1.4</td>
<td>2.5</td>
<td>11.0</td>
<td>10.1</td>
</tr>
<tr>
<td>5112</td>
<td>5.0</td>
<td>0.9</td>
<td>0.9</td>
<td>0.5</td>
<td>0.5</td>
<td>9.2</td>
</tr>
<tr>
<td>32442</td>
<td>9.2</td>
<td>1.2</td>
<td>3.1</td>
<td>11.4</td>
<td>12.4</td>
<td>10.7</td>
</tr>
<tr>
<td>38196</td>
<td>5.0</td>
<td>2.4</td>
<td>2.4</td>
<td>9.9</td>
<td>9.9</td>
<td>24.6</td>
</tr>
<tr>
<td>43092</td>
<td>5.0</td>
<td>3.8</td>
<td>3.8</td>
<td>5.4</td>
<td>5.4</td>
<td>4.1</td>
</tr>
<tr>
<td>7038</td>
<td>5.0</td>
<td>2.1</td>
<td>2.1</td>
<td>4.7</td>
<td>4.7</td>
<td>12.4</td>
</tr>
<tr>
<td>23280</td>
<td>5.0</td>
<td>1.8</td>
<td>1.8</td>
<td>3.6</td>
<td>3.6</td>
<td>41.1</td>
</tr>
<tr>
<td>1019</td>
<td>15.9</td>
<td>1.7</td>
<td>2.4</td>
<td>3.1</td>
<td>9.3</td>
<td>18.0</td>
</tr>
<tr>
<td>17340</td>
<td>5.0</td>
<td>2.5</td>
<td>2.5</td>
<td>9.3</td>
<td>9.3</td>
<td>48.8</td>
</tr>
<tr>
<td>44084</td>
<td>13.9</td>
<td>1.2</td>
<td>1.6</td>
<td>1.1</td>
<td>7.2</td>
<td>8.1</td>
</tr>
<tr>
<td>39004</td>
<td>5.0</td>
<td>2.1</td>
<td>2.1</td>
<td>4.4</td>
<td>4.4</td>
<td>4.6</td>
</tr>
<tr>
<td>32140</td>
<td>7.4</td>
<td>0.9</td>
<td>2.0</td>
<td>5.0</td>
<td>5.2</td>
<td>20.4</td>
</tr>
<tr>
<td>3060</td>
<td>17.3</td>
<td>1.5</td>
<td>1.9</td>
<td>1.1</td>
<td>10.6</td>
<td>4.9</td>
</tr>
<tr>
<td>10198</td>
<td>5.0</td>
<td>1.8</td>
<td>1.8</td>
<td>3.6</td>
<td>3.6</td>
<td>8.3</td>
</tr>
<tr>
<td>43129</td>
<td>10.9</td>
<td>2.4</td>
<td>2.4</td>
<td>2.3</td>
<td>10.8</td>
<td>10.9</td>
</tr>
<tr>
<td>30176</td>
<td>10.0</td>
<td>1.4</td>
<td>1.7</td>
<td>3.0</td>
<td>11.7</td>
<td>9.4</td>
</tr>
<tr>
<td>10228</td>
<td>5.0</td>
<td>1.2</td>
<td>1.2</td>
<td>8.2</td>
<td>8.2</td>
<td>11.1</td>
</tr>
<tr>
<td>35086</td>
<td>9.6</td>
<td>1.7</td>
<td>3.2</td>
<td>4.4</td>
<td>8.4</td>
<td>13.4</td>
</tr>
<tr>
<td>30116</td>
<td>9.6</td>
<td>0.7</td>
<td>1.0</td>
<td>4.0</td>
<td>5.4</td>
<td>4.3</td>
</tr>
<tr>
<td>42264</td>
<td>9.1</td>
<td>0.8</td>
<td>2.8</td>
<td>4.4</td>
<td>8.7</td>
<td>11.3</td>
</tr>
<tr>
<td>54004</td>
<td>5.0</td>
<td>4.5</td>
<td>4.5</td>
<td>10.4</td>
<td>10.4</td>
<td>21.3</td>
</tr>
<tr>
<td>30067</td>
<td>32.1</td>
<td>0.6</td>
<td>2.7</td>
<td>1.0</td>
<td>15.8</td>
<td>8.0</td>
</tr>
</tbody>
</table>
Looking at the results from field office 30067, this DMU is highly inefficient with maximum improvement been required in the Registered Volunteers per 1K HH, where the value needs to increase from 1.0 to 15.8.

Table 5.35 Peers in Node 2 for Total Field Office Operations

<table>
<thead>
<tr>
<th>Field office</th>
<th>Peer</th>
</tr>
</thead>
<tbody>
<tr>
<td>33181</td>
<td>46372,36308,35106</td>
</tr>
<tr>
<td>42296</td>
<td>46441,32004</td>
</tr>
<tr>
<td>20060</td>
<td>46441,20064</td>
</tr>
<tr>
<td>32442</td>
<td>46441,32004</td>
</tr>
<tr>
<td>1019</td>
<td>23280,46441</td>
</tr>
<tr>
<td>44084</td>
<td>23280,38196,46441</td>
</tr>
<tr>
<td>32140</td>
<td>39004,23280,46441</td>
</tr>
<tr>
<td>3060</td>
<td>43092,46441</td>
</tr>
<tr>
<td>43129</td>
<td>39004,43092,46441</td>
</tr>
<tr>
<td>30176</td>
<td>23280,46441</td>
</tr>
<tr>
<td>35086</td>
<td>23280,46441,32004</td>
</tr>
<tr>
<td>30116</td>
<td>23280,46441</td>
</tr>
<tr>
<td>42264</td>
<td>39004,43092,46441</td>
</tr>
<tr>
<td><strong>30067</strong></td>
<td><strong>43092,46441</strong></td>
</tr>
</tbody>
</table>

Table 5.35 also shows that field offices 46441 and 43092 are the peers for field office 30067. These results are meaningful since both peers operate in harsher environments according to population (smaller communities) and use relatively less resources (income) to create more or comparable capabilities. Regarding the % minority variable, 46441 presents a lower percentage minority (34.46%) and 43092 presents a much higher percentage minority (88.96%) illustrating the properties of the Banker and Morey formulation. The weighted average of both DMUs will produce the ideal projection to the frontier of DMU 30067. For instance, DMU 43092 outperforms by far to DMU 30067 in more than one output — e.g. 5.2 volunteers per
thousand households vs. 1 volunteer per thousand households—while at the same time using much lesser resources than 30067.

5.3.3 Node 3: (Service Delivery) Results for Total Field Office Operations

To perform the corresponding analysis for this node this study faced the constraint that not all necessary data were available. For instance, the data for the service quality index from all lines of service were unavailable since most field offices for this sample were not part of the pilot test of the survey instrument. In order to illustrate the applicability of Model I to different units of analysis, these unavailable data needed to be randomly generated. The generation on a computer of random variables from different continuous and discrete distributions is a common procedure during the testing and validation phase of optimization models, and recently of DEA formulations as well (e.g. Pedraja-Chaparro et. al., 1999 Rubinstein and Kreimer, 1988; Ruggiero, 1998). For this purpose, the random generator in MS Excel was used. As a result, continuous values between a minimum of 6 (minimum possible rate for each service quality construct) and a maximum of 36 were generated for 32 observations.

In this node, field offices were ranked based on the Ruggiero variable (population) ranging from the largest to the smallest. Based on such a ranking, field office 30067 had the largest population and field office 35106 had the smaller population size. The Banker & Morey variable (Population Density) was also included in the model assuming that the spread of the service delivery area was an important factor affecting service delivery operations.

Inputs for this node were total paid staff, volunteers, disaster staff (DSV), and health & safety instructors. Outputs were the measures of service delivery units or market penetration when available and reliable in addition to program dollars for all other lines of service where a standard measure of service delivery was not found, such as health & safety reach (enrollees as % of population), disaster families assisted and program expenses for all other lines of service.

The results obtained and the target recommendations for the various field offices are shown in Table 5.36 below. This table lists the field office code, the overall score for this node and the actual and target values for each of the four output variables. All field offices that have a
score of 4.0 are efficient in all the four output dimensions. 14 field offices were deemed as under performing in this node.

The 18 field offices deemed efficient or best performer are the ones that appear as peers for the inefficient field offices. As it can be seen, the first field office is efficient because of the unavailability of any other field office to be compared with. Further down in the table, each field office is allowed to be compared with all the field offices above it. It is seen for instance, that field office 33181, which was inefficient in the previous node, is efficient in this node.

Looking at the results from field office 20064, which are highly inefficient with maximum improvement require in the program expenses for other lines of service (OTH_PRGEXP_PHH) which also is a proxy measure of the volume of operations in other lines of service than Disaster and Health & Safety training. The value of this indicator needs to be increased from 0.7 to 6.3 dollars per household. Also the target for service quality should go from 8 to the maximum of 36. From Table 5.37, it can be seen that field office 17340 is its main peer with (weight = 0.45). Other peers include DMUs 54004, 23280, and 30067. All four peers operate in much larger communities, making service delivery more difficult. Regarding the population density as a measure of area covered, DMU 20064 operates in a small area with very high density (1624 inhabitants per sq. mile). Its peers operate all at different levels, one with less density (DMU 17340) and another DMU with more than twice inhabitants per square feet (DMU 30067). Relatively speaking, all peers are outperforming DMU 20064. So, it can be concluded that the results of this node are meaningful.

Another important validation factor for these results was the agency’s policy regarding Health & Safety reach. It is a desirable standard for all chapters to reach at least 5% of their population with these types of services (instruction/courses). Agency officials were pleased to see the DEA results that prescribed targets around 5%, with a maximum of 8.74 prescribed to one of the units with highest number of paid staff and instructors. All of this makes sense since this field office (DMU 8004) would have more resources to increase its community reach.
<table>
<thead>
<tr>
<th>DMU</th>
<th>Score</th>
<th>DSV Enrollees per 10K HH</th>
<th>Registered Volunteers per 1K HH</th>
<th>Total Paid staff per 100K HH</th>
<th>Health &amp; Safety Instructors per 1K HH</th>
<th>H&amp;S Reach</th>
<th>Disaster Families assisted per 10K HH</th>
<th>Program Exp. For other lines of service per HH</th>
<th>SQ_INDEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>30067</td>
<td>4.00</td>
<td>0.70</td>
<td>2.53</td>
<td>10.08</td>
<td>1.33</td>
<td>1.75</td>
<td>5.47</td>
<td>2.08</td>
<td>25.00</td>
</tr>
<tr>
<td>54004</td>
<td>4.00</td>
<td>0.64</td>
<td>0.98</td>
<td>8.05</td>
<td>1.13</td>
<td>1.13</td>
<td>4.73</td>
<td>4.73</td>
<td>1.00</td>
</tr>
<tr>
<td>42264</td>
<td>4.00</td>
<td>2.41</td>
<td>2.35</td>
<td>10.87</td>
<td>3.33</td>
<td>3.34</td>
<td>8.62</td>
<td>8.62</td>
<td>0.63</td>
</tr>
<tr>
<td>30116</td>
<td>4.00</td>
<td>0.89</td>
<td>0.54</td>
<td>9.24</td>
<td>1.26</td>
<td>1.26</td>
<td>2.65</td>
<td>2.42</td>
<td>1.60</td>
</tr>
<tr>
<td>35086</td>
<td>4.98</td>
<td>2.52</td>
<td>9.26</td>
<td>48.78</td>
<td>4.24</td>
<td>6.42</td>
<td>7.25</td>
<td>6.24</td>
<td>1.93</td>
</tr>
<tr>
<td>10228</td>
<td>4.00</td>
<td>2.21</td>
<td>5.92</td>
<td>21.79</td>
<td>5.26</td>
<td>4.03</td>
<td>4.05</td>
<td>6.96</td>
<td>2.12</td>
</tr>
<tr>
<td>30176</td>
<td>7.03</td>
<td>2.43</td>
<td>9.92</td>
<td>24.64</td>
<td>8.18</td>
<td>5.96</td>
<td>5.75</td>
<td>1.73</td>
<td>19.00</td>
</tr>
<tr>
<td>43129</td>
<td>5.84</td>
<td>0.96</td>
<td>4.45</td>
<td>13.43</td>
<td>4.17</td>
<td>2.63</td>
<td>4.48</td>
<td>6.00</td>
<td>0.78</td>
</tr>
<tr>
<td>10198</td>
<td>4.00</td>
<td>0.56</td>
<td>7.19</td>
<td>7.03</td>
<td>3.02</td>
<td>1.94</td>
<td>7.35</td>
<td>3.57</td>
<td>0.80</td>
</tr>
<tr>
<td>3060</td>
<td>4.00</td>
<td>0.82</td>
<td>4.28</td>
<td>17.39</td>
<td>2.71</td>
<td>3.99</td>
<td>4.99</td>
<td>3.60</td>
<td>21.00</td>
</tr>
<tr>
<td>32140</td>
<td>8.10</td>
<td>2.15</td>
<td>4.73</td>
<td>12.35</td>
<td>9.09</td>
<td>3.59</td>
<td>4.67</td>
<td>7.07</td>
<td>0.46</td>
</tr>
<tr>
<td>39004</td>
<td>4.00</td>
<td>1.49</td>
<td>1.11</td>
<td>4.89</td>
<td>2.21</td>
<td>1.22</td>
<td>6.98</td>
<td>6.98</td>
<td>0.33</td>
</tr>
<tr>
<td>44084</td>
<td>4.00</td>
<td>0.71</td>
<td>4.01</td>
<td>4.28</td>
<td>2.95</td>
<td>4.88</td>
<td>2.05</td>
<td>1.47</td>
<td>18.00</td>
</tr>
<tr>
<td>17340</td>
<td>4.00</td>
<td>2.23</td>
<td>4.59</td>
<td>11.48</td>
<td>4.62</td>
<td>11.54</td>
<td>9.33</td>
<td>5.12</td>
<td>25.00</td>
</tr>
<tr>
<td>1019</td>
<td>4.00</td>
<td>1.88</td>
<td>6.86</td>
<td>30.45</td>
<td>5.27</td>
<td>13.16</td>
<td>7.03</td>
<td>6.46</td>
<td>34.00</td>
</tr>
<tr>
<td>23280</td>
<td>4.00</td>
<td>0.58</td>
<td>5.91</td>
<td>21.65</td>
<td>4.13</td>
<td>4.62</td>
<td>3.05</td>
<td>14.00</td>
<td>0.60</td>
</tr>
<tr>
<td>7038</td>
<td>6.89</td>
<td>1.79</td>
<td>5.14</td>
<td>15.45</td>
<td>6.57</td>
<td>7.07</td>
<td>0.85</td>
<td>2.74</td>
<td>24.00</td>
</tr>
<tr>
<td>43092</td>
<td>8.22</td>
<td>0.81</td>
<td>8.21</td>
<td>11.11</td>
<td>3.36</td>
<td>3.36</td>
<td>5.96</td>
<td>1.63</td>
<td>27.00</td>
</tr>
<tr>
<td>38196</td>
<td>4.00</td>
<td>0.91</td>
<td>4.99</td>
<td>20.37</td>
<td>3.92</td>
<td>7.36</td>
<td>3.76</td>
<td>3.76</td>
<td>1.63</td>
</tr>
<tr>
<td>32442</td>
<td>4.45</td>
<td>1.84</td>
<td>3.63</td>
<td>41.09</td>
<td>5.36</td>
<td>10.19</td>
<td>2.47</td>
<td>3.44</td>
<td>34.00</td>
</tr>
<tr>
<td>5112</td>
<td>4.00</td>
<td>3.26</td>
<td>4.25</td>
<td>31.27</td>
<td>1.65</td>
<td>7.26</td>
<td>0.00</td>
<td>27.00</td>
<td>0.00</td>
</tr>
<tr>
<td>20060</td>
<td>8.18</td>
<td>0.65</td>
<td>1.57</td>
<td>10.28</td>
<td>1.33</td>
<td>1.34</td>
<td>0.96</td>
<td>0.96</td>
<td>20.00</td>
</tr>
<tr>
<td>30140</td>
<td>6.38</td>
<td>1.16</td>
<td>11.44</td>
<td>10.70</td>
<td>2.54</td>
<td>2.54</td>
<td>7.06</td>
<td>1.30</td>
<td>20.00</td>
</tr>
<tr>
<td>DMU</td>
<td>Score</td>
<td>DSV Enrollees per 10K HH</td>
<td>Registered Volunteers per 1K HH</td>
<td>Total Paid staff per 100K HH</td>
<td>Health &amp; Safety Instructors per 1K HH</td>
<td>H&amp;S Reach</td>
<td>Disaster Families assisted per 10K HH</td>
<td>Program Exp. For other lines of service per HH</td>
<td>SQ_INDEX</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>--------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>----------------------------------------</td>
<td>-----------</td>
<td>-------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>42296</td>
<td>11.09</td>
<td>1.38</td>
<td>3.03</td>
<td>9.43</td>
<td>9.43</td>
<td>4.88</td>
<td>3.92</td>
<td>1.76</td>
<td>5.28</td>
</tr>
<tr>
<td>46441</td>
<td>12.73</td>
<td>4.46</td>
<td>2.63</td>
<td>10.39</td>
<td>5.14</td>
<td>10.39</td>
<td>5.14</td>
<td>3.53</td>
<td>5.78</td>
</tr>
<tr>
<td>21244</td>
<td>14.26</td>
<td>1.79</td>
<td>1.31</td>
<td>14.90</td>
<td>12.19</td>
<td>3.11</td>
<td>3.11</td>
<td>3.62</td>
<td>4.60</td>
</tr>
<tr>
<td>32004</td>
<td>4.00</td>
<td>3.84</td>
<td>3.84</td>
<td>5.40</td>
<td>4.10</td>
<td>2.76</td>
<td>2.76</td>
<td>5.46</td>
<td>5.46</td>
</tr>
<tr>
<td>33181</td>
<td>4.00</td>
<td>1.65</td>
<td>1.65</td>
<td>3.14</td>
<td>3.14</td>
<td>18.04</td>
<td>18.04</td>
<td>2.26</td>
<td>2.26</td>
</tr>
<tr>
<td>20064</td>
<td>17.64</td>
<td>1.69</td>
<td>1.69</td>
<td>12.47</td>
<td>19.60</td>
<td>19.60</td>
<td>5.63</td>
<td>2.74</td>
<td>6.67</td>
</tr>
<tr>
<td>46372</td>
<td>4.00</td>
<td>0.76</td>
<td>0.76</td>
<td>4.43</td>
<td>4.43</td>
<td>11.33</td>
<td>11.33</td>
<td>2.92</td>
<td>2.92</td>
</tr>
<tr>
<td>40050</td>
<td>4.00</td>
<td>1.76</td>
<td>1.76</td>
<td>3.59</td>
<td>3.59</td>
<td>8.31</td>
<td>8.31</td>
<td>6.16</td>
<td>6.16</td>
</tr>
<tr>
<td>8004</td>
<td>9.71</td>
<td>2.46</td>
<td>1.98</td>
<td>8.32</td>
<td>7.13</td>
<td>30.31</td>
<td>16.04</td>
<td>5.01</td>
<td>5.01</td>
</tr>
<tr>
<td>35106</td>
<td>4.00</td>
<td>1.15</td>
<td>1.15</td>
<td>1.14</td>
<td>1.14</td>
<td>8.14</td>
<td>8.14</td>
<td>6.19</td>
<td>6.19</td>
</tr>
</tbody>
</table>

Cont’d of Table 5.36
The application of this node had the purpose of testing the model formulations and illustrating its potential results. These results had no practical implications to the agency under study for real data was not available for this node. Data for generic Outcomes 1 and 2 needed to be randomly generated with the same procedure described for node 3. Outcomes 1 and 2 were assumed to be the percentage of effectiveness achieved by the field office as perceived by the clients or customers in for the two desirable outcomes. Further, the input variable SQ_INDEX_INV is the complement of SQ_INDEX from node 3 and was calculated by subtracting the field office index value from the maximum possible service quality index, which is 36.

The results obtained and the target recommendations for the various field offices are as shown in Table 5.38 below. The table lists the field office code, the overall score for this node and the actual and target values for each of the two outcome variables. In this node, the field offices were ranked based on % minority as the environmental harshness variable ranging from the largest to the smallest. Based on such a ranking, field office 43092 had the largest % minority (over 88%). There was no internal environmental variable for this node.

<table>
<thead>
<tr>
<th>Field office</th>
<th>Peer</th>
</tr>
</thead>
<tbody>
<tr>
<td>35086</td>
<td>54004,30116</td>
</tr>
<tr>
<td>30176</td>
<td>10228,30116,30067</td>
</tr>
<tr>
<td>43129</td>
<td>10228,54004</td>
</tr>
<tr>
<td>32140</td>
<td>30116,54004,30067</td>
</tr>
<tr>
<td>7038</td>
<td>23280,17340,44084,54004,30067</td>
</tr>
<tr>
<td>43092</td>
<td>43092,1019,54004</td>
</tr>
<tr>
<td>32442</td>
<td>23280,17340,30116,54004,30067</td>
</tr>
<tr>
<td>20060</td>
<td>30116,54004,30067</td>
</tr>
<tr>
<td>30140</td>
<td>17340,44084,39004,30067</td>
</tr>
<tr>
<td>42296</td>
<td>17340,44084,54004,30067</td>
</tr>
<tr>
<td>46441</td>
<td>17340,30116,30067</td>
</tr>
<tr>
<td>21244</td>
<td>17340,54004,30067</td>
</tr>
<tr>
<td>20064</td>
<td>23280,17340,1019,54004,30067</td>
</tr>
<tr>
<td>8004</td>
<td>36308,1019,44084,3060</td>
</tr>
</tbody>
</table>

Table 5.37 Peers in Node 3 for Total Field office Operations

5.3.4 Node 4 (Customer Effectiveness Node) Results for Total Field Office Operations

The application of this node had the purpose of testing the model formulations and illustrating its potential results. These results had no practical implications to the agency under study for real data was not available for this node. Data for generic Outcomes 1 and 2 needed to be randomly generated with the same procedure described for node 3. Outcomes 1 and 2 were assumed to be the percentage of effectiveness achieved by the field office as perceived by the clients or customers in for the two desirable outcomes. Further, the input variable SQ_INDEX_INV is the complement of SQ_INDEX from node 3 and was calculated by subtracting the field office index value from the maximum possible service quality index, which is 36.

The results obtained and the target recommendations for the various field offices are as shown in Table 5.38 below. The table lists the field office code, the overall score for this node and the actual and target values for each of the two outcome variables. In this node, the field offices were ranked based on % minority as the environmental harshness variable ranging from the largest to the smallest. Based on such a ranking, field office 43092 had the largest % minority (over 88%). There was no internal environmental variable for this node.
All field offices that have a score of 2.0 are efficient in both of the outcome dimensions. A little more than half the number of DMUs were deemed inefficient (18) for this node. All inefficient field offices presented inefficiencies only in outcome 2. Of course, these values are relative to the values generated by the random procedure. However, there are several DMUs with scores very close to 2, so that the prescribed targets are generally very close to the actual values in the outcome side. The maximum target for outcome 1 was 92% and in outcome 2 it was 96%.
<table>
<thead>
<tr>
<th>DMU</th>
<th>Score</th>
<th>Phi1</th>
<th>Phi2</th>
<th>SQ_INDEX</th>
<th>INV</th>
<th>Outcome1</th>
<th>Outcome2</th>
<th>MINORITY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
<td>Target</td>
<td>Actual</td>
<td>Target</td>
<td>Actual</td>
<td>Target</td>
<td>Actual</td>
<td>Target</td>
</tr>
<tr>
<td>43092</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>25.00</td>
<td>97.00</td>
<td>97.00</td>
<td>56.50</td>
<td>88.96</td>
</tr>
<tr>
<td>54004</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>9.00</td>
<td>92.00</td>
<td>92.00</td>
<td>55.50</td>
<td>70.77</td>
</tr>
<tr>
<td>20064</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>28.00</td>
<td>43.00</td>
<td>43.00</td>
<td>34.00</td>
<td>66.84</td>
</tr>
<tr>
<td>30067</td>
<td>5.20</td>
<td>4.03</td>
<td>1.17</td>
<td>11.00</td>
<td>23.00</td>
<td>92.63</td>
<td>55.63</td>
<td>57.94</td>
</tr>
<tr>
<td>5112</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>9.00</td>
<td>96.00</td>
<td>96.00</td>
<td>88.50</td>
<td>56.87</td>
</tr>
<tr>
<td>42264</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>66.00</td>
<td>66.00</td>
<td>82.00</td>
<td>44.51</td>
</tr>
<tr>
<td>3060</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>15.00</td>
<td>67.00</td>
<td>67.00</td>
<td>77.00</td>
<td>39.81</td>
</tr>
<tr>
<td>50050</td>
<td>5.73</td>
<td>3.45</td>
<td>2.28</td>
<td>19.00</td>
<td>28.00</td>
<td>96.63</td>
<td>30.00</td>
<td>38.12</td>
</tr>
<tr>
<td>43129</td>
<td>3.24</td>
<td>1.99</td>
<td>1.25</td>
<td>7.00</td>
<td>45.00</td>
<td>89.33</td>
<td>69.50</td>
<td>36.88</td>
</tr>
<tr>
<td>20060</td>
<td>2.81</td>
<td>1.48</td>
<td>1.33</td>
<td>16.00</td>
<td>65.00</td>
<td>96.44</td>
<td>56.00</td>
<td>74.50</td>
</tr>
<tr>
<td>46441</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>29.00</td>
<td>25.00</td>
<td>25.00</td>
<td>45.00</td>
<td>34.46</td>
</tr>
<tr>
<td>30176</td>
<td>2.85</td>
<td>1.06</td>
<td>1.79</td>
<td>17.00</td>
<td>91.00</td>
<td>96.50</td>
<td>40.50</td>
<td>72.50</td>
</tr>
<tr>
<td>32442</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>2.00</td>
<td>73.00</td>
<td>73.00</td>
<td>70.00</td>
<td>32.14</td>
</tr>
<tr>
<td>46372</td>
<td>4.92</td>
<td>3.87</td>
<td>1.05</td>
<td>19.00</td>
<td>25.00</td>
<td>96.63</td>
<td>65.00</td>
<td>31.62</td>
</tr>
<tr>
<td>33181</td>
<td>2.68</td>
<td>1.11</td>
<td>1.57</td>
<td>18.00</td>
<td>87.00</td>
<td>96.56</td>
<td>45.00</td>
<td>70.50</td>
</tr>
<tr>
<td>10198</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>37.00</td>
<td>37.00</td>
<td>98.00</td>
<td>30.21</td>
</tr>
<tr>
<td>1019</td>
<td>4.50</td>
<td>3.30</td>
<td>1.19</td>
<td>2.00</td>
<td>22.00</td>
<td>72.67</td>
<td>70.00</td>
<td>28.36</td>
</tr>
<tr>
<td>10128</td>
<td>3.08</td>
<td>2.06</td>
<td>1.02</td>
<td>22.00</td>
<td>47.00</td>
<td>96.81</td>
<td>61.00</td>
<td>27.44</td>
</tr>
<tr>
<td>8004</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>22.00</td>
<td>99.00</td>
<td>99.00</td>
<td>62.00</td>
<td>25.06</td>
</tr>
<tr>
<td>30140</td>
<td>3.08</td>
<td>1.83</td>
<td>1.25</td>
<td>6.00</td>
<td>86.00</td>
<td>96.56</td>
<td>69.00</td>
<td>24.82</td>
</tr>
<tr>
<td>42296</td>
<td>2.91</td>
<td>1.53</td>
<td>1.39</td>
<td>24.00</td>
<td>64.00</td>
<td>97.67</td>
<td>42.00</td>
<td>21.61</td>
</tr>
<tr>
<td>35086</td>
<td>2.55</td>
<td>1.22</td>
<td>1.33</td>
<td>11.00</td>
<td>79.00</td>
<td>96.46</td>
<td>63.50</td>
<td>21.24</td>
</tr>
<tr>
<td>7038</td>
<td>2.73</td>
<td>1.02</td>
<td>1.72</td>
<td>12.00</td>
<td>95.00</td>
<td>96.69</td>
<td>48.00</td>
<td>82.38</td>
</tr>
<tr>
<td>36308</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>10.00</td>
<td>63.00</td>
<td>63.00</td>
<td>89.00</td>
<td>19.13</td>
</tr>
<tr>
<td>35106</td>
<td>2.27</td>
<td>1.13</td>
<td>1.14</td>
<td>26.00</td>
<td>70.00</td>
<td>79.00</td>
<td>47.00</td>
<td>19.12</td>
</tr>
<tr>
<td>32140</td>
<td>3.08</td>
<td>1.23</td>
<td>1.85</td>
<td>23.00</td>
<td>80.00</td>
<td>98.33</td>
<td>32.50</td>
<td>16.42</td>
</tr>
<tr>
<td>17340</td>
<td>3.47</td>
<td>2.19</td>
<td>1.28</td>
<td>11.00</td>
<td>44.00</td>
<td>96.46</td>
<td>66.00</td>
<td>15.59</td>
</tr>
<tr>
<td>30116</td>
<td>4.83</td>
<td>3.37</td>
<td>1.46</td>
<td>17.00</td>
<td>29.00</td>
<td>97.85</td>
<td>49.50</td>
<td>15.52</td>
</tr>
<tr>
<td>39004</td>
<td>3.75</td>
<td>2.64</td>
<td>1.10</td>
<td>4.00</td>
<td>30.00</td>
<td>79.33</td>
<td>77.00</td>
<td>13.55</td>
</tr>
<tr>
<td>44084</td>
<td>2.44</td>
<td>1.14</td>
<td>1.30</td>
<td>18.00</td>
<td>86.00</td>
<td>98.08</td>
<td>54.00</td>
<td>13.36</td>
</tr>
<tr>
<td>23280</td>
<td>2.08</td>
<td>1.06</td>
<td>1.02</td>
<td>22.00</td>
<td>93.00</td>
<td>99.00</td>
<td>61.00</td>
<td>13.15</td>
</tr>
<tr>
<td>38196</td>
<td>3.61</td>
<td>2.29</td>
<td>1.32</td>
<td>9.00</td>
<td>42.00</td>
<td>96.00</td>
<td>67.00</td>
<td>11.4</td>
</tr>
<tr>
<td>32004</td>
<td>2.38</td>
<td>1.25</td>
<td>1.13</td>
<td>3.00</td>
<td>61.00</td>
<td>76.00</td>
<td>74.50</td>
<td>10.27</td>
</tr>
<tr>
<td>21244</td>
<td>3.62</td>
<td>2.21</td>
<td>1.42</td>
<td>8.00</td>
<td>42.00</td>
<td>92.67</td>
<td>62.00</td>
<td>8.65</td>
</tr>
</tbody>
</table>
It is worth noting that DMUs operating in harsher environment have generally very poor performance in quality and in outcomes and are still deemed efficient due to the lack of DMUs to which they can compare to. This is the case of DMU 35106, for example, which reported low service quality and low outcome achievement. This DMU could only be compared to 43092 which reported a similar performance. Thus both are considered efficient.

Table 5.39 Peers in Node 4 for Total Field office operations

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Peer</th>
</tr>
</thead>
<tbody>
<tr>
<td>3060</td>
<td>54004, 30067, 43092</td>
</tr>
<tr>
<td>40050</td>
<td>5112, 43092</td>
</tr>
<tr>
<td>43129</td>
<td>42264, 5112</td>
</tr>
<tr>
<td>20060</td>
<td>5112, 43092</td>
</tr>
<tr>
<td>30176</td>
<td>5112, 43092</td>
</tr>
<tr>
<td>44084</td>
<td>42264</td>
</tr>
<tr>
<td>46372</td>
<td>5152, 43092</td>
</tr>
<tr>
<td>33181</td>
<td>5112, 43092</td>
</tr>
<tr>
<td>1019</td>
<td>42264, 5112</td>
</tr>
<tr>
<td>10228</td>
<td>5112, 43092</td>
</tr>
<tr>
<td>30140</td>
<td>42264, 5112</td>
</tr>
<tr>
<td>42296</td>
<td>8004, 43092</td>
</tr>
<tr>
<td>35086</td>
<td>8004, 5112</td>
</tr>
<tr>
<td>30176</td>
<td>8004, 36308, 43092</td>
</tr>
<tr>
<td>7038</td>
<td>8004, 7038, 5112</td>
</tr>
<tr>
<td>35106</td>
<td>46441, 43092</td>
</tr>
<tr>
<td>32140</td>
<td>8004, 43092</td>
</tr>
<tr>
<td>32442</td>
<td>10198, 5112, 36308</td>
</tr>
<tr>
<td>17340</td>
<td>8004, 5112</td>
</tr>
<tr>
<td>30116</td>
<td>8004, 5112</td>
</tr>
<tr>
<td>39004</td>
<td>42264, 5112</td>
</tr>
<tr>
<td>44084</td>
<td>8004, 5112</td>
</tr>
<tr>
<td>38196</td>
<td>5112</td>
</tr>
<tr>
<td>21244</td>
<td>42264</td>
</tr>
<tr>
<td>32004</td>
<td>42264, 5112</td>
</tr>
<tr>
<td>23280</td>
<td>8004</td>
</tr>
</tbody>
</table>

In the section that follows the significance of including the variable of environmental harshness for social service production is discussed.
5.4 A Note about the Inclusion of Environmental Variables in the Model

The importance of including the harshness variable “Ruggiero variable” in the formulations was tested by eliminating it from Model I using the pilot data set in its raw (volume) format. As all DMUs in the sample were allowed to be compared with one another, the discriminatory power of DEA was increased. Thus, nine DMUs were found under performing in node 1, as opposed to only two units deemed inefficient when the harshness variable was included as shown in Table 5.40 below.

Table 5.40 Performance Scores for Node 1 without Environmental Harshness

<table>
<thead>
<tr>
<th>DMU</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5132</td>
<td>1.00</td>
</tr>
<tr>
<td>5180</td>
<td>1.00</td>
</tr>
<tr>
<td>5264</td>
<td>1.78</td>
</tr>
<tr>
<td>5503</td>
<td>1.00</td>
</tr>
<tr>
<td>10060</td>
<td>1.00</td>
</tr>
<tr>
<td>12024</td>
<td>3.34</td>
</tr>
<tr>
<td>15080</td>
<td>1.34</td>
</tr>
<tr>
<td>22172</td>
<td>2.28</td>
</tr>
<tr>
<td>22322</td>
<td>2.84</td>
</tr>
<tr>
<td>25150</td>
<td>1.60</td>
</tr>
<tr>
<td>29042</td>
<td>1.00</td>
</tr>
<tr>
<td>35072</td>
<td>3.59</td>
</tr>
<tr>
<td>35168</td>
<td>1.80</td>
</tr>
<tr>
<td>38360</td>
<td>1.00</td>
</tr>
<tr>
<td>45010</td>
<td>2.72</td>
</tr>
</tbody>
</table>

However, a close observation to the characteristics of the peer units for each field office helps to conclude that the comparisons might not always be meaningful if there is no environmental harshness limiting the composite peer group.

For instance, DMU 5264 is under performing. The environmental characteristics of 5264 represent an aggregated household income of 70 billion and a percentage minority 39%. Its peers are 5503 and 10060. DMU 5503 operates in a wealthier community (aggregated household income of 129 billion with 46% minority). Under Model I this field office would have not been allowed to be a peer of 5264 even though their minority composition is similar, simply because...
they operate in different fundraising environments. The later under the premise that one cannot compare one DMU with another one with the probability of twice potential donations than the one under study and make of it its efficient peer. The opposite is possible under the argument that if a DMU performs well in worst operating conditions so it is possible for a DMU operating in better conditions to achieve similar levels of output.

DMU 10060 is the second peer of 5264. This field office operates in a community not as wealthy as 5264 (40 billion aggregated household income) but with much higher minority composition (69%). Therefore, this should be a meaningful peer for 5264. All other cases are similar in that at least one peer operated in a better environment for all inefficient DMUs. However, in general, the differences even if more than ten times in few cases, were always in peers with less weight. The peers with more weight were always units closer to the unit under analysis in both minority composition and aggregated household income. However, even when close in aggregated household income, in many cases the peers operated in slightly better environments. This poses the question of what a similar operating environment is and what is not. Thus, the boundaries of what is comparable or not need to be studied further.

Another problem that arises with this is when those peers are operating in a totally different environment, such as the case of 5503, which is one of the peers of 15080. DMU 5503 operates in an environment with 129.4 billion dollars while 15080 operates in an environment of 10.6 billion dollars. 5503 has a minority population that comprises 46% of its population and 15080 has a minority ratio of 8%. Therefore, they are not comparable units. A similar case happens with the other peer (10060) that operates in a community with 49 billion and 69% minority. However, the other two peers of 5503 are more comparable; particularly 29042 which is the peer with the highest weight (0.85) and operates in an environment with 2 billion dollars of aggregated household income and 2% minority.

Table 5.41 presents the peers for inefficient units in node 1. From the face validity point of view, presenting a DMU with recommendations for improvement based in the performance of so different units poses a problem. Of course, the environmental harshness constraint has per se another set of problems such as that it limits the composite set so much that unless the set contains a large sample with many DMUs, a number of them will be deemed

263
efficient by default. Furthermore, the DMU operating in the harsher environment will always be deemed efficient, as this DMU would not have any other DMU to be compared against. The results of Model I for all nodes are presented in Appendix E.

Table 5.41. Peers for Inefficient Units in Node 1 without Environmental Harshness

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Peer</th>
</tr>
</thead>
<tbody>
<tr>
<td>5264</td>
<td>29042,10060,5503</td>
</tr>
<tr>
<td>12024</td>
<td>38360,29042,5503</td>
</tr>
<tr>
<td>15080</td>
<td>29042,10060,5503</td>
</tr>
<tr>
<td>22172</td>
<td>29042,10060,5503</td>
</tr>
<tr>
<td>22322</td>
<td>38360,29042,5503</td>
</tr>
<tr>
<td>25150</td>
<td>29042,10060,5503</td>
</tr>
<tr>
<td>35072</td>
<td>38360,29042</td>
</tr>
<tr>
<td>35168</td>
<td>29042,10060,5503</td>
</tr>
<tr>
<td>45010</td>
<td>38360,29042,5503</td>
</tr>
</tbody>
</table>

The usefulness of having an environmental harshness constraint needs to be studied case by case. If the environment does not play a fundamental role or a factor that clearly invalidates performance comparisons among DMUs or if under different conditions it does not exist, it is better not to include harshness restrictions in the formulations.

5.5 Conclusions about Models

It is obvious from the illustrative applications presented that the value of Model I is more evident to the decision-maker since it uses real data at each node and comes up with immediate performance improvement recommendations. Conversely, Model II assumes re-engineering or improvements and it helps to determine potential savings or the potential revenue if the recommendations are followed. But at the same time, Model II uses actual data only for node 1 and does not include the actual data in the subsequent nodes. This can be seen as a
drawback because of the failure to account for actual performance. However, Model II assumes some dependency between nodes while Model I does not.

Considering the above, a well informed decision-maker would use both, Model I and II, and judiciously come up with plausible recommendations. One typical approach would be to use Model II at the top level management of the field offices for helping them visualize the field office potential and then proceed to Model I to dig into each node individually to locate and identify the places where there is excessive use of inputs and shortage of outputs or outcomes.

Because of the restrictive nature of the formulations, all models work better when the number of DMUs is relatively large. Also, the nature of the restriction according to the level of environmental harshness makes it necessary that the number of DMUs be large. However, the variables operating under a harsher environment will always tend to be efficient given the reduced number of DMUs in their sample. This is the main drawback of this model. A sensitivity analysis was performed with respect to the environmental variables and its results are presented in section 5.4. In almost all nodes, having both restrictions (the level of harshness and the internal non-controllable environmental variables) produced a set of peers that was more meaningful for nonprofit organizations largely impacted by their operating environment. However, the downside of this implication is of course that more DMUs will be deemed efficient and because of that, no performance improvement recommendations will be offered. The second problem is that DMUs that operate in very similar environments but not necessarily equal environments could potentially be left outside the composite reference set because the environmental harshness constraint, as it has been devised presently, excludes all DMUs that operate in a better environment even if this difference is numerical but not practical. This is perhaps the major problem for this formulation.

Addressing the issue of standardizing the data set before implementing the models, one needs to analyze the operation, the transformation process, the dispersion of the DMUs and other environmental factors or variables influencing the decision. This would help to understand if standardizing controls better for the spread of the data available, and therefore, DMUs that otherwise will be deemed efficient because of their scale of operations will be included in the peer set of others with smaller scales. Thus, the use of volume or indices needs to be evaluated on a case-by-case basis.
In the case of the agency, the models worked better with the standardized data for two reasons. First, these type of data enabled management to compare field offices in the same range of population groups against a standard — which was easily understood by everybody — that equates field office performance notwithstanding the level of operations. Secondly, standardized data gave more discriminatory power to the models.

Regarding validation of the models, this was accomplished through three different ways:

- **Nonprofit expert validation.** The *conceptual model* for nonprofit performance measurement in four areas: income generation (fundraising), capacity creation, service delivery and outcome achievement was presented in several nonprofit forums to a total of more than 100 nonprofit executives who have provided valuable feedback and validated the conceptual framework as one pertinent and applicable for nonprofit organizations. Other organizations, including the international development counterpart for the agency showed interest in the applicability of these models for evaluating country program performance thus illustrating its possible generalizability.

- **DEA Expert validation.** The *DEA formulations* were presented to the DEA track participants at the Institute for Operations Research and Management Science annual meeting held in Miami in November 2001. The feedback gathered from DEA experts was positive and encouraging. Similarly, the formulations and results will be presented at the North American Productivity workshop in June 2002 and at the DEA International Symposium in Moscow, also in June.

- **Decision-maker validation.** The *results* of Model I to the total field office operations were presented in a report and discussed in detail with top agency decision-makers from the evaluation unit at the national headquarters for the agency. These officials confirmed that those results
were both, valid—in most cases consistent with their own experience and their perception of the expectations of individual field offices—and also achievable (realistic).

Further validation of the DEA formulations needs to be accomplished through peer review after submission of technical papers to specialized journals.

The section that follows overviews the Excel Solver code used to run the DEA analysis.

### 5.6 Development of an Algorithm for the New DEA Formulation

The new DEA formulations presented in Chapter 3 were coded using Excel Solver™. The professional version of Excel Solver was used whenever the number of DMUs times the number of input/output variables exceeded 200. The codes are in Appendix D. The run time of the models varies between 10 seconds and two minutes approximately depending on the size of the data set and the number of variables. It can handle up to a total of 100-150 field offices.

### 5.7 Deployment to Field Offices: Performance Improvement Reports, Service Quality and Effectiveness Reports

The deployment of the performance measurement and evaluation system to the DMUs in any DEA study is per se a challenge. Explaining how DEA works and how the results were obtained requires specialized analytical skills. In our case, in order to validate the results and deploy the system to field offices it was necessary to design reports that would communicate the DEA results, the prescribed performance improvement recommendations and the potential benefits to the decision maker in a clear, short and appropriate manner.

An important part of the design of any performance measurement and evaluation system is the deployment and portrayal of data and information. A series of reports that comprise a type of Visual Management System (Clark, 1995) was developed to accomplish this.

The reports created for this research will be sent to the DMUs involved in this study. One of this report is the performance improvement report which is one page long and
contain a benchmarking score, lists field offices to whom DMUs can most meaningfully compare themselves (peer groups), concrete recommendations about where DMUs need to improve and by how much (performance targets) and the best practice frontier for field offices with similar characteristics. An initial draft with explanations of this report is included as Table 5.42

To visualize where DMUs stand in each one of the input/output variables used in the DEA analysis, a graphical trend analysis report like the one shown in the subsequent pages was created for each field office in the analysis. This trend analysis report portrays each indicator in a three-year trend against the results of the peers assigned by DEA to that particular DMU. The best-practice frontier computed by DEA is multi-dimensional therefore; it is difficult or impossible —depending on the number of inputs and outputs—to visualize these data with standard graphical software. The trend analysis report is a modest attempt to provide decision-makers, whom themselves have in many cases limited technical background, some sort of visual display of the indicators that inform the DEA results.

For the reporting of service quality and effectiveness indicators, a detailed graphical report is expected to be produced for each survey report. Each DMU will then receive as many service quality reports as DMUs collect data through the surveys completed by their clients. A sample of this report is shown in Figure 5.2. The left part of the report shown as Figure 5.2 shows the service quality construct in table and stacked bars format (for different cognitive styles). The right part of the report does the same for the effectiveness construct. The service quality index will be displayed and the decision-maker will be able to see the sources of good or bad service quality. At the bottom additional programmatic information collected through the surveys will be included, such as demographics, repeat clients, etc.
### Table 5-42. Performance Report for FIELD OFFICE A

Comparable peers given amount of resources, services and population size: Birmingham, Al; Columbia, SC; Dayton, OH; Wilmington, DE; Tulsa, OK

<table>
<thead>
<tr>
<th>Financial Resources Utilization</th>
<th>Reported</th>
<th>Desirable</th>
<th>% difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management and Fundraising expenses per household</td>
<td>$2.64</td>
<td>$2.64</td>
<td>0</td>
</tr>
<tr>
<td>Donations per household</td>
<td>$2.78</td>
<td>$2.98</td>
<td>$0.20</td>
</tr>
<tr>
<td>Institutional funds per household</td>
<td>$1.25</td>
<td>$3.25</td>
<td>$2.00</td>
</tr>
<tr>
<td>All other income (including products and services) per household</td>
<td>$1.8</td>
<td>$1.8</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Capacity Building</th>
<th>Reported</th>
<th>Desirable</th>
<th>Number that are not being utilized at their full potential</th>
<th>% difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered Volunteers per 1000 households</td>
<td>36</td>
<td>56</td>
<td>20</td>
<td>55%</td>
</tr>
<tr>
<td>Paid Staff per 1000 households</td>
<td>3</td>
<td>2</td>
<td>-1</td>
<td>-50%</td>
</tr>
<tr>
<td>DSV enrollees per 1000 households</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>500%</td>
</tr>
<tr>
<td>H&amp;S Instructors per 1000 households</td>
<td>25</td>
<td>35</td>
<td>10</td>
<td>40%</td>
</tr>
</tbody>
</table>

| Assets | | |
|---------------------------------|-----------|-----------|------------------|--------------|
| | | | | |

<table>
<thead>
<tr>
<th>Ivery indicators:</th>
<th>Reported</th>
<th>Expected</th>
<th>Expected % increase</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>H&amp;S Reach</td>
<td></td>
<td></td>
<td></td>
<td>The percentage of the population reached should increase by 68% to reach 5.5% of the population</td>
</tr>
<tr>
<td>Disaster Families assisted per 100,000 households</td>
<td></td>
<td></td>
<td></td>
<td>Given the amount of resources utilized and population size, field office can increase by 298% its disaster operations***</td>
</tr>
<tr>
<td>Program expenses per household</td>
<td></td>
<td></td>
<td></td>
<td>Field office should increase program expenses for Community Services, International and AFS in 400%</td>
</tr>
</tbody>
</table>

*** Field offices need to evaluate if opportunities exist to increase their service delivery by the amount recommended by the model. A better measure would be the % of local disasters served but that measure is currently not available.
Performance Management Report for Field Office B

CEO: J. C  
2000 Population: 787,000  
Staff: 44 F/T, 23 P/T.  
Board Chairman: J. M.  
Jurisdiction: 2 Counties  
No bases.  
Volunteers: 1,820 Registered, 11,690 Total.

Figure 5.1. Sample of Trend Analysis Report

Program and Services Expenses per Household 1998-2000

Total Donations per Household

Fundraising Expenses per Household

Diversity Gap in Leadership Volunteers

Health and Safety Reach

Families Assisted per 100,000 Households
Figure 5.2 Samples of Service Quality and Effectiveness Report

First Aid and CPR Training - Service Quality and Effectiveness Indicators

<table>
<thead>
<tr>
<th>Service Quality Indicators:</th>
<th>Effectiveness Indicators:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Excellent</strong></td>
<td><strong>Very Good</strong></td>
</tr>
<tr>
<td>Instructor presented information clearly</td>
<td>94.5%</td>
</tr>
<tr>
<td>Instructor was knowledgeable</td>
<td>100.5%</td>
</tr>
<tr>
<td>Level of outcome achievement: enrollees who are willing to provide emergency care</td>
<td>71.2%</td>
</tr>
<tr>
<td>Would recommend this course to a friend</td>
<td>88.7%</td>
</tr>
</tbody>
</table>

Service Quality Indicators

- Overall Satisfaction
- Length of the course
- Cost of the course
- Effectiveness in helping enrollees learn skills
- Quality of course books and videos
- Inclusion of needed information and skills
- Instructor was knowledgeable
- Instructor presented information clearly

Effectiveness Indicators

- Would recommend this course to a friend
- Feels comfortable responding to an emergency
- Feels prepared to provide emergency care using the skills learned in the course
- Feels confident identifying an emergency situation
- CourseTaken:
  - Feels confident identifying an emergency situation: 88.7%
  - Feels prepared to respond to an emergency: 55.1%
  - Feels comfortable responding to an emergency: 71.2%

Taking for Recertification

- Yes: 68.6%
- No: 31.4%

Languages mentioned:
- English: 67%
- Spanish: 5%
- Chinese: 5%
- Hindi: 5%
- Vietnamese: 5%
- Other: 11%

Gender of respondent:
- Male: 56%
- Female: 44%

Race:
- Hispanic or Latino (of any race): 4.2%
- White: 79.6%
- Black: 3.9%
- Asian: 1.8%
- Native Hawaiian: 0.1%

Age of respondent:
- 18 to 29: 13.8%
- 30 to 39: 22.8%
- 40 to 49: 19.0%
- 50 to 59: 26.1%
- 60 to 69: 14.9%
- 70 or older: 10.1%

Table:
- Enrollees who would be interested in taking this course if offered in another language:
- Yes: 88.8%
- No: 13.1%
CHAPTER 6 Conclusions and Recommendations for Future Research

This chapter presents the conclusions and recommendations emanating from this research endeavor. It is structured in four sections. The first section summarizes this research effort and points out important insights of this study above and beyond the mathematical terminology already presented in other chapters. It also points out the types of validity relevant to this research. The second section describes the major contribution of this research to three fields of study: nonprofit organizations, particularly charities dedicated to the delivery of social services; the performance measurement and evaluation field at large, and the Data Envelopment Analysis field of study. This section also includes some general concluding comments. The third section summarizes the limitations of this study. Finally, the last section outlines some recommendations for future research.

6.1 Summary

For service organizations in general, but particularly for nonprofits dedicated to the delivery of social or human services, determining the best way of evaluating their services is generally a very complex problem. Organizational performance is always a multi-dimensional concept. As such, some organizations can perform well in one area or dimension while others will be better in another. In the service delivery industry, however, measuring and evaluating the multi-dimensionality of performance gets even more complicated because of the types of performance dimensions involved. While the rational utilization of resources to deliver the highest amount of service units is important, the way service is delivered —which involves service quality and the subsequent satisfaction of their customers— is also crucial. Furthermore, for social services in particular, effectiveness, or the attainment of pre-determined outcome-based objectives is generally the main motivation for a human services agency to exist (Medina-Borja & Triantis, 2001). Thus, evaluating performance also requires involving a number of variables or indicators of service quality, efficiency and effectiveness. There is also a need to develop a way to integrate all relevant dimensions into a common measure or score in order to
provide managers with the ability to make decisions and benchmark their services against other comparable service providers.

To effectively deal with the multi-dimensionality problem of performance evaluation systems for services, an adaptation of a linear programming technique called Data Envelopment Analysis (DEA) (Charnes et al. 1978) was introduced. Such technique has the capability of computing performance scores and the performance frontier for different branches, units, chapters or other comparable decision-making units (DMUs) of social service organizations. This technique was developed to answer the question of whether the inputs are being transformed into outputs and outcomes in the most efficient way.

There have been few attempts to use linear programming approaches for performance measurement in the social services arena. The difficulty in developing measurement tools that account for all the variables or dimensions of service delivery makes DEA an appropriate option in order to address this extremely relevant problem for service organizations.

The objective of this research was to develop a performance measurement system that collects, evaluates and provides concrete performance improvement recommendations to the decision-makers in the nonprofit sector. Three dimensions of performance were identified in the literature as important for social services, namely, effectiveness, service quality and efficiency.

Effectiveness measures the extent to which the program obtains its goals and objectives fulfilling its mission statement (Schalock, 1995). A program is as effective as to the extent that it accomplishes what it was designed to accomplish—i.e. outcome achievement. Client outcomes of the greatest interest are those related to program goals and objectives (i.e. statements of what a program is trying to accomplish for its client population, Medina-Borja & Triantis, 2001). A six-point Likert scale was used to measure the client's change in perceptions, knowledge and behavioral intentions. The constructs defined for this research are proxy measures of service outcomes.

Service quality is a perceived judgment; resulting from an evaluation process where customers compare their expectations with the service they have received (Gronroos, 1984). Due to this subjective nature, measuring service quality poses difficulties for the service provider. SERVQUAL (Parasuraman, Zeithaml & Berry, 1988 and 1991a) is the most widely
known instrument used to elicit service quality constructs from customers. Some of the constructs measured by this instrument were used to develop ten service quality and effectiveness questionnaires for each line of service —some lines of service had several questionnaires for different services. A scale for service quality performance that includes timeliness, empathy, reliability and assurance and other constructs comparable to those on SERVQUAL was used in addition to hard measures of service performance and one specific item for customer satisfaction. All items measure ex-post perceptions and not expectations; except for one item in the Disaster Services questionnaire in which the expectation for service that the client had prior to a disaster event was elicited.

The most recent scholarly research on nonprofit organizations continues to stress the accountability issue, while recommends the measurement of organizational capacity and organizational effectiveness. A framework to measure performance in nonprofits in four stages or nodes was developed based on the most current scholarly and practitioner trends reported in the literature. The nodes represent the most important production functions for nonprofit organizations dedicated to social service delivery. These are:

1. The financial node (fundraising or income generation)
2. The capacity creation node
3. The service delivery node and
4. The effectiveness node.

Two basic DEA formulations were developed to represent the framework above: Model I is a four stage model that carries the actual values of output variables of one node to the successive node, and Model II is a formulation that carries the projections —i.e. the recommended targets— of the optimal values from one node to the other. This last formulation assumes that the DMUs have undergone a reengineering effort and that their indicators are at their maximum potential. All formulations are non-radial, have a preferential weight structure and include two types of environmental variables. One variable accounts for the harshness of the operating environment in which the DMU has to engage in income generation activities (fundraising, grants and contracts management) and service delivery. The second type of environmental variable seeks to ensure that other non-controllable factors that affect the
operating environment are considered for the performance evaluations. In this way, the reference set for a DMU under evaluation is such that the virtual DMU for an inefficient unit presents environmental factors equal to the one under evaluation. The constraints that express these environmental considerations restrict considerably the solution space. However, for large sets of DMUs its inclusion greatly improves the comparability of the peer set.

Data for process inputs and outputs other than service quality and effectiveness were obtained through the Information System department of the agency selected as the study site. Variable selection was performed according to a very structured process delineated in Chapter 4 for attaining the best results from DEA.

To demonstrate generalizability of the models and mathematical formulations, data for two different units of analysis were obtained. First, Model I and Model II were applied to 15 disaster services departments. Sensitivity analysis of the models regarding the use of two types of data formats: volume data and standardize (ratio) data was performed. Second, data for analyzing total field office operations (not only disaster but all lines of service and administration) was obtained for the three first nodes of 32 field offices. These field offices correspond to a category of field offices serving populations between 750K and 1.25 million. The service quality and effectiveness data was randomly generated since these field offices had not yet surveyed their clients at the time of analysis. Only Model I was applied to the later data set.

It was confirmed that the DEA framework works better in organizations with a large number of comparable DMUs, since the greater the number of DMUs the better the discriminatory power of these models. Also, through the discussions with agency officials and other experts in social service operations, it was established that for organizations with a lot of branches and social objectives it is more meaningful to use ratios – such as donations per household, or number of disaster volunteers per 1000 households. DEA helps develop the concept of best-practice frontier which informs at which level those performance standards should be met. For decision making purposes, it is also more meaningful to know, for example, that in the case of a major disaster the organization needs to be able to deploy X number of trained volunteers per Y number of households in the community. Therefore, the DEA framework becomes a planning and resource allocation tool as well.
The modifications to the traditional DEA approach allowed us to point out possible causes for reduced service quality, effectiveness and efficiency. Some of them can be attributed to lack of funding, insufficient human resources, etc. As social service nonprofit organizations continue to face new evaluation challenges and taxpayers and funding sources demand accountability and results to continue their support, the relevance of an integrated performance measurement system increases. This research effort in particular not only collects data but interprets it, transforms it into information and depicts this information for decision makers in usable format, but also provides concrete performance improvement recommendations that are based in proven analytical tools. Furthermore, the approach can also been used as a planning tool for resource allocation. With these results, decision-makers will be able to take more informed actions to improve performance on all fronts.

In closing, this research was able to accomplish the goal of developing and testing an integrated, multi-dimensional performance measurement system for evaluating performance in nonprofit settings. This pioneering DEA formulation has even been taken one step further by having had the unique opportunity for implementing it, collecting information and evaluating performance metrics at one of the largest national social service agencies in the United States. In the DEA community large scale implementation has not been done and the pitfalls and recommendations for success is an area that is pretty much lacking.

6.1.1 Validity and Generalizability of this Study

Any time a researcher attempts to translate a concept or construct into a functioning and operating reality (the operationalization) concerns about how well the translation was performed arise. This issue is relevant for both treatments and measures (Trochim, 2002). This section addresses the issue of validity for the present research study.

Let’s recall that this research has several components, namely (i) an operational and conceptual model translated into a (ii) mathematical model programmed as a linear programming solution algorithm; and (iii) a series of service quality and effectiveness constructs included in the surveys used to elicit data to feed the mathematical model. Hence, the issue of validity needs to be addressed from at least three different perspectives. I will examine first the overall validity of the conceptual and operational model used for measuring performance in
social service provision. Second, the actual validation of the mathematical model (formulations) and the verification of the algorithms or programs used to compute the results of the model. Third, the validity and reliability of the service quality and effectiveness constructs used for the survey instruments that elicit service quality and outcome data.

6.1.1.1. Content and Face Validity of the Study

According to Trochim (2002) both face and content validity are attempting to assess the degree to which the researcher accurately translated a construct into the operationalization. Trochim groups both content and face validity under the label “translation validity.” Translation validity focuses on whether the operationalization is a good reflection of the construct—it assumes you have a good detailed definition of the construct and that you can check the operationalization against it.

In content validity, you essentially check the operationalization against the relevant content domain for the construct. This approach assumes that you have a good detailed description of the content domain, something that's not always true. For instance, content validity was approached by this study by checking the validity of the three dimensions of performance for social service with an extensive review of literature in chapter 2. That review spelled out lots of criteria for the content that were relevant for the three constructs. For instance, the SERVQUAL dimensions and the types of input/output indicators relevant for social services. All three dimensions of performance, efficiency, service quality and effectiveness have been advocated by a number of authors in the nonprofit sector (Forbes, 1998; Letts et al., 1999; Light, 2000; Martin & Kettner, 1996, 1997; Newman et al., 1987; Nyhan & Martin, 1999; Patti, 1988; Sola and Prior, 2001). The fact that there exist four basic important nodes in nonprofit production so that these three dimensions should be measured against these four areas has been also validated extensively in the literature. All these have been reviewed in chapter 2 and 3.

Also, the service quality constructs were both, meeting the basic SERVQUAL dimensions for service quality (Parasuraman et al, 1993, 1994) and in line to previous research on service quality items internal to the organization.

In face validity, you look at the operationalization and see whether "on its face" it seems like a good translation of the construct. According to Trochim (2002), because of the
subjective nature of this type of validity, this will not be convincing to others and is usually qualified as the weakest way to try to demonstrate construct validity. However, as this author says, not because it is weak evidence it means that it is wrong evidence, as any researcher needs to rely on his/her subjective judgment throughout the research process. Trochim suggests that one can improve the quality of face validity assessment considerably by making it more systematic. This can be done by sending the constructs to a carefully selected sample of experts on the subject area and obtaining favorable judgements so that this type of validity gains more weight. As part of this study this type of assessment was performed. Regarding the face validity of the service quality and effectiveness constructs they were object of extensive evaluation and approval by evaluation experts in the Agency under study. They all reported back their judgment that the service quality and effectiveness measures appear to be a good measure of quality and desirable outcomes of the service. Outcomes were deemed as the most appropriate for each service.

Face validity was also performed on the results obtained by the mathematical model. For operations research models in general, face validity is the sometimes-subjective approval by the decision-maker of the model and is a very important step of model validation. DEA researchers are not different and the importance of validating the results is discussed extensively in the next section.

Another type of validity relevant to this research is predictive validity. In predictive validity, we assess the operationalization's ability to predict something it should theoretically be able to predict (Trochim, 2002). For instance, for the operationalization of performance in social services the three dimensions of service quality, efficiency and effectiveness input to the DEA engine did in fact depicted (ranked) the performance of the field units. The targets obtained were desirable potential targets in each one of the four nodes. There were no unexpected results. There is a high correlation between the performance scores and the outputs in each node.

6.1.1.2. Mathematical Model Validation

As explained in chapter 3, a model is an analytical representation or quantification of a real system and the ways in which phenomena occur within that system, used to predict or
assess the behavior of the real system under specified (often hypothetical) conditions. The process of determining whether a model is an adequate representation of the real system being modeled is called model validation. This is done by comparing the predictions of the model with observations of the real system (International Atomic Energy Agency, 2002). According to the IAEA there is some controversy about the extent to which model validation can be achieved, particularly in relation to modeling effects or phenomena expected to occur in the long term. Model verification, on the other hand, is the process of determining whether a computational model correctly implements the intended conceptual model or mathematical model. Model verification will often be a part of the broader process of validation.

In the case of this research, model validation has been performed in several ways. This section will not attempt to compare DEA with other potential techniques to compute efficiency/performance estimators but concentrates on the validation of the developed models only.

First, the formulations for Models I and II are extensions, adaptations or combinations of published DEA formulations for the treatment of special cases (Banker and Morey, 1986; Athassanopoulos, 1999, 2000; Ruggiero, 1996, 1998, 2000; Soteriou and Zenios, 1999). All are extensions from the original DEA formulations disseminated by the seminal papers of Charnes et al (1978) and Banker et al (1984). As such, the DEA technique is a robust and thoroughly peer reviewed model. To address the issue of the quality of the particular DEA model developed for this study—i.e. how suitable is the formulation developed to the task at hand and how good is it—Pedraja-Chaparro et al (1999) presented a set of criteria for validating DEA results. Unfortunately, of all four criteria presented only one can be quantified in real life applications. This one addresses the objective of DEA of identifying poorly performing DMUs and their relatively efficient peers. These authors assume that all DMUs are likely to be inefficient with respect to a true (albeit unknown) production frontier and suggest that a basic measure of the model’s quality would be the “proportion of DMUs deemed inefficient” by DEA (p. 639). All other three criteria refer to an absolute real efficiency and attempt to correlate the DEA results to those of the absolute true efficiency, which in most cases is impossible to identify and obtain. In that respect, even with a extremely reduced sample size, Model I and Model II were able to identify at least 10% of inefficient units even with the environmental harshness
restrictive nature of both models. This proportion increased substantially in the case of normalized (standardized) data.

According to Pedraja-Chaparro et al (1999), DEA researchers implicitly suggest the pre-eminence of robustness above other dimensions of model validation. Other authors have also reviewed the issues of validation and model specification (Boussofiane et al, 1991; Dyson et al, 2001; Golany and Roll 1989). In this study two different units of analysis and sets of data were used in order to demonstrate the applicability of the model to different situations and to show how robust the mathematical model was to different applications. The formulations worked well in both cases.

These same authors presented a number of potential model misspecifications found in the literature that DEA researchers should address. The principal causes of misspecification are:

1. Omission of a relevant factor (input or output) and or inclusion of an irrelevant factor. The factors used in the models have undergone extensive tests of statistical significance and face validity as described in sections 4.2 and 4.3 of this document. As such, the chances of having omitted a significant performance factor in social service provision are minimal.

2. An incorrect returns to scale assumption (for which several model specifications were tested in this research). This action of testing a variety of model specifications is relatively common in the DEA community.

3. And data inadequacies, particularly shortage of observations during the application. All these issues were analyzed before as the two applications used in this study has an illustrative purpose only (see chapter 4 and 5 of this document).

In all three cases of misspecification above, testing the face validity of the model with decision makers is a step for which the DEA community gives great importance (Dyson et al, 2001; Pedraja-Chaparro et al, 1999) particularly because DEA is mostly an “application driven theory” research field (Dyson et al, 2001 p. 245).
To verify the performance of the computational algorithm, simpler DEA formulations with published results were coded using Excel solver in order to check the comparability of results. Also, commercial DEA software\(^7\) was used in the BCC and CCR basic formulations and then with DEA Solver and the results were identical. These tests verify the applicability of Excel Solver to DEA but of course, no results were available from other software for the particular case of the formulations developed for this research. The results, however, were obviously correct by observation, as expected.

6.1.1.3. Survey Instruments: Construct Validity and Reliability

Regarding the quality of the measurement (as in the measurements obtained through customer surveys for this study) there are a number of validity types that are typically addressed:

**Construct validity** refers to the degree to which inferences can legitimately be made from the operationalizations in the study and how those relate to the theoretical constructs on which those operationalizations were based. Construct validity is the approximate conclusion that your operationalization accurately reflects its construct. Trochim (2002) divides the issues associated with construct validity into two categories that he calls the "land of theory" and the "land of observation." The land of theory comprises all of the ideas, theories, hunches and hypotheses that the researcher has about the world and your attempt to explain or articulate this to others. In the land of theory one finds the constructs of the measures or concepts that one believes is trying to measure. The land of observation consists of what you see happening in the world around you and the public manifestations of that world. In the land of observation the researcher finds his/her actual program, treatment, measurement or observational procedures. Presumably, he or she constructed the land of observation based on theory (Trochim 2002). According to Flyn and Pearcy (2001)

There are two broad ways of looking at the idea of construct validity. One assures construct validity by defining the construct so precisely that the researcher can operationalize it in a straightforward manner. Then, proponents of this approach believe that you have either operationalized the construct correctly or you haven't. The construct of performance in nonprofit

\(^7\) DEA Solver™
social service provision has been discussed extensively in the literature and comprises technical efficiency, service quality and effectiveness. In that, the operational model has been validated by theory. The service quality construct is also an adaptation of the SERVQUAL instrument, measuring the same constructs with slightly different items.

The second perspective believes that there are no black-and-white constructs as concepts are more or less related to each other (Trochim, 2002). The meaning of terms or constructs differs relatively, not absolutely. In that respect, a service quality measure might measure a lot of the construct of service quality but it may not capture all of it. Trochim (2002) recommends that in order to establish construct validity the researcher needs to focus on meeting the following conditions:

- Set the construct to be operationalized (e.g., service quality) within a semantic net (or "net of meaning").
- Provide direct evidence that the operationalizations look like what they should theoretically look like.
- Provide evidence that your data support your theoretical view of the relations among constructs. For example, relationship among service quality and customer satisfaction.

For this research, it is believed that the service quality construct chosen measures to a great extent service quality perceptions in each service questionnaire. This belief is based on a high correlation between the service quality items and the customer satisfaction item of over 0.90 in most cases. Also, based in construct validity as mentioned before (support from the literature) and the reliability of the constructs obtained during the pilot test (close to 0.9 in some cases). According to Flynn and Pearcy (2001) construct validity is the most important type of validity in surveys, over content and predictive validity. These authors also said that refining the constructs/scales using first the coefficient Alpha Cronbach and then applying factor analysis to make sure that the number of constructs equals the number of factors is a methodological procedure that strengthens the validity of the instrument.

In criterion-related validity, the behavior of the operationalization is examined given your theory of the construct. It assumes that the operationalization should function in
predictable ways in relation to other operationalizations based upon the researcher’s theory of the construct. The pilot tests conducted confirmed the behavior expected in the responses.

6.2 Contribution of this Research

This research study has contributed to the body of knowledge in several ways in addition to lending itself for an actual implementation of DEA in the real world. Specific contributions to the performance measurement field have been accomplishing as follows:

1. This research represents a unique attempt to combine areas that have not been fully explored in the performance measurement and evaluation field. Social workers and industrial engineers have priority had few opportunities to join their research efforts. It required that the researcher addressed issues in an interdisciplinary way.

2. This multi-dimensional performance measurement system will allow public and private social services agencies to evaluate their overall performance taking into account not only process measures but the impact that their services have on their beneficiaries.

3. The development of a performance measurement and evaluation framework for nonprofit organizations working in social service delivery is perhaps one of the biggest contributions. This framework collects performance indicators relevant for nonprofits, evaluates them, provides performance improvement recommendations, and reports those results to decision-makers.

4. The inclusion of important environmental variables that are determinants of service delivery in social service provision has been fundamental in this framework. An augmentation of the Ruggiero (1998) and Banker and Morey (1986) approaches to handle environmental variables is crucial for
the implementation of DEA in production systems directly impacted by these type of variables.

5. Also, this research represents one of the few attempts to incorporate outcome data into a DEA model. While DEA has been used to evaluate service quality, customer satisfaction and technical efficiency of public sector organizations, few studies have combined outcomes to calculate DEA performance scores. Also, the creation of tools to elicit outcome data is a contribution to the field.

6. The actual development of the DEA models to calculate scores of multi-dimensional performance for similar DMUs is per se, a valuable contribution to the field.

7. A first attempt to develop information portrayals (Visual Measurement Systems) to convey the output of the DEA analysis to its actual clients is also a contribution. This is probably one of the first times that DEA reports are actually implemented in the field. This required translating the DEA output into information, analyzing the real meaning of the results and the best way to actually present this information to decision-makers—who may not necessarily be mathematically skilled. An initial attempt to visualize the DEA reports in graphical way has also been attempted.

The overall conclusions for both models developed are shown in Table 6.1 below.
<table>
<thead>
<tr>
<th>Model</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Normal data</th>
<th>Standardized data</th>
</tr>
</thead>
</table>
| I     | • Uses real data at each node  
• Provides performance improvement recommendations based on the current situation.  
• Guides the focus of interventions.  
• Locates the areas in need of improvement.  
• Provides benchmarking at each one of the four performance areas: -financial or income generation; capacity creation; service delivery and effectiveness.  
• Clusters DMUs according to environmental harshness. | • The four nodes are treated separately. No interaction is assumed to generate the recommendations  
• Not good for ranking overall DMUs performance (over four nodes).  
• Harshness constraint is very restrictive. DMUs under harsher environments deemed efficient by default. | • Accounts for returns to scale.  
• Meaningful for a variety of organizations regardless of nature of operations.  
• Peers of inefficient units will be relatively close in the scale of operations  
• Very large units might be deemed efficient by default as the scale of operations will prevent DEA to find suitable peers for comparison. | • Possible to evaluate DMUs in a wide range of operations based on a meaningful overall standard  
• Should only be used in cases where the nature of operations puts meaning to normalized measures  
• Model has more discriminatory power  
• Each application problem needs to be analyzed separately to make sure that normalizing the data is meaningful  
• May not always be possible due to the type of data  
• Analysis needs to be performed with relatively homogeneous DMUs to avoid problems with returns to scale. |
| II    | • Provides a picture of potential savings and potential revenue for the inefficient DMUs after improvement targets have been achieved.  
• Serves as a planning tool to forecast future resource needs after re-engineering.  
• Accounts for the dependency between nodes.  
• Last performance score | • Assumes re-engineering or improvements have been done in all those DMUs that appeared inefficient in the previous node (which might be an unrealistic scenario).  
• Does not account for actual performance in Nodes 2, 3 and 4 | • Same as above | • Same as above |
Considering the above, a well-informed decision-maker would use both, Model I and II, and judiciously come up with plausible recommendations. One typical approach would be to use Model II at the top level management of the field offices for helping them visualize the field office potential and then proceed to Model I to dig into each node individually to locate and identify the places where there is excessive use of inputs and

6.3 Limitations of this Research

The limitations of this study are several and can be classified in at least four different realms: application limitations, mode limitations, data-related limitations; and, instrument-related limitations.

6.3.1 Limitations of this Application in the Nonprofit Sector

The present study was developed having nonprofit social service agencies as the unit of analysis. As it turns out, DEA requires a relatively large number of DMUs in order to have robust discriminating power between performing and underperforming units. A common DEA rule of thumb is to have DMUs at least twice the number of the inputs and outputs in a determined node. Others believe that best results are achieved with at least 40 DMUs. In several instances, many nonprofit organizations don't have access to data from that many comparable DMUs, mainly because most agencies outside of larger federations of agencies rarely share performance data. The following is a summary of the limitations of the application to the nonprofit sector:

- A key feature in order to have robust discriminatory power to evaluate data requires at least 40 decision making units. Since there are very few social service organizations with more than 40 projects, programs, departments, field offices or any other comparable and homogeneous unit of analysis, the performance system proposed will have trouble being applied in most organizations for internal benchmarking.
• External benchmarking in the nonprofit sector is very limited since nonprofits only disclose publicly very limited and sometimes incomplete internal performance information.

• The technological sophistication required for the level of analysis suggested in this study might make it very difficult for many organizations not used to measuring performance to make readily use of this tool.

• Not all nonprofits, operate in the production model suggested by the conceptual model in chapter 3. However, the DEA framework is flexible enough to accommodate variance for other nonprofit settings.

6.3.2 Mathematical Model Limitations

Mathematical models attempt to represent reality in mathematical terms. The DEA formulations developed in chapter 3 are not exception. As such, this representation of reality has limitations.

- The discriminatory power of the model is significantly reduced when the number of DMUs is small (rule of thumb usually is less than 40 DMUs). As such, very few organizations can apply it.

- Not all nuances of performance are included and some important aspects might be left out.

- Environmental considerations greatly reduce the potential of DMUs to find other units that can be considered efficient peers.

6.3.3 Data limitations

The data used for this study had several limitations. The most important are:

- The number of DMUs available without missing data was very reduced (less than 20 for Disaster operations and 31 for total field office operations).
• Not all field offices had a sample size adequate to make valid inferences about service quality and effectiveness. As such, the data was taken as illustrative purposes only.

• Although field offices were provided with procedures for random sample, it is not guaranteed that all field offices in this study applied the surveys in an adequate way.

• The operational data is reported annually. Mistakes in the data are probably but unlikely.

• DEA cannot handle missing data and as such, some DMUs were deleted because of a variable with missing data.

6.3.4 Limitations of the Survey Instruments

The instruments were pilot-tested extensively and the results very satisfactory. Nonetheless, they have several limitations:

• In the service quality side of the surveys, some researchers might argue that a 6-point scale has limitations and that some of the point labels are not equally spaced. As explained in chapter 4 other researchers argue against the existence at all of labels that represent equally spaced concepts.

• The service quality construct measures perceptions and not expectations.

• In some questionnaires the outcome side of the surveys operationalize some outcomes with only one item.

• Outcomes were operationalized as immediate outcomes (changes, perception, intentions, perceived knowledge). Intermediate and long term outcomes were not included.

• Space was a limiting factor as the agency did not want to use questionnaires with more than two pages letter size.
6.4 Recommendations for Future research

The current research can be extended and further investigated with respect to one or more of its components, namely, performance measurement theory, service quality, and innovative approaches in DEA.

With respect to service quality and performance measurement theory, new ways to combine the data collected through survey instruments into a single service quality indicator need to be investigated. The suitability of such indicators to DEA analysis also needs further research, particularly as to how most meaningfully include them so that the DEA recommendations with respect to those variables can be better interpreted and applied.

Also, the DEA results in this research study were not suitable for correlation analysis since the sample of DMUs was very small. Therefore, any statistical inferences were neither significant nor appropriate. However, it is of extreme interest to gather a larger sample of DMUs to run these models to see if it is at all possible to infer conclusions as to the direction and strength of the relationships among financial strength, operational capacity, efficiency, service quality and effectiveness in the nonprofit sector. This can be done by correlations, as follows: efficiency scores to the different performance indicators, efficiency scores from each node to each other and performance indicators to each other.

In the service quality arena, the service quality and effectiveness data collected is very rich (almost 6 thousand respondents in total) and suitable for a number of important analysis as to the relationship between the different dimensions of service quality and the final rating of customer satisfaction; service quality and outcome achievement (effectiveness). Also, research related to the development of new service quality constructs more suitable for social services.

With respect to the augmentation of DEA for social services and nonprofit sectors in general, the inclusion of environmental factors using the Ruggiero (1998) approach requires that the cut off for the DMUs in each frontier to be not as radical. The discriminatory power of the formulation is dramatically affected by this cutoff. Different mathematical approaches can be
explored to solve this problem. A moving range or an intelligent algorithm that allows the inclusion of DMUs within a certain range of the cutoff can be one solution. The algorithm would make the judgment as to whether the operating environment of the next DMU ranked is still relatively the same as the one being evaluated or not and therefore could whether be included or not in the analysis.

A fuzzy math formulation that allows the inclusion of all DMUs in the sample but with different membership function as to the environmental level cutoff needs to be investigated. Fuzzy formulations have been developed to cope with different DEA problems (e.g. Girod, 1996; Kabnukar, 2001; Sheth, 1999). This particular fuzzy approach would assume that the crisp cutoff assigned by the Ruggiero model could be translated into imprecise ranges for the harshness variable that could be assigned by the decision-maker. These fuzzy cutoffs will associate membership functions as to where a DMU can be included in the composite reference set of the DMU under evaluation.

Another solution for incorporating environmental factors in DEA evaluation could be to cluster the DMUs with statistical techniques using all environmental variables and then running DEA only within the clusters. The viability of this or other approaches needs to be investigated.

While DEA is a useful, innovative and powerful performance evaluation tool, it is still not widely used in the real world, as it should be. This is partly because of the inherent difficulties in explaining this technique, as well as its results, to lay—non-mathematically proficient—people. DEA visualization and reporting techniques need to be further researched by combining computer science and human factors research for the development of user friendly tools that could expand the understanding of DEA among decision makers in all types of organizations.

Finalizing, nonprofit and public sector performance measurement continues to play an important role. Perhaps as important as in the early 90s.. Social service agencies usually have program evaluation departments or contract their programs evaluations to show that their programs have had the desired effects. This is the mind set of most agencies. However, this system does not show causality and does not guarantee that the data was collected in the most appropriate way. It is a managerial performance improvement framework that borrows some
program evaluation concepts and constructs relies on the trust and eager willingness of agencies to improve performance and show accountability, stewardship and effectiveness.

These differences might be not fully understood in the nonprofit sector and as such, it is possible that many organizations are still not ready to face the challenge of implementation of a system that they do not appreciate and do not understand. Or it can be that such a system is not appropriate for them. Research is needed to devise a performance evaluation framework suitable for those agencies that does not fit the DEA profile discussed in the limitations section.
Bibliographic References


Governmental Accounting Standards Board. (1994). Concept Statement No.2 of the Governmental Accounting Standards Board on concepts related to service efforts and accomplishments reporting. Norwalk, CT.


International Atomic Energy Agency (2002) "IAEA’s glossary pages”
http://www.iaea.or.at/ns/CoordiNet/safetypubs/iaeaglossary/glossaryhomepage.htm


Stamatis, D.H. (1996). *Total quality service: principles, practice, and implementation.* Delray Beach, Fla.: St. Lucie,  


Todd, Rebecca and Ramanathan, K. (1994) “Perceiving social needs, outcomes measurement and budgetary responsiveness in a nonprofit setting: some empirical evidence” *Accounting Review*, v69 n1, Jan 1994, p.122  


Appendix A: The Outcome-Based Objectives Development Process

This tool is based on the United Way requirements for outcome evaluation of funded agencies. I have modified this approach to include two additional steps to assure the evaluation of measures obtained against the desired outcomes. The cycle begins with the internal evaluation of the mission and vision of the program and ultimately compares the obtained outcomes with the desirable outcomes. Desirable outcomes might be stated in the form of outcome objectives for the program or service. These outcome objectives undergo periodic revision as a result of the control loop established. This strategic revision of the program can be better appreciated in Figure A-1.

A Nine Step Approach to Develop a Performance Measurement System for Social service organizations

The following is a description of this nine-step approach to develop outcome objectives for social service organizations depicted in Figure 1 (Triantis & Medina-Borja, 1996; Medina-Borja, 1996.)

Step 1. Review the Agency and Program Missions and Relate them to the Services
An outcome objective, to be valid and credible must be linked to and supported by the agency and program’s mission. Outcome objectives defined for a program must have a direct relationship to both the agency’s mission and to the individual program’s mission and goals. In any case, to be consistent, an agency cannot claim an outcome objective totally extraneous to its mission.

Step 2: Identify the Needs Associated with Each Service
Each program and service must respond to an existing need or problem in the population it serves in order to provide a real benefit to its clients. Sometimes an agency may have been running programs for a long time. The agency may have changed as well as its client’s needs. However, the way the service is provided and its goals and objectives may have not been revised. In some cases, the need is not there any longer. An agency needs to be capable of clearly identifying such problems and needs and able to explain how its programs and services are helping the population to overcome them.

Step 3: Define Goals for Each Service and relate them to the Agency & Program Missions
Officials define the program’s goals for each service provided. In many cases, agencies already have a goal for the services they provide. In that case, that goal is used. The program’s goals should state the motivation for providing a service --what do we want to achieve by providing this service. Agencies should identify local needs and set up local program’s goals. Because sometimes services have different clients and solve different problems, it is better to develop at least one goal for each service. This group of goals will therefore be the program’s goals.
**Step 4: Understand & Identify Each Service Delivery Process**

A clear identification of each service’s delivery process is needed to later identify valid outcome and quality indicators. A Deming Flow Process Diagram is developed at this stage. The objective of this step is to get a better understanding of the service delivery process by identifying:

- Stakeholders or Clients (From Step 2)
- Process or Value Adding Activities
- All the resources (Inputs), utilized during the process and the cost associated with it (when available).
- All the products (outputs) resulting from the process and the number of units of service produced per unit of time (when available). Each unit of service should solve or address at least one client need in order to be meaningful.
- Process performance results of the service delivery process. These results will serve as the service’s outcomes when the purpose of the service is not to generate a change in the population served but to provide an infrastructure. Process Performance Results are the measurable and qualitative results of the service delivery process. Some of the indicators collected in a Customer Satisfaction Survey that measure the quality of the service can be considered performance results of the process.

**Step 5: Identify Service’s Outcomes**

In this step all information identified in the previous steps is gathered to define outcome objectives for the service. This step asks the following questions:

1. Who or what will be changed? Who or what will be impacted?

2. After the service is finished: What is going to happen with your client as a result of the service? What type of effects/changes will occur?

The effect of a service on an individual can vary over time. Immediate and intermediate level outcomes are generally easier to identify and to measure.

The results obtained from steps 4 and 5 —inputs, outputs, process performance results and service outcomes— must be linked in a Logic Model in a causal manner.

**Step 6: Identify Indicators for your Outcomes**

Once service’s outcomes are identified, then agencies need to think about a way to measure the outcome or the effect caused. The purpose of having “indicators” is to determine to what extent the program is accomplishing what it was designed to accomplish. Since many outcomes are abstract concepts, such as “alleviating human suffering” or “improving ability to find a job,” indicators where measurement is possible need to be found. To get this type of information agencies may need to ask the client —to find out if the family has in fact changed, if his/her suffering has been alleviated, if the likelihood of violence or other types of family crises have been reduced. To do that, a measurement tool, or instrument will need to be designed. An initial approach is to examine the tools that agencies already have and ask which, if any, will answer the question: How much change will occur? To what extent? and When will the change take place?
A variety of tools and approaches may be currently available. Some of them are customer surveys, interviews, client records, public county’s records, etc. Indicators must be reliable, well founded (valid) and acceptable (credible). After defining them, agencies need to check if the metric relates to the outcome and to the program’s goal. To be useful, indicators should help evaluate the program, not only define some characteristics of it.

**Step 7: Develop Program’s Outcome Objectives**

At this point all the information necessary to write outcome objectives has been gathered. Agencies should do this for each service separately since all services are different and have different goals. In general terms, the term objective is a measurable achievement over a period of time (normally one or two years) that when achieved, will denote progress in reaching your goal. A good objective will state who or what will change; by how much; and by what date. Because each service addresses different needs and each has different clients, each service has different outcomes and several outcome objectives. This is resource consuming. For this reason, it is recommended to select a maximum of two to three outcomes for each service.

After this exercise, agency’s officials might want to return to the strategic planning group and suggest changes to the missions and goals that will reflect the positive effects of your services. In some cases, when the source of data is not obvious, it is recommended to add a third element to the outcome objective: a description of the data collection tool or an explanation of how you plan to collect the data to measure performance.

Also, in this step a measure of performance that utilizes the identified indicators to measure outcome achievement needs to be developed. These measures go beyond the existing performance standards and measures. In this step is where other metrics, related to other dimensions of performance such as service quality and technical efficiency, could be introduced. Then, an outcome objective will answer the following:

- What will be change on the client as a result of the service?
- To what extent?
- By how much and in what period of time?
- What are the characteristics of the procedure used to achieve this outcome? (Timeliness, customer satisfaction level, methodology, etc.)
- What are the levels of efficiency (outcomes/input and outputs/input) planned by the agency for the achievement of such outcomes and how these compare with the levels actually obtained?

**Step 8: Collect and Analyze Outcome Data from clients**

The use of public data has many limitations related to the link between the indices to the service provided. Even if public data is available, the problems targeted by some programs are so specific that only in rare instances it will be possible to find published indicators that measure what we want to assess. To prove that the desired outcome is related to determined program, the agency has to design and apply a specific measurement instrument that will track the changes in the actual clients who are receiving the program’s services.

This step comprehends three important elements:
1. Develop an instrument to collect data (design, test, and refinement)
2. Collect Data
3. Develop a tool (statistical, linear programming, etc.) to analyze and benchmark the data
All three elements are of great importance and present a great deal of difficulty for most agencies. A good instrument will test the clients before entering the program or before the service is provided and will monitor their progress after the program or service has been received. If the objective of the programs is to provide services for clients in an ongoing fashion, the provision of infrastructure is the service and outcome data here is more related to customer satisfaction. In these cases, measurement tools such as surveys and client interviews should be designed to collect appropriate data.

**Step 9: Compare measured outcomes to desired outcomes**

The last step in the outcome-evaluation cycle is to compare the obtained outcomes to the planned or desired outcomes. This is done through the revision of the agencies outcome objectives and comparing the performance targets stated for outcomes, service quality and efficiency. When a program or service is not attaining its objectives two paths are possible:

1. To modify elements of the process used (changes in the conceptual or behavioral model, changes in the way the service was provided, such as timeliness, demeanor of the change agent, etc.).
2. To modify the outcome-objectives and service goals to reflect the reality.

Both actions are radical steps because, first, it is difficult for the service provider to realize that a procedure he or she believes to work has its pitfalls, and second, because the power to modify goals and objectives resides very high in the agencies hierarchy. A commitment at all levels of the organization to do something and to continuously improve the service delivery system is an unavoidable condition to successfully implement outcome-based evaluation as a PDSA cycle. The whole concept of performance measurement and evaluation is non-sense if agencies are not willing to recognize problems and take steps to change. The process outlined here is just an instrument that guarantees that data is used for decision-making.

Now that I have described all steps of the OBOD process I can provide an theoretical justification for this methodology by first establishing a common ground for planning improvement processes. I will also compare the outcome-based process to other specific strategic improvement planning methodologies.
Figure A-1. A Nine-Step Process that Facilitates the Definition of a Program’s Outcome Objectives and the Revision of Organizational Guiding Principles. (Adapted from Triantis & Medina-Borja, 1996)
Appendix B: Service Quality and Effectiveness Disaster Services Final Survey and Guidelines for Sampling and Survey Administration
SERVICE QUALITY QUESTIONNAIRE
Disaster Services

The is interested in learning about your experience in receiving Disaster Services. One member of your household should fill out this questionnaire on behalf of all affected parties. Please help us by answering the following questions.

Please completely fill the bubbles next to your answer(s). Use black or blue ink only.

EXAMPLE: Will you fill in bubbles completely?
- YES
- NO

GENERAL INFORMATION

For questions 1 and 2 fill only the one bubble that best represents your answer.

1. How did you first learn about this service?
- On Site (worker)
- Fire/rescue squad
- Police
- Through a friend or family member
- On TV
- Do not remember
- Other

2. Where did your family spend the first night after your disaster occurred?
- Own home
- Family or friend’s home
- Hotel/motel paid by the f
- Hotel/motel NOT paid by f
- Other

3. In what ways did the help you? (FILL ALL THAT MAY APPLY)
- Food/restaurant meals
- Clothing
- Lodging - hotel costs
- Lodging other than hotel - rent, shelter, other
- Medicines, eyeglasses
- Furniture/appliances
- Names and phone numbers of other groups that can help
- Counseling
- Information about preparing my family and home for future disaster
- Did not help
- Other

ABOUT THE SERVICE YOU RECEIVED

4. When did the first get in touch with you after your disaster? (FILL ONLY ONE)
- Within 2 hours
- 3 to 4 hours
- 5 to 24 hours
- 25 to 48 hours
- More than 48 hours

5. Where did you FIRST get help? (FILL ONLY ONE)
- On scene
- By phone
- At Red Cross office

6. Was the help you got from 1...
- More than you expected
- What you expected
- Less than you expected
- Did not know what to expect

Comments:

Would you contact this service if you were to need further disaster help in the future?
- Definitely would
- Probably would not
- Probably would not
- Definitely would not

PLEASE CONTINUE ON BACK ——
3. Thinking about your experience with the assistance services on a scale from 1 to 6 (with 1 being poor and 6 excellent), please rate each of the following.

<table>
<thead>
<tr>
<th>Ease of Contacting</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Does not apply</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helpfulness of a Worker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>worker's knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timeliness of Assistance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convenience of Assistance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usefulness of names and phone numbers of other groups that can help</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helpfulness in beginning your recovery back to your normal life</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effectiveness in meeting your most immediate needs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall satisfaction with disaster help</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ABOUT YOUR READINESS TO FACE A DISASTER

9. How prepared are you? Before the disaster.

Yes | No | Not Sure | Does not apply
---|----|----------|----------------|

Every floor of my home had at least one working smoke alarm.

My family had a plan for emergencies including ways to escape and an agreed-upon location to meet outside the home.

My family practiced our emergency plan.

My family had a disaster supplies kit (e.g., an emergency kit, a battery-powered radio, a flashlight with extra batteries, bottled water and food).

10. Anything else you want to tell us about the services you received from the provider is not sufficient.

feel free to attach a piece of paper if the space provided is not sufficient

ABOUT YOURSELF

The survey wants to know more about you and your specific needs to better tailor our programs and services to your communities. Please provide us with the following information:

Number of people living in your home (include yourself)?

1. You 2. You and spouse 3. You and spouse or 2 or more

Did you own or rent your home?

Did you own or rent your home?

O Own O Rent O Other

Your home was a:

Single Family Home

Did you own or rent your home?

Did you own or rent your home?

O Own O Rent O Other

Your home was a:

Single Family Home

Did you own or rent your home?

Did you own or rent your home?

O Own O Rent O Other

Your home was a:

Single Family Home

Did you own or rent your home?

Did you own or rent your home?

O Own O Rent O Other

Was your home insured?

Yes | No | Do not know
---|----|------------|

Chapter Code (Please fill in the bubbles)

0 1 2 3 4 5 6 7 8 9

Type of disaster:

Fire

Flood

Hurricane

Tornado

Earthquake

Other

Date Disaster Occurred (month and year only)

Month / Year

Thank you!
Data Collection Logistics

We have developed ten different surveys to reach the major client groups field offices serve.

Current thinking suggests that field offices are best suited to present clients with these surveys, either in person or via mail.

These standard questions have been developed with intensive field office, lines of service and customer input and have been shown to correlate strongly with satisfaction and with customer behaviors that will strengthen The agency services (for example, favorable recommendations or repeat business). These questions will provide you with the ability to make changes or improvements in the way services are delivered.

The most important points outlined in this document have been summarized at the end on the "Points to Remember" section. Please detach or copy that section and give it to the assigned technical liaison personnel and/or to the person in charge of making the copies and distributing the surveys. This is extremely important. Also, make sure your technical liaison reads this guidance. This will facilitate the process at both ends, field office and national headquarters.

For all surveys, field offices need to copy the two-page questionnaires included in the package as a ONE double-sided page. PLEASE DO NOT OVERLOOK THIS STEP. Clients will appreciate a one–sheet questionnaire and this format greatly facilitates the processing of forms at national headquarters. This step is extremely important, since it will enable us to process a great volume of questionnaires with the help of technology.

Also, please make sure that the “FOR FIELD OFFICE USE ONLY” section is appropriately completed by your field office personnel BEFORE the questionnaires are distributed to clients and enrollees. Assign a person to do this job, and instruct him or her to write as carefully as possible inside the boxes, avoiding touching the sides of the boxes with handwriting. This is the only means to assign correct results to your field office.

You will also need to prepare the mailings (client labels, self-addressed envelopes, etc.). Clients in turn will mail them with a self-addressed envelope to national headquarters, in Falls Church, Va. You can also ask your clients to send the surveys to you, and you will collect them and mail them to national headquarters. The mailing address is the same shown below.

For surveys associated with a course or presentation, the instructor or presenter will collect all surveys, bundle them and return them to the field office. Field office officials should wait until they collect at least 50 surveys from each type of course, and then mail them together to national headquarters.

For all surveys, the mailing address at national headquarters is

Administration and Evaluation
Field office Services Network
Attn: Alexandra Medina-Borja

The following sections describe the data collection logistic issues related to the different lines of service. Each of these agency services presents different issues in contacting respondents and obtaining service quality and effectiveness measures.

Disaster Services:

Disaster Services clients can be difficult to survey. Previous experiences with surveying disaster clients have taught us that unless you contact these clients immediately after the service is completed, you should expect to reach only about 20 percent of them, with many surveys returned as undeliverable. We suggest that you contact your clients as soon as possible. Include a copy of the disaster services questionnaire in any follow up materials that your
**field office personally distributes to them.** Include a self-addressed envelope and mention to them that whenever they feel ready, the Agency will greatly appreciate their comments regarding the service. This will give them time to recover before feeling compelled to respond to our survey.

**When giving the questionnaire to your clients, stress the questionnaire is anonymous and that we will use their opinions only to enhance and improve our services.**

Another concern is that some disaster clients may be recent immigrants or others may have limited English-language skills. This may limit the number of responses you get back unless you use surveys translated into the clients’ language. If a translation other than Spanish is needed for a reasonable number of clients, contact Administration and Evaluation at the Field office Services Network.

Field offices with few Disaster Services will probably need to give a survey to all clients to get back a reasonable number for drawing valid conclusions. A report will not be produced by national headquarters until you have received at least 50 surveys, because at least that many are needed to produce reliable results. However, in cases of very small field offices with less than 50 disaster operations per year, a report will be produced and assumed it is valid with less than 50 responses.

**Armed Forces Emergency Services:**

Heads of family who initiated AFES emergency services need to respond to this survey. Expected response rates for emergency communications clients are in the order of 30 percent. These clients also can be difficult to contact. Some move soon after a death. Others may have placed the emergency communications request from the home of a relative. The same guidelines as above need to be followed. **Whenever possible, give the survey personally to the client as discussed above under Disaster Services.**

**Health and Safety Course Takers:**

Five Health and Safety service quality surveys have been designed as follows:

- First aid, CPR and AED
- Aquatics – Lifeguard training
- Aquatics - Water safety
- Caregiving
- HIV/AIDS education

All five questionnaires have a special box at the end marked “FOR FIELD OFFICE USE ONLY.” You will need to fill in the information required in those fields, especially the field office code, date and type of instructor (volunteer, paid or third party). We ask participants to write the name of the instructor as well, so that recommendations to improve course delivery as well as praise can be directed to the appropriate party.

The questionnaires are to be handed to all course participants after the end of each course or presentation. Participants need to hand the completed survey back to their instructor, who then needs to deliver all completed surveys to the field office. Field offices in turn need to send the package to national headquarters. Before mailing them back to national headquarters, make sure you have at least 50-100 forms for each one of the four H&S areas for which we have different survey forms.

If after a few weeks your field office would have thousands of surveys, use the sampling technique discussed in the next subsection.

**Community Services**

One general survey for community services has been developed. Given that field offices nationwide offer a wide variety of Community Services, developing only one standard questionnaire was a difficult task. Some questions
needed to be general because of this. Each client needs to interpret the question in light of the type of service received.

Give the survey to all clients or head of family who receive a service classified as community services (transportation, meals-on-wheels, shelter, daycare, etc.). If your service is not classified in the first question, make sure your clients fill the “other” category and write down what the service is called.

Some clients in this category of services have learning disabilities, speak a foreign language or have visual and/or writing difficulties. Do not survey clients for whom reading and completing the questionnaire will be a big burden.

Kindly ask your clients to take the survey home, fill it out and return it to you (or mail it directly to us) whatever the field office feels is more cost-effective. The goal is to get the best possible return rate.

The procedure to be followed in this case is the same as the one for Disaster Services, except that you can actually HAND the survey to your clients at the time of the service (e.g. after pick up, or after the meal is served, etc.). Again, if after a few weeks your field office would have thousands of surveys, use the sampling technique discussed in the next subsection.

**International Services**

This package includes two surveys for International Services:

- **Tracing and other International social Services** (to be sent to customers who initiated tracing, The agency Message, or other International Social Services). Send the survey to all customers who initiated a tracing case or The agency message and live in the United States (even if those cases are still open or the person was not found).

The procedure to be followed in the case of Tracing and other International Social Services is the same as the one for AFES emergency communications clients. The difference is that clients who initiated a tracing service might have not yet received notice of conclusion of the case. The survey addresses this issue and asks clients’ expectations regarding length of service. **Again, whenever possible give the survey personally to the client as discussed for Disaster services.**

- **International Services courses/ presentations** (to be handed to participants at International Services courses or presentations), including:
  - Orientation to International Services course,
  - International Humanitarian Law course,
  - Tracing and Other International Social Services course,
  - International Relief and Development course, and
  - International Services presentation.

In the case of International Services courses and presentations, the procedure to be followed is similar to the one for Health and Safety courses. Instructors should complete the “For The agency Field office/ Instructor Use Only” section. Instructors will hand the questionnaires to participants, who in turn will complete and return them to the instructor who will package all returned surveys in an envelope and deliver to the field office. Since the frequency of scheduled international services courses and presentations is generally much lower than for Health and safety courses, field offices need to mail surveys to national headquarters immediately after a course or presentation has been completed. They do not need to accumulate 50 surveys. Any number of participants will be accepted since the numbers for these courses tend to be smaller.

**Number of Customers to Contact**

How many customers will you need to contact to obtain 100 completed responses?
Despite their relative convenience and low cost, mail surveys typically have lower response rates than other survey methods. Begin by assuming that you will have to mail out three or four questionnaires to get one response back. As time goes on, you can correct this estimate to reflect your experience.

For forms that are distributed immediately following a class or presentation, you may receive 80 percent or 90 percent of the forms if customers are not in a hurry and if someone is assigned to collect the forms.

Responses from 200 to 300 people for each service delivery survey will be adequate for most field offices with a population over 50,000. Field offices serving smaller populations (under 50,000) may use smaller samples and should aim for a minimum of 50 completed surveys of each type before sending them back to national headquarters.

This is done to have an unbiased result of the quality of your services. Having few respondents can skew the results to either side of the distribution. For instance, if you have only one respondent and this client rated your service as either “excellent” or “poor”, then there is no way to know on average what the true rating of service quality is at your field office. The results from this sole client will prevail as true for all clients. The larger the number of respondents, the more representative your results would be of the whole population.

Large field offices should aim to have at least 100 usable surveys for each type of survey, although more would be a better number for each fiscal year. These numbers translate into at least 800 clients surveyed across all lines of service, a number that represents a good sample from which inferences about service quality and effectiveness can be made.

If you mail the surveys, and your The agency field office serves fewer than 300 customers, you will probably need to ask for information from all customers. If your field office serves more than 300 customers, you should obtain completed responses for at least 100 customers. The table in the next section and the flow process charts will help you discriminate your sample size.

**Sampling**

**For field offices with fewer than 300 clients/enrollees in any given service:**

If you have fewer than 300 clients in Disaster, AFES or Tracing Services, then survey all your clients (census).

If you have fewer than 300 enrollees in any given type of course or presentation, then survey all your enrollees until the end of the fiscal year and send the total number of surveys collected. Hopefully, it will be at least 50 of each.
For field offices with more than 300 clients/enrollees in any given service: (see the table below)

<table>
<thead>
<tr>
<th>Service</th>
<th>More than 300</th>
<th>500</th>
<th>300</th>
<th>50-100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual reported number of clients/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>clients/enrollees</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disaster Services</strong></td>
<td>More than 300</td>
<td>500</td>
<td>300</td>
<td>50-100</td>
</tr>
<tr>
<td><strong>AFES</strong></td>
<td>More than 300</td>
<td>1,000</td>
<td>300</td>
<td>50-100</td>
</tr>
<tr>
<td><strong>Tracing Services</strong></td>
<td>More than 300</td>
<td>500</td>
<td>300</td>
<td>50-100</td>
</tr>
<tr>
<td><strong>First Aid, CPR and AED</strong></td>
<td>More than 300</td>
<td>1,000</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>Aquatics</strong></td>
<td>More than 300</td>
<td>1,000</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>HIV/AIDS</strong></td>
<td>More than 300</td>
<td>1,000</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>Community Services</strong></td>
<td>More than 300</td>
<td>1,000</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>Mission Related Caregiving</strong></td>
<td>More than 300</td>
<td>1,000</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>International Courses and Presentations</strong></td>
<td>More than 300</td>
<td>500</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

**SAMPLE SIZE FOR THE PILOT TEST**

**Clients in the sample**
- Send to not more than
- But not fewer than
- Minimum number of valid responses needed

Random Sample Procedure

For the agency field offices that serve more than 300 customers with a given service: You may decide to mail surveys to a sample of customers (rather than to all customers). If so, you will need to draw a random sample. This is the standard procedure for drawing a random sample:

5. Find out how many total customers you have.

6. Assign a different number to each case, starting 1, 2, etc. Generate random numbers with a random number program, (you can use Excel, Lotus, Access, or simply a random numbers table provided at the end of most basic statistics textbooks). In MS Excel you can use the function =INT(RAND()*TOTAL) being TOTAL the number of cases you have. To generate a table of random numbers, copy the function to a range of cells in the spreadsheet and use it to select your cases. Sort as many random numbers as you need. Be advised that each time you open the spreadsheet, the random numbers will be recalculated. So, if you want to keep the numbers for your records, then copy the table into a Word document. For example, if you wish to mail questionnaires to 400 customers out of a total of 2,000, the program should draw 400 random numbers (without replacement of those numbers already selected) and sort these numbers in order from lowest to highest. Copy =INT(RAND()*2000) to 400 cells and select those cases.
7. Count through the customers using the sorted list of random numbers to identify the ones that will be included in the sample.

8. Transfer customer names and addresses to the envelopes.

Procedure to Apply the Surveys

There is a substantial change in the way these surveys will be processed and analyzed.

- First, to make sure that there is standardization across field offices, all clients are being asked the same set of questions nationwide. This means all field offices must use the same set of surveys to allow for fair comparison of results across field offices.
- Second, because these forms will be scanned at national headquarters, all field offices must either copy the forms in this package or download them from the ICP’s Web page and print them. The format cannot be changed or the form will be unusable with the scanning machine. **Forms need to be copied on a one double-sided page,** so that all questionnaires fit one page only. **Do not reduce or enlarge copies.**
- Third, when sending the surveys to clients by mail, you have to include a self-addressed envelope so that the surveys can come directly to national headquarters, be scanned and the results for each field office analyzed. If you want to use self-addressed envelopes directed to your field office, gather the responses and mail them to national headquarters periodically.
- When the forms are collected on-site, after Health and Safety courses or International Services courses and presentations, instructors need to deliver the bundle of completed surveys to the field office. Field offices in turn need to mail the forms to national headquarters.

We are asking a number of demographic questions to better know our customers. The zip code and date of the service is requested for mapping purposes (meaning, that we actually have software to create maps). Maps are used to identify things like regions of the country where satisfaction is higher, or where the The agency was consistently late on a disaster scene, or where CPR training is more popular, etc. We are asking the date to track seasonally distributed services, if any.

SPECIAL INSTRUCTIONS IN HOW TO FILL THE BOXES IN THE “FOR FIELD OFFICE USE ONLY” SECTION:

We are using state-of-the-art scanning equipment that is capable of recognizing handwritten characters. To increase the dependability of important information, however, such as your field office code, a few instructions need to be followed.

Whenever the software is not sure of the value of any character in the boxes, it will prompt an operator to decide which character to input. However, we will save time and resources if the majority of the forms can be scanned automatically, without human intervention. To improve our chances, please note the following points:

- Numbers or letters should not touch the sides of the boxes.
- Numbers and letters should be clear and distinct from each other.
- Letters should be capitalized.
- Do not overwrite. If you make an error, please erase it or ask for a new form.
- Write your comments only for those questions where a rectangular box is provided
- Write within the box after filling the corresponding bubble on the side.
- Do not scribble on the form.
The table below shows characters easily verifiable and gives examples of characters and styles not easily verifiable by the software:

<table>
<thead>
<tr>
<th>Numbers</th>
<th>Easily verifiable</th>
<th>Not Verifiable</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 0 0</td>
<td>Ø Ø Ø</td>
</tr>
<tr>
<td>1</td>
<td>1 1 1</td>
<td>1 / 1</td>
</tr>
<tr>
<td>2</td>
<td>2 2 2</td>
<td>2 2 2</td>
</tr>
<tr>
<td>3</td>
<td>3 3 3</td>
<td>3 3 3</td>
</tr>
<tr>
<td>4</td>
<td>4 4 4</td>
<td>4 4 4</td>
</tr>
<tr>
<td>5</td>
<td>5 5 5</td>
<td>5 5 5</td>
</tr>
<tr>
<td>6</td>
<td>6 6 6</td>
<td>6 6 6</td>
</tr>
<tr>
<td>7</td>
<td>7 7 7</td>
<td>7 7 7</td>
</tr>
<tr>
<td>8</td>
<td>8 8 8</td>
<td>8 8 8</td>
</tr>
<tr>
<td>9</td>
<td>9 9 9</td>
<td>9 9 9</td>
</tr>
</tbody>
</table>

The next pages present a process map or flowchart of the process your field office needs to follow to survey your clients. They graphically describe each step from selecting the sample to mailing the survey to national headquarters.
Select questionnaire to process

Confirm chapter population

Obtain number of clients for that service

Is number of clients greater than 300?

NO

Send / give questionnaires to all clients (census)

YES

Sample size per service/ questionnaire >= 300 (max 10% of clients served)

Generate random numbers or use random number table to select sample

Review the survey forms and fill out the information required under the section “For...

Copy enough survey forms (two-sided)

Add mailing labels to each envelope

Write mailing labels with clients’ addresses

National headquarters processes surveys

National headquarters prepares summary and distributes report back to field offices

Mail surveys to clients

See attached Random Numbers Table

Read important instructions attached

Order business-size envelopes to mail the forms

Order return envelopes with national headquarters’ address to include in the package

Include one return envelope in the envelope with each survey (do not staple)
Select questionnaire to process

Obtain number of enrollees for the class/presentation

Read important instructions attached

Review the survey forms and fill out the information required under the section “For Field Use Only”

Copy enough survey forms (two-sided)

Give forms to instructor

Instructor distributes after class to all enrollees

Instructor returns completed surveys to field office

Field of. bundles the returned forms (surveys) and mails to national headquarters

National headquarters prepares summary and distributes report back to field offices.

Field officer repeats process for each class until end of fiscal year

PROCESS MAP FOR SURVEYS RELATED TO HEALTH AND SAFETY AND INTERNATIONAL SERVICES COURSES
Select questionnaire to process

Obtain number of clients served per year

Read important instructions attached

Review the survey forms and fill out the information required under the section “For field use only”

Copy enough survey forms (two-sided)

Give forms to the worker or person responsible for distributing to clients responsible

Worker distributes after service (for community services) or during follow up (for disaster victims) to all clients

If the survey is for Disaster

Order return envelopes with NHQ’s address to include in the package

Include one return envelope with each survey (do not staple)

Field officer bundles the returned forms (surveys) and mails to NHQ

For CS: Worker returns completed surveys to chapter
For DS: client returns to chapter/NHQ

National headquarters prepares summary and distributes report back to field offices

National headquarters processes surveys

Field office repeats process for each service/client
Copying and Adapting the Questionnaires

For the Indicators of Field office Performance, all field offices should use the same survey questionnaires. The surveys can be downloaded and printed by your local field office from the ICPs Web page at https://l

You can also copy the surveys from this booklet.

NOTE: IT IS ABSOLUTELY NECESSARY THAT YOU PRODUCE DOUBLE-SIDED HIGH-QUALITY COPIES OF THE QUESTIONNAIRES. EVERY QUESTIONNAIRE MUST FIT ON ONE TWO-SIDED LETTER SIZE PAGE. DO NOT REDUCE OR ENLARGE

ALSO, THE ORIENTING SQUARES AT THE CORNERS OF THE FORMS NEED TO BE PRESERVED WHILE COPYING. MAKE SURE THAT YOUR COPIES ARE NOT SKEWED AND THAT ALL FOUR BLOCKS SHOW UP IN YOUR COPIES.

Because national headquarters will scan the surveys and analyze the data for you using automated procedures, you cannot add questions to address local issues. Please contact us if you have special needs and we will work with you to establish a separate survey. All questionnaires need to be a printout from the web page or a high quality copy of the attached questionnaires. Failure to do that will compromise the “scanability” of the forms and national headquarters’s ability to analyze your data.

Mail Survey Checklist

1. Decide how many customers you will contact.
2. Review the survey forms and fill out the information required under the section “For Field office Use Only.” Be careful so that your handwriting is kept inside the boxes and does not touch their borders.
3. Copy enough survey forms. You need to copy every survey form in two-sided format so that each questionnaire fits on one page.
4. Order standard business-size envelopes (#10) to mail the questionnaires to customers.
5. Order smaller return envelopes so that customers can return the survey to you or to national headquarters. It is best to pay the postage. Asking the customer to put a stamp on the envelope will reduce the likelihood of a response.
6. You also may wish to order mailing labels if you can print mailing labels more easily than you can address envelopes.

NOTE: If a service was provided to a family, you will need to decide which person in the family should receive the questionnaire. One approach might be to survey the primary applicant (the person who requested the service) when this person can be identified. Otherwise, you might alternate between selecting the female or male head of household.

NOTE: For Health and Safety Services questionnaires, one option is to write in the name of the course, the date, the name of the instructor, and the location before the questionnaire is mailed to the instructor. Alternatively, the instructor may provide this information to the customers as they complete these questionnaires. In any case, please make sure that all information relative to your field office is filled into the appropriate spaces.
7. Staple a return envelope to each survey form so that the questionnaire and the return envelope don’t become separated when the customer opens the letter.

8. Log in surveys as received using the tracking form on page 17. You can also copy them if you wish to make a parallel analysis in your field office, however, this is NOT NECESSARY, as Administration and Evaluation has technical staff and equipment dedicated to do the analysis for you. We can also provide you with your raw data if requested. If you choose to copy forms, PLEASE SEND ORIGINALS to national headquarters.

9. Mail the surveys in bundles of 50-100 to national headquarters and make sure that each one of the surveys has the complete information about field office, course and instructor (where applicable)
Appendix C: Data Sets
Data set 1 from Pilot Disaster Services Departments

<table>
<thead>
<tr>
<th>ECODE</th>
<th>AGGHINC</th>
<th>PCT_Minority</th>
<th>POP00</th>
<th>HH_1999</th>
<th>DISASEXP</th>
<th>FNDRSN_E</th>
<th>FRAISREV</th>
<th>FRVOLS_REG</th>
<th>DAY_ASSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5132</td>
<td>$14,722,932</td>
<td>56%</td>
<td>742,355</td>
<td>102,837</td>
<td>$258,556.00</td>
<td>186,613</td>
<td>$513,425</td>
<td>26</td>
<td>800.30</td>
</tr>
<tr>
<td>5180</td>
<td>$121,962,619</td>
<td>66%</td>
<td>5,737,588</td>
<td>1,698,041</td>
<td>$5,531,606.00</td>
<td>2,039,211</td>
<td>$7,812,053</td>
<td>181</td>
<td>375.97</td>
</tr>
<tr>
<td>5264</td>
<td>$70,424,969</td>
<td>39%</td>
<td>2,810,321</td>
<td>1,541,012</td>
<td>$1,326,767.00</td>
<td>1,022,436</td>
<td>$4,463,398</td>
<td>68</td>
<td>503.91</td>
</tr>
<tr>
<td>5503</td>
<td>$129,404,072</td>
<td>46%</td>
<td>4,475,547</td>
<td>2,404,178</td>
<td>$4,604,701.00</td>
<td>1,379,087</td>
<td>$10,613,322</td>
<td>165</td>
<td>609.73</td>
</tr>
<tr>
<td>10060</td>
<td>$41,682,798</td>
<td>69%</td>
<td>2,278,320</td>
<td>1,135,074</td>
<td>$1,587,571.00</td>
<td>398,730</td>
<td>$3,910,004</td>
<td>211</td>
<td>297.33</td>
</tr>
<tr>
<td>12024</td>
<td>$10,247,825</td>
<td>10%</td>
<td>1,251,752</td>
<td>256,538</td>
<td>$733,253.00</td>
<td>348,803</td>
<td>$713,950</td>
<td>75</td>
<td>244.06</td>
</tr>
<tr>
<td>15080</td>
<td>$10,621,037</td>
<td>8%</td>
<td>463,430</td>
<td>258,027</td>
<td>$213,911.00</td>
<td>123,067</td>
<td>$755,368</td>
<td>20</td>
<td>1093.92</td>
</tr>
<tr>
<td>22172</td>
<td>$7,202,271</td>
<td>12%</td>
<td>342,220</td>
<td>177,780</td>
<td>$267,473.00</td>
<td>134,954</td>
<td>$507,013</td>
<td>42</td>
<td>1318.40</td>
</tr>
<tr>
<td>22322</td>
<td>$97,514,301</td>
<td>30%</td>
<td>4,081,169</td>
<td>2,055,642</td>
<td>$2,204,434.00</td>
<td>1,049,960</td>
<td>$2,494,295</td>
<td>25</td>
<td>825.14</td>
</tr>
<tr>
<td>25150</td>
<td>$34,804,630</td>
<td>17%</td>
<td>1,448,301</td>
<td>786,095</td>
<td>$1,754,722.00</td>
<td>406,969</td>
<td>$1,993,832</td>
<td>30</td>
<td>790.32</td>
</tr>
<tr>
<td>29042</td>
<td>$2,269,208</td>
<td>1%</td>
<td>133,082</td>
<td>70,152</td>
<td>$74,780.00</td>
<td>34,097</td>
<td>$183,513</td>
<td>12</td>
<td>250.96</td>
</tr>
<tr>
<td>35072</td>
<td>$34,996,065</td>
<td>15%</td>
<td>1,540,943</td>
<td>778,808</td>
<td>$2,517,810.00</td>
<td>806,883</td>
<td>$1,046,267</td>
<td>18</td>
<td>934.80</td>
</tr>
<tr>
<td>35168</td>
<td>$33,562,930</td>
<td>29%</td>
<td>1,454,014</td>
<td>710,111</td>
<td>$1,423,864.00</td>
<td>508,955</td>
<td>$2,285,015</td>
<td>260</td>
<td>471.45</td>
</tr>
<tr>
<td>38360</td>
<td>$80,765,331</td>
<td>29%</td>
<td>3,211,087</td>
<td>1,727,033</td>
<td>$2,776,825.00</td>
<td>1,168,798</td>
<td>$7,170,390</td>
<td>338</td>
<td>374.12</td>
</tr>
<tr>
<td>45010</td>
<td>$1,843,303</td>
<td>2%</td>
<td>99,491</td>
<td>52,226</td>
<td>$27,375.00</td>
<td>66,901</td>
<td>$183,513</td>
<td>12</td>
<td>250.96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ECODE</th>
<th>DSVOLS_REG</th>
<th>DSTOTSTAFF</th>
<th>DSHR Enroll</th>
<th>Dis_fam_assited</th>
<th>% Minor leadership</th>
<th>PROFESSION</th>
<th>DISEXP_HH</th>
</tr>
</thead>
<tbody>
<tr>
<td>5132</td>
<td>311</td>
<td>3</td>
<td>49</td>
<td>79</td>
<td>14.71</td>
<td>5</td>
<td>2.51</td>
</tr>
<tr>
<td>5180</td>
<td>1122</td>
<td>31</td>
<td>262</td>
<td>574</td>
<td>34.30</td>
<td>197</td>
<td>3.26</td>
</tr>
<tr>
<td>5264</td>
<td>891</td>
<td>6.04</td>
<td>115</td>
<td>248</td>
<td>17.57</td>
<td>6</td>
<td>0.86</td>
</tr>
<tr>
<td>5503</td>
<td>758</td>
<td>17</td>
<td>410</td>
<td>855</td>
<td>31.38</td>
<td>50</td>
<td>1.92</td>
</tr>
<tr>
<td>10060</td>
<td>1907</td>
<td>11</td>
<td>92</td>
<td>246</td>
<td>22.17</td>
<td>75</td>
<td>1.40</td>
</tr>
<tr>
<td>12024</td>
<td>623</td>
<td>8</td>
<td>50</td>
<td>162</td>
<td>5.59</td>
<td>1</td>
<td>2.86</td>
</tr>
<tr>
<td>15080</td>
<td>131</td>
<td>2.35</td>
<td>30</td>
<td>144</td>
<td>17.46</td>
<td>14</td>
<td>0.83</td>
</tr>
<tr>
<td>22172</td>
<td>506</td>
<td>4.5</td>
<td>77</td>
<td>101</td>
<td>8.24</td>
<td>5</td>
<td>1.50</td>
</tr>
<tr>
<td>22322</td>
<td>292</td>
<td>7</td>
<td>80</td>
<td>1676</td>
<td>30.89</td>
<td>64</td>
<td>1.07</td>
</tr>
<tr>
<td>25150</td>
<td>378</td>
<td>19</td>
<td>119</td>
<td>564</td>
<td>3.70</td>
<td>79</td>
<td>2.23</td>
</tr>
<tr>
<td>29042</td>
<td>127</td>
<td>0.5</td>
<td>11</td>
<td>62</td>
<td>4.55</td>
<td>7</td>
<td>1.07</td>
</tr>
<tr>
<td>35072</td>
<td>1070</td>
<td>16.01</td>
<td>78</td>
<td>815</td>
<td>35.74</td>
<td>83</td>
<td>3.23</td>
</tr>
<tr>
<td>35168</td>
<td>351</td>
<td>8</td>
<td>107</td>
<td>369</td>
<td>32.05</td>
<td>96</td>
<td>2.01</td>
</tr>
<tr>
<td>38360</td>
<td>636</td>
<td>16</td>
<td>87</td>
<td>1078</td>
<td>12.81</td>
<td>34</td>
<td>1.61</td>
</tr>
<tr>
<td>45010</td>
<td>50</td>
<td>2</td>
<td>49</td>
<td>17</td>
<td>.00</td>
<td>1</td>
<td>0.52</td>
</tr>
<tr>
<td>ECODE</td>
<td>DSVOLS_REG_1000HH</td>
<td>DSTOTSTAFF_10KHH</td>
<td>DSHR_10KHH</td>
<td>Dis Fam Ass_10KHH</td>
<td>Timeliness</td>
<td>SQ_Index</td>
<td>Recovery</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>------------</td>
<td>-------------------</td>
<td>------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>5132</td>
<td>3.02</td>
<td>0.29</td>
<td>5</td>
<td>8</td>
<td>0.00</td>
<td>25.25</td>
<td>25.00</td>
</tr>
<tr>
<td>5180</td>
<td>0.66</td>
<td>0.18</td>
<td>2</td>
<td>3</td>
<td>40</td>
<td>29.74</td>
<td>60</td>
</tr>
<tr>
<td>5264</td>
<td>0.58</td>
<td>0.04</td>
<td>1</td>
<td>2</td>
<td>55.00</td>
<td>31.85</td>
<td>66.70</td>
</tr>
<tr>
<td>5503</td>
<td>0.32</td>
<td>0.07</td>
<td>2</td>
<td>4</td>
<td>54</td>
<td>30.48</td>
<td>64.3</td>
</tr>
<tr>
<td>10060</td>
<td>1.68</td>
<td>0.10</td>
<td>1</td>
<td>2</td>
<td>59.00</td>
<td>31.68</td>
<td>77.40</td>
</tr>
<tr>
<td>12024</td>
<td>2.43</td>
<td>0.31</td>
<td>2</td>
<td>6</td>
<td>35</td>
<td>20.8</td>
<td>40</td>
</tr>
<tr>
<td>15080</td>
<td>0.51</td>
<td>0.09</td>
<td>1</td>
<td>6</td>
<td>70</td>
<td>32.77</td>
<td>77.3</td>
</tr>
<tr>
<td>22172</td>
<td>2.85</td>
<td>0.25</td>
<td>4</td>
<td>6</td>
<td>75.00</td>
<td>33.00</td>
<td>50.00</td>
</tr>
<tr>
<td>22322</td>
<td>0.14</td>
<td>0.03</td>
<td>0</td>
<td>8</td>
<td>58.00</td>
<td>31.09</td>
<td>60.90</td>
</tr>
<tr>
<td>25150</td>
<td>0.48</td>
<td>0.24</td>
<td>2</td>
<td>7</td>
<td>60.00</td>
<td>31.88</td>
<td>73.10</td>
</tr>
<tr>
<td>29042</td>
<td>1.81</td>
<td>0.07</td>
<td>2</td>
<td>9</td>
<td>40.00</td>
<td>24.88</td>
<td>37.50</td>
</tr>
<tr>
<td>35072</td>
<td>1.37</td>
<td>0.21</td>
<td>1</td>
<td>10</td>
<td>60</td>
<td>31.24</td>
<td>68.6</td>
</tr>
<tr>
<td>35168</td>
<td>0.49</td>
<td>0.11</td>
<td>2</td>
<td>5</td>
<td>65.00</td>
<td>31.78</td>
<td>80.40</td>
</tr>
<tr>
<td>38360</td>
<td>0.37</td>
<td>0.09</td>
<td>1</td>
<td>6</td>
<td>40.00</td>
<td>30.16</td>
<td>58.20</td>
</tr>
<tr>
<td>45010</td>
<td>0.96</td>
<td>0.38</td>
<td>9</td>
<td>3</td>
<td>35</td>
<td>25.67</td>
<td>66.7</td>
</tr>
</tbody>
</table>
## Dataset 2: 32 Field Offices

<table>
<thead>
<tr>
<th>ECODE</th>
<th>POP</th>
<th>AGHII</th>
<th>MINORITY</th>
<th>POP_DEN_PHH</th>
<th>DONATIONS_PHH</th>
<th>MGT_FNDRS_PHH</th>
<th>INST_INC_PHH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1019</td>
<td>1,013,636</td>
<td>22,692,714</td>
<td>28.36</td>
<td>220.17</td>
<td>4.54</td>
<td>1.48</td>
<td>6.92</td>
</tr>
<tr>
<td>3060</td>
<td>1,054,770</td>
<td>18,562,591</td>
<td>39.81</td>
<td>41.61</td>
<td>3.23</td>
<td>1.11</td>
<td>0.38</td>
</tr>
<tr>
<td>5112</td>
<td>888,134</td>
<td>16,113,954</td>
<td>56.87</td>
<td>109.49</td>
<td>1.14</td>
<td>1.06</td>
<td>0.01</td>
</tr>
<tr>
<td>7038</td>
<td>983,848</td>
<td>25,149,846</td>
<td>19.47</td>
<td>568.60</td>
<td>1.76</td>
<td>2.41</td>
<td>3.99</td>
</tr>
<tr>
<td>8004</td>
<td>762,227</td>
<td>18,568,037</td>
<td>25.06</td>
<td>219.16</td>
<td>5.22</td>
<td>1.99</td>
<td>4.15</td>
</tr>
<tr>
<td>10198</td>
<td>1,071,025</td>
<td>21,609,041</td>
<td>30.21</td>
<td>233.51</td>
<td>1.83</td>
<td>0.68</td>
<td>2.14</td>
</tr>
<tr>
<td>10228</td>
<td>1,097,227</td>
<td>30,082,581</td>
<td>27.44</td>
<td>349.68</td>
<td>6.40</td>
<td>2.40</td>
<td>0.63</td>
</tr>
<tr>
<td>17340</td>
<td>1,019,628</td>
<td>21,615,931</td>
<td>15.59</td>
<td>354.14</td>
<td>2.57</td>
<td>2.39</td>
<td>10.71</td>
</tr>
<tr>
<td>20060</td>
<td>863,876</td>
<td>33,153,404</td>
<td>36.69</td>
<td>1750.86</td>
<td>3.66</td>
<td>1.66</td>
<td>0.94</td>
</tr>
<tr>
<td>20064</td>
<td>786,865</td>
<td>18,998,771</td>
<td>66.84</td>
<td>1624.08</td>
<td>2.04</td>
<td>0.67</td>
<td>1.08</td>
</tr>
<tr>
<td>21244</td>
<td>803,045</td>
<td>18,408,453</td>
<td>8.65</td>
<td>819.77</td>
<td>0.74</td>
<td>0.52</td>
<td>0.75</td>
</tr>
<tr>
<td>23280</td>
<td>991,264</td>
<td>27,205,643</td>
<td>13.15</td>
<td>1104.96</td>
<td>4.24</td>
<td>3.18</td>
<td>3.52</td>
</tr>
<tr>
<td>30067</td>
<td>1,197,781</td>
<td>27,383,475</td>
<td>57.94</td>
<td>3918.16</td>
<td>3.96</td>
<td>0.91</td>
<td>1.09</td>
</tr>
<tr>
<td>30116</td>
<td>1,124,763</td>
<td>33,351,813</td>
<td>15.52</td>
<td>1014.67</td>
<td>2.47</td>
<td>1.13</td>
<td>0.14</td>
</tr>
<tr>
<td>30140</td>
<td>860,209</td>
<td>32,423,947</td>
<td>24.82</td>
<td>3649.59</td>
<td>3.81</td>
<td>0.81</td>
<td>0.92</td>
</tr>
<tr>
<td>30176</td>
<td>1,078,658</td>
<td>31,100,108</td>
<td>32.59</td>
<td>1808.62</td>
<td>3.05</td>
<td>1.81</td>
<td>1.40</td>
</tr>
<tr>
<td>32004</td>
<td>789,219</td>
<td>17,195,866</td>
<td>10.27</td>
<td>188.65</td>
<td>3.46</td>
<td>1.63</td>
<td>0.83</td>
</tr>
<tr>
<td>32140</td>
<td>1,037,746</td>
<td>21,035,992</td>
<td>16.42</td>
<td>422.57</td>
<td>1.79</td>
<td>2.59</td>
<td>3.24</td>
</tr>
<tr>
<td></td>
<td>32442</td>
<td>33181</td>
<td>35086</td>
<td>35106</td>
<td>36308</td>
<td>38196</td>
<td>39004</td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>909,719</td>
<td>787,169</td>
<td>1,104,473</td>
<td>757,019</td>
<td>772,395</td>
<td>918,696</td>
<td>1,032,606</td>
</tr>
<tr>
<td></td>
<td>37,180,338</td>
<td>19,939,570</td>
<td>22,834,007</td>
<td>14,863,459</td>
<td>15,661,175</td>
<td>1,104,473</td>
<td>22,672,842</td>
</tr>
<tr>
<td></td>
<td>2037.44</td>
<td>716.45</td>
<td>778.51</td>
<td>583.62</td>
<td>257.34</td>
<td>318.48</td>
<td>919.92</td>
</tr>
<tr>
<td></td>
<td>4.62</td>
<td>4.76</td>
<td>4.29</td>
<td>13.05</td>
<td>4.99</td>
<td>4.39</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>1.29</td>
<td>3.91</td>
<td>2.28</td>
<td>3.18</td>
<td>3.38</td>
<td>2.15</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>3.72</td>
<td>4.76</td>
<td>4.83</td>
<td>7.82</td>
<td>5.86</td>
<td>3.69</td>
<td>1.28</td>
</tr>
<tr>
<td>ECODE</td>
<td>ALLOTH</td>
<td>INC_PH</td>
<td>H&amp;S_REACH</td>
<td>DS_DSHR_P10KHH</td>
<td>VL_REG_P1KHH</td>
<td>PD_TOTAL_P100KHH</td>
<td>HS_INST_P1KHH</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>-----------</td>
<td>----------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>---------------</td>
</tr>
<tr>
<td>1019</td>
<td>2.00</td>
<td>5.27</td>
<td>1.65</td>
<td>3.14</td>
<td>18.04</td>
<td>2.26</td>
<td>338</td>
</tr>
<tr>
<td>3060</td>
<td>1.50</td>
<td>1.70</td>
<td>1.49</td>
<td>1.11</td>
<td>4.89</td>
<td>2.21</td>
<td>468</td>
</tr>
<tr>
<td>5112</td>
<td>1.40</td>
<td>1.65</td>
<td>0.89</td>
<td>0.54</td>
<td>9.24</td>
<td>1.26</td>
<td>62</td>
</tr>
<tr>
<td>7038</td>
<td>3.45</td>
<td>6.57</td>
<td>2.15</td>
<td>4.73</td>
<td>12.35</td>
<td>9.09</td>
<td>506</td>
</tr>
<tr>
<td>8004</td>
<td>1.74</td>
<td>1.99</td>
<td>0.82</td>
<td>4.28</td>
<td>17.39</td>
<td>2.71</td>
<td>376</td>
</tr>
<tr>
<td>10198</td>
<td>0.93</td>
<td>1.94</td>
<td>1.76</td>
<td>3.59</td>
<td>8.31</td>
<td>6.16</td>
<td>609</td>
</tr>
<tr>
<td>10228</td>
<td>1.93</td>
<td>4.03</td>
<td>1.18</td>
<td>8.21</td>
<td>11.11</td>
<td>1.59</td>
<td>980</td>
</tr>
<tr>
<td>17340</td>
<td>4.81</td>
<td>4.62</td>
<td>2.52</td>
<td>9.26</td>
<td>48.78</td>
<td>4.24</td>
<td>502</td>
</tr>
<tr>
<td>20060</td>
<td>0.89</td>
<td>1.33</td>
<td>0.70</td>
<td>2.53</td>
<td>10.08</td>
<td>1.33</td>
<td>433</td>
</tr>
<tr>
<td>20064</td>
<td>1.30</td>
<td>2.74</td>
<td>0.56</td>
<td>7.19</td>
<td>7.03</td>
<td>3.02</td>
<td>161</td>
</tr>
<tr>
<td>21244</td>
<td>1.85</td>
<td>3.62</td>
<td>0.65</td>
<td>1.57</td>
<td>10.28</td>
<td>2.57</td>
<td>35</td>
</tr>
<tr>
<td>23280</td>
<td>4.25</td>
<td>7.16</td>
<td>1.84</td>
<td>3.63</td>
<td>41.09</td>
<td>15.29</td>
<td>414</td>
</tr>
<tr>
<td>30067</td>
<td>2.55</td>
<td>1.75</td>
<td>0.64</td>
<td>0.98</td>
<td>8.05</td>
<td>1.13</td>
<td>302</td>
</tr>
<tr>
<td>30116</td>
<td>2.31</td>
<td>2.65</td>
<td>0.71</td>
<td>4.01</td>
<td>4.28</td>
<td>2.95</td>
<td>242</td>
</tr>
<tr>
<td>30140</td>
<td>1.54</td>
<td>2.54</td>
<td>0.58</td>
<td>5.91</td>
<td>21.65</td>
<td>4.13</td>
<td>473</td>
</tr>
<tr>
<td>30176</td>
<td>1.99</td>
<td>2.82</td>
<td>1.38</td>
<td>3.03</td>
<td>9.43</td>
<td>4.88</td>
<td>350</td>
</tr>
<tr>
<td>32004</td>
<td>2.72</td>
<td>5.46</td>
<td>2.23</td>
<td>4.59</td>
<td>11.48</td>
<td>6.82</td>
<td>330</td>
</tr>
<tr>
<td>32140</td>
<td>2.19</td>
<td>3.59</td>
<td>0.91</td>
<td>4.99</td>
<td>20.37</td>
<td>3.92</td>
<td>318</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>32442</td>
<td>1.83</td>
<td>5.36</td>
<td>1.16</td>
<td>11.44</td>
<td>10.70</td>
<td>7.69</td>
<td>307</td>
</tr>
<tr>
<td>33181</td>
<td>3.11</td>
<td>11.24</td>
<td>2.21</td>
<td>5.92</td>
<td>21.79</td>
<td>5.26</td>
<td>314</td>
</tr>
<tr>
<td>35086</td>
<td>2.66</td>
<td>6.42</td>
<td>1.73</td>
<td>4.45</td>
<td>13.43</td>
<td>4.17</td>
<td>563</td>
</tr>
<tr>
<td>35106</td>
<td>3.37</td>
<td>7.40</td>
<td>1.88</td>
<td>6.86</td>
<td>30.45</td>
<td>7.12</td>
<td>353</td>
</tr>
<tr>
<td>36308</td>
<td>1.96</td>
<td>9.76</td>
<td>3.26</td>
<td>4.25</td>
<td>31.27</td>
<td>2.83</td>
<td>404</td>
</tr>
<tr>
<td>38196</td>
<td>4.25</td>
<td>5.38</td>
<td>2.43</td>
<td>9.92</td>
<td>24.64</td>
<td>8.18</td>
<td>493</td>
</tr>
<tr>
<td>39004</td>
<td>1.32</td>
<td>1.22</td>
<td>2.14</td>
<td>4.39</td>
<td>4.59</td>
<td>2.56</td>
<td>251</td>
</tr>
<tr>
<td>40050</td>
<td>2.39</td>
<td>5.43</td>
<td>1.79</td>
<td>5.14</td>
<td>15.45</td>
<td>3.36</td>
<td>92</td>
</tr>
<tr>
<td>42264</td>
<td>0.98</td>
<td>3.34</td>
<td>0.76</td>
<td>4.43</td>
<td>11.33</td>
<td>2.92</td>
<td>468</td>
</tr>
<tr>
<td>42296</td>
<td>1.43</td>
<td>1.76</td>
<td>1.79</td>
<td>7.22</td>
<td>14.90</td>
<td>3.11</td>
<td>312</td>
</tr>
<tr>
<td>43092</td>
<td>0.76</td>
<td>2.85</td>
<td>3.84</td>
<td>5.40</td>
<td>4.10</td>
<td>2.76</td>
<td>579</td>
</tr>
<tr>
<td>43129</td>
<td>1.32</td>
<td>2.63</td>
<td>2.41</td>
<td>2.35</td>
<td>10.87</td>
<td>3.33</td>
<td>151</td>
</tr>
<tr>
<td>44084</td>
<td>2.20</td>
<td>4.88</td>
<td>1.15</td>
<td>1.14</td>
<td>8.14</td>
<td>6.19</td>
<td>123</td>
</tr>
<tr>
<td>46372</td>
<td>1.83</td>
<td>3.27</td>
<td>2.46</td>
<td>8.32</td>
<td>30.31</td>
<td>5.01</td>
<td>480</td>
</tr>
<tr>
<td>46441</td>
<td>1.78</td>
<td>3.53</td>
<td>1.69</td>
<td>12.47</td>
<td>19.60</td>
<td>5.63</td>
<td>509</td>
</tr>
<tr>
<td>54004</td>
<td>3.57</td>
<td>3.26</td>
<td>4.46</td>
<td>10.39</td>
<td>21.34</td>
<td>7.34</td>
<td>125</td>
</tr>
</tbody>
</table>
## APPENDIX D: Excel Solver Codes

![Excel Solver Interface](image.png)

The Excel Solver is a tool used for solving optimization problems. In this document, the solver is used to find the maximum value of a certain objective function subject to specific constraints.

### Excel Solver Parameters

- **Max Change**: 0
- **Change**: 1
- **Max Iterations**: 1
- **Max Time**: 1

### Objective Function

The objective function is to maximize

\[
\text{Objective} = 0.4565 \times 0.4565 + 0.4565 \times 0.4565 + 0.4565 \times 0.4565
\]

### Constraints

1. Donations \( z_1 \) subject to \( z_1 \leq 1 \)
2. PCI_Min_Thru \( z_2 \) subject to \( z_2 \leq 1 \)
3. PCI_Min_Event \( z_3 \) subject to \( z_3 \leq 1 \)
4. PCI_Min_Min \( z_4 \) subject to \( z_4 \leq 1 \)
5. \( z_5 \) subject to \( z_5 \leq 1 \)

### Input Data

- Donations: 100, 200, 300
- PCI_Min_Thru: 0.4565
- PCI_Min_Event: 0.4565
- PCI_Min_Min: 0.4565
- PCI_Min_Mix: 0.4565

### Solver Results

- Optimal Solution: Value of the objective function is maximized to 0.4565.

---

*Note: The image and code are placeholders and do not reflect the actual Excel Solver interface or data.*
Sub Macro1()
    Application.ScreenUpdating = False
    Sheets("Sheet2").Select
    Range("A1").Select
    For n = 1 To 56
        Sheets("Model").Select
        Range("B3").Select
        Selection.Value = n
        Range("C12:C67").Select
        Selection.Value = 0
        Dim RowofCellToAdd
        RowofCellToAdd = Str(n + 11)
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        Dim CellToAdd
        CellToAdd = "C" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        Range("C7").Select
        Selection.Value = 1
        Range("C8:" & CellToAdd).Select
        Selection.Value = 0
        SolverAdd (CellToAdd)
        SolverSolve (True)
        Range("C2:C67").Select
        Selection.Copy
        Sheets("Sheet2").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
                        False, Transpose:=False
        Range("A1").Offset(0, n).Select
    Next n
    For n = 1 To 56
        RowofCellToAdd = Str(n + 11)
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        CellToAdd = "C" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        SolverDelete (CellToAdd)
    Next n
End Sub
Sub Macro1()
Application.ScreenUpdating = False
Sheets("Sheet2").Select
  Range("A1").Select
  For n = 1 To 15
    Sheets("Data & Model").Select
    Range("J3").Select
    Selection.Value = n
    Range("K16:K31").Select
    Selection.Value = 0
    Dim RowofCellToAdd
    RowofCellToAdd = Str(n + 16)
    LengofRowofCellToAdd = Len(RowofCellToAdd)
    Dim CellToAdd
    CellToAdd = "K" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
    Range("K7:K10").Select
    Selection.Value = 1
    Range("K11:" & CellToAdd).Select
    Selection.Value = 0
    SolverAdd (CellToAdd)
    SolverSolve (True)
    Range("K3:K31").Select
    Selection.Copy
    Sheets("Sheet2").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
      False, Transpose:=False
    Range("A1").Offset(0, n).Select
  Next n
  For n = 1 To 15
    RowofCellToAdd = Str(n + 16)
    LengofRowofCellToAdd = Len(RowofCellToAdd)
    CellToAdd = "K" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
    SolverDelete (CellToAdd)
  Next n
End Sub
<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
<th>Q</th>
<th>R</th>
<th>S</th>
<th>T</th>
<th>U</th>
<th>V</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>PROG CODE</td>
<td>ECODE</td>
<td>Pop</td>
<td>DGER ENROL</td>
<td>disc req vol</td>
<td>STAFF</td>
<td>ASSETS</td>
<td>Timeline</td>
<td>SQ_index</td>
<td>Dist_Fam_assist</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>45210</td>
<td>95491</td>
<td>49</td>
<td>50</td>
<td>2</td>
<td>250.86</td>
<td>35</td>
<td>25.67</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>Obj</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>DGER ENROL</td>
<td>4.56-14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>disc req vol</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Dgeries TOT PAID STAFF</td>
<td>1.3E-14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Dgies TOT PAID STAFF</td>
<td>0</td>
<td>disc req vol</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Timeliness</td>
<td>0</td>
<td>Dgeries TOT PAID STAFF</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>SQ_index</td>
<td>0</td>
<td>Dgeries TOT PAID STAFF</td>
<td>250.86</td>
<td>250.86</td>
<td>250.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Dist_Fam_assist</td>
<td>0</td>
<td>Timeliness</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
</tr>
<tr>
<td>18</td>
<td>21</td>
<td>4.56-14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
</tr>
</tbody>
</table>

**Solver Parameters**

**Objective**
- **Set Cell**: R15
- **Equal To**: Min
- **Subject To the Constraints**:
  - R12: $P12 = $R12: $R15
  - R13: $P13 = $R13: $R15

**Changing Variable Cells**
- $P12
- $P13
- $P14
- $P15

347
Sub Macro1()
    Application.ScreenUpdating = False
    Sheets("Sheet2").Select
    Range("A1").Select
    For n = 1 To 15
        Sheets("Data & Model").Select
        Range("M3").Select
        Selection.Value = n
        Range("N17:N31").Select
        Selection.Value = 0
        Dim RowofCellToAdd
        RowofCellToAdd = Str(n + 16)
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        Dim CellToAdd
        CellToAdd = "N" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        Range("N7:N9").Select
        Selection.Value = 1
        Range("N10:" & CellToAdd).Select
        Selection.Value = 0
        SolverAdd (CellToAdd)
        SolverSolve (True)
        Range("N3:N31").Select
        Selection.Copy
        Sheets("Sheet2").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
            False, Transpose:=False
        Range("A1").Offset(0, n).Select
        Next n
    For n = 1 To 15
        RowofCellToAdd = Str(n + 16)
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        CellToAdd = "N" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        SolverDelete (CellToAdd)
        Next n
    End Sub
<table>
<thead>
<tr>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
<th>Q</th>
<th>R</th>
<th>S</th>
<th>T</th>
<th>U</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PROG CODE</td>
<td>ECODE</td>
<td>POL_Majority</td>
<td>Timeliness</td>
<td>SQ_Index</td>
<td>Dis_Fam</td>
<td>Recovery</td>
<td>Needs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>BMU</td>
<td>15</td>
<td>23042</td>
<td>0.012905625</td>
<td>42</td>
<td>4.91</td>
<td>62</td>
<td>37.5</td>
<td>37.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>max Obj</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>iPhi</td>
<td></td>
<td>2.1876</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Timeliness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>SQ_Index</td>
<td></td>
<td>0.2202</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Dis_Fam_basic</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Recovery</td>
<td></td>
<td>0</td>
<td>Timeliness</td>
<td>42</td>
<td>42 &lt;=</td>
<td>42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Needs</td>
<td></td>
<td>0</td>
<td>SQ_Index</td>
<td>4.6169</td>
<td>4.91 &lt;=</td>
<td>4.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>s1</td>
<td></td>
<td>0.00 Dis_Fam_basic</td>
<td>62</td>
<td>62 =</td>
<td>62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>s2</td>
<td></td>
<td>0</td>
<td>Recovery</td>
<td>80.10</td>
<td>80.10 &gt;=</td>
<td>80.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>s3</td>
<td></td>
<td>0</td>
<td>Needs</td>
<td>0.452</td>
<td>0.452 &gt;=</td>
<td>0.452</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>s4</td>
<td></td>
<td>0</td>
<td>VP</td>
<td>1</td>
<td>1 &lt;=</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>s5</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>s7</td>
<td></td>
<td>0.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>s9</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>s10</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>s11</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>s12</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>s13</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>s14</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>s15</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sub Macro1()
    Application.ScreenUpdating = False
    Sheets("Sheet2").Select
    Range("A1").Select
    For n = 1 To 15
        Sheets("Data & Model").Select
        Range("K3").Select
        Selection.Value = n
        Range("L14:L28").Select
        Selection.Value = 0
        Dim RowofCellToAdd
        RowofCellToAdd = Str(n + 13)
        Dim LengofRowofCellToAdd
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        Dim CellToAdd
        CellToAdd = "L" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        Range("L7:L8").Select
        Selection.Value = 1
        Range("L9:" & CellToAdd).Select
        Selection.Value = 0
        SolverAdd (CellToAdd)
        SolverSolve (True)
        Range("L3:L28").Select
        Selection.Copy
        Sheets("Sheet2").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
        Range("A1").Offset(0, n).Select
        Next n
    Sheets("Data & Model").Select
    For n = 1 To 15
        RowofCellToAdd = Str(n + 13)
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        CellToAdd = "N" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        SolverDelete (CellToAdd)
        Next n
    End Sub
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PROG_ID</td>
<td>CODE</td>
<td>ASHM</td>
<td>MINORITY</td>
<td>MGT_FNORS_PHH</td>
<td>DONATIONS_PHH</td>
<td>INST_INV_PHH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>31</td>
<td>33442</td>
<td>34567</td>
<td>29.69</td>
<td>1.146253812</td>
<td>4.024131816</td>
<td>3.736495413</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>GMU</td>
<td>31</td>
<td>33442</td>
<td>34567</td>
<td>29.69</td>
<td>1.146253812</td>
<td>4.024131816</td>
<td>3.736495413</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>MGT_FNORS</td>
<td>0</td>
<td>MINORITY</td>
<td>28.62</td>
<td>29.69</td>
<td>29.69</td>
<td>1.146253812</td>
<td>1.146253812</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>DONATIONS</td>
<td>0</td>
<td>MGT_FNORS</td>
<td>1.434131816</td>
<td>4.024131816</td>
<td>4.024131816</td>
<td>4.024131816</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>INST_INV</td>
<td>0</td>
<td>MGT_FNORS</td>
<td>0</td>
<td>1.00</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>MINORITY</td>
<td>0</td>
<td>MGT_FNORS</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subject Parameters:

Opt-Cell: $z_1$

Constraints:

- MGT_FNORS $\leq$ 1.146253812
- DONATIONS $\leq$ 4.024131816
- INST_INV $\geq$ 1.00

Solve:

- $z_1 = 0.6322$
- $z_2 = 0.01$
- $z_3 = 0.01$
- $z_4 = 0.01$
- $z_5 = 0.01$
- $z_6 = 0.01$
- $z_7 = 0.01$
- $z_8 = 0.01$
- $z_9 = 0.01$
- $z_{10} = 0.01$
- $z_{11} = 0.01$
- $z_{12} = 0.01$
- $z_{13} = 0.01$
- $z_{14} = 0.01$
- $z_{15} = 0.01$
- $z_{16} = 0.01$
- $z_{17} = 0.01$
- $z_{18} = 0.01$
- $z_{19} = 0.01$
- $z_{20} = 0.01$
Sub Macro2()
' Keyboard Shortcut: Ctrl+r
'To generate the Peer set and the Peer Weights
'  Dim Chapter, Peer_is
  Application.ScreenUpdating = False
  Set NewSheet = Worksheets.Add
  NewSheet.Name = "Peers"
  Sheets("Peers").Select
  Range("A1").Select
  ActiveCell.Value = "Chapter"
  ActiveCell.Offset(0, 1).Value = "Peer"
  ActiveCell.Offset(0, 2).Value = "Peer Weight"
  Sheets("Model").Select
  NoofChapters = Cells(1, 1)
  For n = 1 To NoofChapters
    Sheets("Temp").Select
    Chapter = Cells(n + 1, 1)
    For m = 1 To NoofChapters
      Sheets("Temp").Select
      If (n = m Or Cells(n + 1, m + 1) > 0.001) Then
        Peer_is = Cells(1, m + 1)
        PeerWeight = Cells(n + 1, m + 1)
        Sheets("Peers").Select
        ActiveCell.Value = Chapter
        ActiveCell.Offset(0, 1).Value = Peer_is
        ActiveCell.Offset(0, 2).Value = PeerWeight
        ActiveCell.Offset(1, 0).Select
        Sheets("Temp").Select
      End If
    Next m
  Next n
End Sub
<table>
<thead>
<tr>
<th></th>
<th>PROD_ID</th>
<th>CODE</th>
<th>CITY</th>
<th>STATE</th>
<th>POP</th>
<th>MINORITY</th>
<th>ALLINC_PH</th>
<th>DS_DISH_P</th>
<th>VL_REG_P</th>
<th>PD_TOTAL_P</th>
<th>HS_INST_P</th>
<th>DAY_ASET</th>
<th>NGAP</th>
<th>VHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>34</td>
<td>30067</td>
<td>Fairfield</td>
<td>NJ</td>
<td>1197761</td>
<td>57.94</td>
<td>7.8012548</td>
<td>0.036000075</td>
<td>0.975201003</td>
<td>0.046850802</td>
<td>1.131214</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>MINORITY</td>
<td>6</td>
<td>ALINC_P</td>
<td>DS_DISH_P</td>
<td>VL_REG_P</td>
<td>PD_TOTAL_P</td>
<td>HS_INST_P</td>
<td>DAY_ASET</td>
<td>NGAP</td>
<td>VHS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>x</td>
<td>24</td>
<td>z1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>24</td>
<td>z2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>24</td>
<td>z3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>24</td>
<td>z4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>24</td>
<td>z5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td>24</td>
<td>z6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>24</td>
<td>z7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>19</td>
<td>24</td>
<td>z8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>24</td>
<td>z9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>21</td>
<td>24</td>
<td>z10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>22</td>
<td>24</td>
<td>z11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>23</td>
<td>24</td>
<td>z12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Solver Parameters**

- **Set Cell**: Opt
- **Equal To**: 0
- **By Changing Variable Cells**: Opt
- **Subject to the Constraints**: Opt

**Solver Options**

- **Add Variable**: Opt
- **Add Constraint**: Opt
- **Add Variable**: Opt
- **Add Constraint**: Opt

**Solver Results**

- Optimal Value: 0
- Variables: Opt
- Constraints: Opt

**Solver Statistics**

- Objective: Opt
- Variables: Opt
- Constraints: Opt

**Solver Summary**

- Optimal Value: 0
- Variables: Opt
- Constraints: Opt
Sub Macro1()
    Application.ScreenUpdating = False
    Sheets("Sheet2").Select
    Range("A2").Select
    Sheets("Model").Select
    NoofChapters = Cells(1, 1)
    For n = 1 To NoofChapters
        Sheets("Model").Select
        Range("B3").Select
        Selection.Value = n
        Range("C21:C54").Select
        Selection.Value = 0
        Dim RowofCellToAdd
        RowofCellToAdd = Str(n + 20)
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        Dim CellToAdd
        CellToAdd = "C" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        Range("C7:C12").Select
        Selection.Value = 1
        Range("C13:" & CellToAdd).Select
        Selection.Value = 0
        SolverAdd (CellToAdd)
        SolverSolve (True)
        Range("C3:C54").Select
        Selection.Copy
        Sheets("Sheet2").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=True
        Range("A2").Offset(n, 0).Select
    Next n
    For n = 1 To NoofChapters
        Sheets("Model").Select
        RowofCellToAdd = Str(n + 20)
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        CellToAdd = "C" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        SolverDelete (CellToAdd)
    Next n
End Sub
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>34</td>
<td>PROG_ID</td>
<td>ECODE</td>
<td>CITY</td>
<td>STATE</td>
<td>POP</td>
<td>POP_DEN</td>
<td>POP_Q_HH</td>
<td>POI TOTAL PK HH</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>001</td>
<td>34</td>
<td>3512</td>
<td>0008</td>
<td>OH</td>
<td>75709</td>
<td>5936242837</td>
<td>1.15041103</td>
<td>1.13514992</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>max</td>
<td>OH</td>
<td>4</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>Ph</td>
<td>1</td>
<td>2</td>
<td>Phi</td>
<td>1</td>
<td>3</td>
<td>Phi</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>POP DEN</td>
<td>0</td>
<td>12</td>
<td>POP Q HH</td>
<td>0</td>
<td>13</td>
<td>POP Q HH</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>8</td>
<td>19</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>22</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>9</td>
<td>24</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>26</td>
<td>0</td>
<td>27</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>10</td>
<td>29</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>32</td>
<td>0</td>
<td>33</td>
</tr>
</tbody>
</table>

**Subset Parameters**

- **Equal R**: Faux Faux
- **Variable**
- **Add**: Add, Remove, Modify
- **Constraint**: Standard Linear, Quadratic
- **Options**: Standard Linear, Quadratic

**Model**

- **Start**: Start, Stop, Continue
- **Per Capita**: Per Capita, Add, Remove, Modify

**Results**

- **Solve**: Solve, Close, Options

---

355
Sub Macro1()
    Application.ScreenUpdating = False
    Set NewSheet = Worksheets.Add
    NewSheet.Name = "Temp"
    Sheets("Temp").Select
    Range("A2").Select
    Sheets("Model").Select
    NoofChapters = Cells(1, 1)
    For n = 1 To NoofChapters
        Sheets("Model").Select
        Range("B3").Select
        Selection.Value = n
        Range("C20:C53").Select
        Selection.Value = 0
        Dim RowofCellToAdd
        RowofCellToAdd = Str(n + 19)
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        Dim CellToAdd
        CellToAdd = "C" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        Range("C7:C10").Select
        Selection.Value = 1
        Range("C11:" & CellToAdd).Select
        Selection.Value = 0
        SolverAdd (CellToAdd)
        SolverSolve (True)
        Range("C3:C53").Select
        Selection.Copy
        Sheets("Temp").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
            False, Transpose:=True
        Range("A2").Offset(n, 0).Select
    Next n
    For n = 1 To NoofChapters
        Sheets("Model").Select
        RowofCellToAdd = Str(n + 19)
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        CellToAdd = "C" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        SolverDelete (CellToAdd)
    Next n
End Sub
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PROG_ID</td>
<td>EGODDE</td>
<td>AQHHI</td>
<td>MINORITY</td>
<td>H55_F</td>
<td>DS_FAM8_P</td>
<td>OTH_FRGEXP_PH</td>
<td>SQ_INDEX</td>
<td>Outcome</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>34</td>
<td>32422</td>
<td>31960128</td>
<td>3214</td>
<td>6.3564</td>
<td>10.1949916</td>
<td>2.41022659</td>
<td>2</td>
<td>52</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>Obj</td>
<td>0.00006666</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>Phi</td>
<td>1.0986</td>
<td>1.43187</td>
<td>1.0986</td>
<td>1.43187</td>
<td>1.0986</td>
<td>1.43187</td>
<td>1.0986</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>Phi</td>
<td>1.0986</td>
<td>1.43187</td>
<td>1.0986</td>
<td>1.43187</td>
<td>1.0986</td>
<td>1.43187</td>
<td>1.0986</td>
</tr>
<tr>
<td>7</td>
<td>11</td>
<td>MINORITY</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>H55_REACH</td>
<td>2.56885</td>
<td>2.56885</td>
<td>2.56885</td>
<td>2.56885</td>
<td>2.56885</td>
<td>2.56885</td>
<td>2.56885</td>
</tr>
<tr>
<td>9</td>
<td>13</td>
<td>DS_FAM8_P</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>OTH_FRGEXP_PH</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>11</td>
<td>15</td>
<td>SQ_INDEX</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>12</td>
<td>16</td>
<td>Outcome1</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>13</td>
<td>17</td>
<td>Outcome2</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>14</td>
<td>18</td>
<td>z1</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>15</td>
<td>19</td>
<td>z2</td>
<td>0.98649</td>
<td>0.98649</td>
<td>0.98649</td>
<td>0.98649</td>
<td>0.98649</td>
<td>0.98649</td>
<td>0.98649</td>
</tr>
<tr>
<td>20</td>
<td>21</td>
<td>z3</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>22</td>
<td>23</td>
<td>z4</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>24</td>
<td>25</td>
<td>z5</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>26</td>
<td>27</td>
<td>z6</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>28</td>
<td>29</td>
<td>z7</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>30</td>
<td>31</td>
<td>z8</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>32</td>
<td>33</td>
<td>z9</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
</tbody>
</table>

Solver Parameters:
- Equal Rest: 0
- Type of: 0
- Subject to the Constraints:
  - SQ_INDEX: 0
  - Outcome1: 0
  - Outcome2: 0
- Constraints: 0
Sub Macro1()
    Application.ScreenUpdating = False
    NewSheet = Worksheets.Add
    NewSheet.Name = "Temp"
    Sheets("Temp").Select
    Range("A2").Select
    Sheets("Model").Select
    NoofChapters = Cells(1, 1)
    For n = 1 To NoofChapters
        Sheets("Model").Select
        Range("B3").Select
        Selection.Value = n
        Range("C18:C51").Select
        Selection.Value = 0
        Dim RowofCellToAdd
        RowofCellToAdd = Str(n + 17)
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        Dim CellToAdd
        CellToAdd = "C" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        Range("C7:C10").Select
        Selection.Value = 1
        Range("C11:" & CellToAdd).Select
        Selection.Value = 0
        SolverAdd (CellToAdd)
        SolverSolve (True)
        Range("C3:C51").Select
        Selection.Copy
        Sheets("Temp").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
            False, Transpose:=True
        Range("A2").Offset(n, 0).Select
    Next n
    For n = 1 To NoofChapters
        Sheets("Model").Select
        RowofCellToAdd = Str(n + 17)
        LengofRowofCellToAdd = Len(RowofCellToAdd)
        CellToAdd = "C" & Right(RowofCellToAdd, LengofRowofCellToAdd - 1)
        SolverDelete (CellToAdd)
    Next n
End Sub
## APPENDIX E: Sensitivity Analysis Model I without Environmental Harshness Constraint for Pilot Data

(volume)

<table>
<thead>
<tr>
<th>ECODE</th>
<th>SCORE</th>
<th>FNDRSN_E</th>
<th>FRAISREV</th>
<th>AGGHHINC</th>
<th>Pct_Minori</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Actual</td>
<td>Slack</td>
<td>Target</td>
<td>Actual</td>
</tr>
<tr>
<td>5132</td>
<td>1.00</td>
<td>186,613.00</td>
<td>0.00</td>
<td>$186,613.00</td>
<td>513,425.00</td>
</tr>
<tr>
<td>5180</td>
<td>1.00</td>
<td>2,039,211.00</td>
<td>0.00</td>
<td>$2,039,211.00</td>
<td>7,812,053.00</td>
</tr>
<tr>
<td>5264</td>
<td>1.78</td>
<td>1,022,436.00</td>
<td>0.00</td>
<td>$1,022,436.00</td>
<td>4,463,398.00</td>
</tr>
<tr>
<td>5503</td>
<td>1.00</td>
<td>1,379,087.00</td>
<td>0.00</td>
<td>$1,379,087.00</td>
<td>10,613,322.00</td>
</tr>
<tr>
<td>10060</td>
<td>1.00</td>
<td>398,730.00</td>
<td>0.00</td>
<td>$398,730.00</td>
<td>3,910,004.00</td>
</tr>
<tr>
<td>12024</td>
<td>3.34</td>
<td>348,803.00</td>
<td>0.00</td>
<td>$348,803.00</td>
<td>713,950.00</td>
</tr>
<tr>
<td>15080</td>
<td>1.34</td>
<td>123,067.00</td>
<td>0.00</td>
<td>$123,067.00</td>
<td>755,368.00</td>
</tr>
<tr>
<td>22172</td>
<td>2.28</td>
<td>134,954.00</td>
<td>0.00</td>
<td>$134,954.00</td>
<td>507,013.00</td>
</tr>
<tr>
<td>22322</td>
<td>2.84</td>
<td>1,049,960.00</td>
<td>0.00</td>
<td>$1,049,960.00</td>
<td>2,494,295.00</td>
</tr>
<tr>
<td>25150</td>
<td>1.60</td>
<td>406,969.00</td>
<td>0.00</td>
<td>$406,969.00</td>
<td>1,993,832.00</td>
</tr>
<tr>
<td>29042</td>
<td>1.00</td>
<td>34,097.00</td>
<td>0.00</td>
<td>$34,097.00</td>
<td>270,419.00</td>
</tr>
<tr>
<td>35072</td>
<td>3.59</td>
<td>806,883.00</td>
<td>199,910.09</td>
<td>$606,972.91</td>
<td>1,046,267.00</td>
</tr>
<tr>
<td>35168</td>
<td>1.80</td>
<td>508,955.00</td>
<td>0.00</td>
<td>$508,955.00</td>
<td>2,285,015.00</td>
</tr>
<tr>
<td>38360</td>
<td>1.00</td>
<td>1,168,798.00</td>
<td>0.00</td>
<td>$1,168,798.00</td>
<td>7,170,390.00</td>
</tr>
<tr>
<td>45010</td>
<td>2.72</td>
<td>66,901.00</td>
<td>0.00</td>
<td>$66,901.00</td>
<td>183,513.00</td>
</tr>
</tbody>
</table>
VITA

Alexandra Medina-Borja was born in Quito, Ecuador in the Andean region of South America. At the age of 2 and until the age of 17, she moved along with her parents to Central America where she attended school in San Salvador, El Salvador. After finishing high school back in her native Quito, she attended the School of Engineering at the Universidade Federal de São Carlos in São Paulo, Brazil, where in 1989 she got her B.S. in Production Engineering with a minor in Materials Engineering. Always attracted by the management of change and the application of quality and performance management tools, she held from 1990 to 1994 a position as National Expert in Quality Control and Product Development in a United Nations’ International Trade Centre in Ecuador. In 1993, she was invited by the Japanese government to be part of the “Ship for the World Youth Program” and sailed around the world as an ambassador of good will. In 1994, she left her job with the United Nations to pursue her Masters of Engineering Administration at the Virginia Tech’s Northern Virginia Center. She graduated in 1995 and in 1996 was admitted into the Ph.D. program in Industrial and Systems Engineering and moved with her husband Armando to Blacksburg, Virginia. There their two children Felipe and Camile were born in 1997 and 1999 while she was completing her doctoral academic requirements. In 1999, she took a break to dedicate time to baby Camile while the family moved to Tegucigalpa, Honduras, as part of a relief operation after Hurricane Mitch. Since January of 2001, Alexandra holds a Senior Research Associate position in one of the largest social service organizations in the United States where she developed and implemented a performance measurement system. She is a Certified Quality Manager by the American Society for Quality. What Alexandra enjoys the most is spending time with her kids, dancing to Latin music with her husband and traveling. One of her most precious memories are those shared with the many good friends from all over the world that she has met throughout her life. For all the experiences and cultures to what she has been exposed she considers herself an extremely privileged human being.