ABSTRACT

LEE, CHONGMYOUNG. The Adoption of Outcome Measures in Nonprofit Organizations: Empirical Analysis of Human Service Nonprofits. (Under the direction of Dr. Richard Clerkin).

The measurement of nonprofit performance has received increasing scholarly attention. According to the logic model, nonprofit organizations' performances are generally measured by inputs, processes, outputs, and outcomes. Literature consistently has found that there is a variation in the adoption of outcome measurement. This study investigates what factors explain this variation in the adoption of outcome measurement in human service nonprofits, using agency theory, institutional theory, and the measures of organizational capacity. Also, I examine the actual use of the outcome information at a preliminary level.

To accomplish the purposes of this study, I collected the data via an online survey – Qualtrics. I sent out my survey link to 828 human service nonprofit organizations: 182 of P32 (Foster Care Organizations), 467 of P43 (Family Violence Shelters), and 179 of P46 (Family Counseling Organizations). 317 organizations consented to participate in my survey (38.29%), but 54 out of the 317 organizations did not finish the survey. Finally, 263 (31.76%) of the responses were usable in data analysis. Using binary logistic regression, the survey data was analyzed.

The results of logistic regression yield several meaningful findings. First, the risk aversion of nonprofits negatively influences the adoption of outcome measurement. Second, the institutional factors – funding from the United Way (Coercive), imitation of the other largest organizations' practices (Mimetic), and accreditation from external parties

(Normative isomorphism) – are positively related to the adoption of outcome measurement. Lastly, human resource capacity and political capacity of organizations have a positive and significant influence on the adoption of outcome measurement.

This study extends the performance measurement research by considering the risk implication of agency theory. In addition, this research indicates how the outcome measurement of human service nonprofits is shaped by institutional factors. Lastly, this study implies the importance of human resource capacity and political capacity in the adoption of the outcome measurement of nonprofits.

The Adoption of Outcome Measures in Nonprofit Organizations: Empirical Analysis of Human Service Nonprofits

by
Chongmyoung Lee

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Public Administration

Raleigh, North Carolina

2014

APPROVED BY:

_____	_____
Richard M. Clerkin, Ph.D. Committee Chair	James E. Swiss, Ph.D.
_____	_____
Mary Tschirhart, Ph.D.	Mark A. Wilson, Ph.D.

DEDICATION

This dissertation is dedicated to God and my family.

To God, since I received the calling to study public administration in 1999, God has always stayed with me and guided my way.

To my wife, Gina Song, without your backing and devotion, completing this dissertation would not have been possible.

To my parents and sister, your unconditional love has always been my greatest support and that is why I stand here.

Thank you all.

BIOGRAPHY

Chongmyoung Lee earned his bachelor's degree in Public Administration in 2006 from Sung Kyun Kwan University, Seoul, South Korea. After graduation, he worked for the National Assembly of the Republic of Korea, working first-hand in policy making and budgeting processes.

After years of work in the National Assembly, he came to the USA to study public administration at the graduate level. He received a Master's degree of Public Service and Administration in 2010 from Texas A&M University with a concentration of nonprofit management.

He began the Ph.D. program in the Department of Public Administration at North Carolina State University in 2010.

His teaching and research interests include quantitative method, performance measurement and performance management, inter-sectoral collaboration, organizational theory, public finance, and public and nonprofit management.

ACKNOWLEDGMENTS

I would like to thank my committee members, Drs.Richard Clerkin, James Swiss, Mary Tschirhart, and Mark Wilson, for their guidance, support, and expertise. I have always enjoyed learning under their direction.

I offer my special thanks to Dr.Richard Clerkin who served not only as the chair of my dissertation committee but also as my advisor. His professionalism, sincere guidance, and mentoring have helped me finish this dissertation and become a better researcher.

Lastly, I gratefully acknowledge my Ph.D. cohort, particularly Sitthichai Tanthasith. Their friendship was a source of my energy and motivation throughout the Ph.D program.

TABLE OF CONTENTS

LIST OF TABLES viii

LIST OF FIGURES x

CHAPTER 1: INTRODUCTION 1
1.1 Introduction 1
1.2 Logic Model 1
1.3 The Adoption of Outcome Measurement 2
1.4 The Use of Outcome Information 3
1.5 Contribution of the Dissertation 4
1.6 Summary and Preview of Upcoming Chapters 7

CHAPTER 2: LITERATURE REVIEW AND HYPOTHESES 9
2.1 Introduction 9
2.2 Background 9
2.2.1 Current State of Performance Measurement in Nonprofits 9
2.2.2 The Value of Balanced Use of Performance Measures 12
2.2.3 The Use of Outcome Information 13
2.2.4 Innovation Adoption and Diffusion 15
2.3 Literature Search Strategy 19
2.3.1 Search Terms and Criteria for Inclusion/Exclusion 19
2.3.2 Literature Review Databases 19
2.4 Defining Outcome Measurement 20
2.5 Agency Theory 21
2.5.1 Risk Aversion 22
2.5.1.1 Risk Aversion and Outcome Measurement 24
2.6 Institutional Theory 28
2.6.1 Coercive Isomorphism 29
2.6.1.1 Coercive Isomorphism and Outcome Measurement 31
2.6.2 Mimetic Isomorphism 32
2.6.2.1 Mimetic Isomorphism and Outcome Measurement 34
2.6.2.2 Uncertainty 37
2.6.3 Normative Isomorphism 39
2.6.3.1 Normative Isomorphism and Outcome Measurement 40
2.7 Organizational Capacity 43
2.7.1 Organizational Capacity and Outcome Measurement 47
2.8 Other Factors Impacting Outcome Measurement 53
2.8.1 Organizational Age 54
2.8.2 Organizational Size 56

2.10 Conclusion 57
2.10.1 Summary 57
2.10.2 Preview of Chapter 3 58

CHAPTER 3: METHODLOGY 59
3.1 Introduction 59
3.2 Research Design 59
3.2.1 Database 59
3.2.2 Sampling Method: Census 60
3.3 Operationalization of Variables 62
3.3.1 Dependent Variable: Adoption of Outcome Measurement 62
3.3.2 Independent Variable: Risk Aversion 65
3.3.3 Independent Variable: Mandate 67
3.3.4 Independent Variable: Mimetic Pressure 67
3.3.5 Independent Variable: Professionalism 69
3.3.6 Independent Variable: Organizational Capacity 70
3.3.7 Control Variables: Organizational Age and Size 73
3.4 The Extent of Outcome Information Use on Selected Dimensions 73
3.5 Pilot Study 74
3.6 Data Analysis 78
3.7 Summary 78

CHAPTER 4: RESULTS 79
4.1 Introduction 79
4.2 Survey response rate 79
4.3 Data quality check 80
4.4 Representatives of respondent group 82
4.4.1 Total expenses 82
4.4.2 Regional distribution 85
4.5 Data analysis 90
4.5.1 Descriptive statistics 90
4.5.2 Outcome measurement rate by time 114
4.5.3 Missing Data Imputation 115
4.5.3.1 Comparing the group with and without missing information 116
4.5.3.2 Multiple Imputation 118
4.5.4 Analysis Models 119
4.5.4.1 Ordinal Regression 119
4.5.4.2 Poisson Regression 120
4.5.4.3 Binomial Regression 121
4.5.4.4 Logistic Regression Model 123
4.5.5 Hypothesis testing and discussion 130

CHAPTER 5: IMPLICATION AND CONCLUSION 145
5.1 Introduction 145
5.2 Use of outcome information 145
5.3 Implication 157
5.4 Limitations 163
5.5 Future Research 165

REFERENCES 170

APPENDICES 206

APPENDIX A: DOCUMENTS USED IN SURVEY 207
A-1: Institution Review Board Approval 207
A-2: Informed Consent Form 208
A-3: Survey Instrument 210

LIST OF TABLES

Table 2.1: Literature Review Databases 20
Table 4.1: Survey response rate 80
Table 4.2: Total expenses of respondents and population group 82
Table 4.3: One sample t-test of all groups 83
Table 4.4: One sample t-test, P32 83
Table 4.5: One sample t-test, P43 84
Table 4.6: One sample t-test, P46 85
Table 4.7: Cross tabulation, Region, Population and Respondents 86
Table 4.8: Chi-Square test, All Groups 86
Table 4.9: Cross tabulation, Region, Population and Respondents (P32) 87
Table 4.10: Chi-Square test, P32 87
Table 4.11: Cross tabulation, Region, Population and Respondents (P43) 88
Table 4.12: Chi-Square test, P43 88
Table 4.13: Cross tabulation, Region, Population and Respondents (P46) 89
Table 4.14: Chi-Square test, P46 89
Table 4.15: Extent of outcome measurement 90
Table 4.16: Outcome indictors of P32 (Foster care organization) 91
Table 4.17: Outcome indicator of P43 (Family violence shelters) 92
Table 4.18: Outcome indicator of P46 (Family counseling organizations) 93
Table 4.19: Number of outcome measures (Count data) 94
Table 4.20: Number of outcome measures (Dichotomous) 94
Table 4.21: Risk aversion, survey item1 95
Table 4.22: Risk aversion, survey item2 96
Table 4.23: Risk aversion, survey item3 96
Table 4.24: Reliability Statistics of Risk Aversion 97
Table 4.25: Performance information, survey item1 98
Table 4.26: Performance information, survey item2 99
Table 4.27: Coercive isomorphism, survey item1 100
Table 4.28: Coercive isomorphism, survey item2 100
Table 4.29: Mimetic isomorphism, survey item1 101
Table 4.30: Mimetic isomorphism, survey item2 101
Table 4.31: Reliability Statistics of Mimetic Isomorphism 102
Table 4.32: Uncertainty, survey item1 102
Table 4.33: Uncertainty, survey item2 103
Table 4.34: Reliability Statistics of Uncertainty 103
Table 4.35: Professionalism, survey item1 104
Table 4.36: Professionalism, survey item2 105
Table 4.37: Professionalism, survey item3 105
Table 4.38: Financial capacity (ER), FY2011 106

Table 4.39: Human Resource Capacity 107
Table 4.40: Infrastructure Capacity 107
Table 4.41: Administrative Capacity 108
Table 4.42: Political support, survey item1 109
Table 4.43: Political support, survey item2 109
Table 4.44: Political support, survey item3 110
Table 4.45: Reliability Statistics of Political Support 110
Table 4.46: Political Negotiation 111
Table 4.47: Reliability Statistics of Political Capacity 111
Table 4.48: Network Capacity 112
Table 4.49: Organizational Age 112
Table 4.50: FTE 113
Table 4.51: Total expenses 113
Table 4.52: Outcome Measurement Training 114
Table 4.53: The Adoption of Outcome Measurement by Period 114
Table 4.54: Independent Samples Test (Total Expenses) 117
Table 4.55: Independent Samples Test (FTE) 117
Table 4.56: Independent Samples Test (Organizational Age) 117
Table 4.57: Little's MCAR test 118
Table 4.58: Descriptive statistics of outcome measurement rate 120
Table 4.59: Kolmogorov-Smirnov Test 121
Table 4.60: Probability of each outcome measure 123
Table 4.61: Recoding scheme 124
Table 4.62: Logistic Regression Model 127
Table 4.63: Similarity between results from the regression with and without missing data 129
Table 4.64: Hosmer and Lemeshow Test 129
Table 4.65: Omnibus Tests of Model Coefficients 130
Table 4.66: Adjusted B coefficient and odds ratio of accreditation 136
Table 4.67: ER of respondents and population group 137
Table 4.68: One sample t-test, ER (all groups) 138
Table 4.69: One sample t-test, ER (P32) 138
Table 4.70: One sample t-test, ER (P43) 139
Table 4.71: One sample t-test, ER (P46) 139
Table 5.1: Use of outcome information 146
Table 5.2: Rotated Component Matrix, Outcome Information Use 149
Table 5.3: Correlation between outcome information use and risk aversion 151
Table 5.4: Correlation between outcome information use and institutional factors 153
Table 5.5: Correlation between outcome information use and organizational capacity 156

LIST OF FIGURES

Figure 2.1: Risk and adoption of outcome measurement 28
Figure 4.2: Outcome Measurement Rate by Period 115
Figure 4.3: Overall summary of missing values 116

CHAPTER 1: INTRODUCTION

1.1 Introduction

Financial and competitive pressures within the nonprofit sector have led to an increased emphasis on performance measurement. As a result of a growing emphasis on accountability in government and philanthropic funding, nonprofits are under increasing pressure to demonstrate excellence in performance in order to secure financial resources similar to their public and private counterparts (Moxham, 2009; Cairns et al., 2005). In conjunction, the current economic climate has had an adverse effect on the amount of public donations made to the nonprofit sector (McLean & Brouwer, 2010) suggesting that government grants and contracts will play a more prominent role in the general financial portfolio of the nonprofit sector. As a result, the pressure for nonprofit organizations to demonstrate their performance in order to secure continued funding is acute (Martikke, 2008). Correspondingly, the measurement of nonprofit performance has received increasing scholarly attention.

1.2 Logic Model

One of the main performance measurement tools is the logic model, which is defined as "a tool used to help identify the linkage between program activities and outputs to desired outcomes" (The Urban Institute, 2002, p.3). According to the logic model, nonprofit organizations' performances are generally measured by inputs, processes, outputs, and outcomes (The Urban Institute, 2002). Input measures include the number of staff required

for an activity, diversity of revenue streams, and net surplus of financial reserves. Process measures include organizational capacity, employee education, and board performance. Output can be measured by the number of people served, on time service deliveries, and the number of services provided. Lastly, outcome measures include improved condition of people served, increased skill and knowledge of people served, and customer satisfaction. Literature argues that the balanced use of performance measures is crucial (e.g., Perrin, 1998; Kaplan & Norton, 1996; Moore, 2003). According to Moore (2003) "Nonprofit managers are probably going to need a mix of outcome, output, process and input measures to allow them to recognize value in what they are doing, and find ways to improve their performance" (p.15).

1.3 The Adoption of Outcome Measurement

There are a number of obstacles to outcome measurement in the nonprofit sector. For instance, measuring outcomes is too expensive and too slow to provide necessary information about how the organization is performing to managers (Curristine, Lonti, & Joumard, 2007; Moore, 2003). Furthermore, as Moxham (2009) insisted, "As the majority of the performance criteria stipulated by the funders focused on short-term measures, such as targets and evidence of attendance, there is little incentive for nonprofits to demonstrate the effects of their services. Yet such evidence, often referred to as outcome or impact, is seen by some as key to voluntary sector accountability" (p. 6). While output indicators are comparatively easy to measure, outcomes are hard to measure because the outcome measurement requires a long time to achieve and is somewhat intangible.

The absence of outcome measures, however, can bring about goal displacement. If nonprofits use only output measures, once they achieve the outputs, they may not care about longer term performance, that is, the outcome. Bohte and Meier (2000) also argue that excessive emphasis on the output measures is problematic because "attaching importance to these measures may lead to goal displacement as agencies focus on generating numbers that please political officials, rather than devoting their energies to achieving more meaningful policy outcomes" (p.174). This goal displacement can lead to mission drift in organizations. That is to say, the nonprofits may move away from their original mission.

Some organizations prefer output indicators to outcome because outputs can be easily measured. On the other hand, other nonprofits adopt the outcome measurement because its absence can bring about goal displacement. In this sense, Thomson (2010) found that there is a variation in the extent of outcome measurement across states and metropolitan areas. My study investigates what factors explain this variation in the adoption of outcome measures, focusing on human service nonprofits.

1.4 The Use of Outcome Information

The main purpose of this study is to examine the adoption of outcome measures. My subsequent research question is how the nonprofits use the outcome information produced by the outcome measurement. Nonprofit organizations have invested substantial resources for measuring their performance. According to prior studies, however, not all nonprofits make use of the benefits of performance data (e.g., LeRoux & Wright, 2010). The substantive

usefulness of performance measurement can be examined by investigating the use of performance information (Propper & Wilson, 2003; Van Dooren, 2008).

Julnes and Holzer (2001) argue that "The utilization of performance-measurement information is a process in which at least two stages can be identified - adoption (development of measures) and implementation (actual use)" (p.703). In this dissertation, I mainly focus on the adoption of the outcome measurement. Thus, I examine the second question – actual use of the outcome information – at only a preliminary level. I summarize the research questions of my study as follows.

Research questions:

1. What factors influence the adoption of outcome measurement in human service nonprofits?
2. How do human service nonprofits use the outcome information produced by the outcome measurement?

1.5 Contribution of the Dissertation

This dissertation has seven main contributions to the field by filling the gap in the literature as follows. First, the literature consistently has found that the adoption of outcome measurement varies across the organizations. However, the reason that there is a variation in the adoption of outcome measurement has not been investigated in-depth (e.g., Zimmermann & Stevens, 2006). In my dissertation, I investigate the factors which influence the adoption of the outcome measures in nonprofits.

Second, there are empirical studies that examine the factors influencing the use of outcome measures (e.g., MacIndoe & Barman, 2012; Barman & MacIndoe, 2012; Thomson, 2010; Julnes & Holzer, 2001; Melkers & Willoughby, 2005). I investigate the factors which have not been examined in the prior studies (e.g., risk aversion, political capacity, human resource capacity, network capacity).

Third, although there are studies that examine the use of outcome measures, some of the studies do not have a theoretical discussion to support their findings (e.g., LeRoux & Wright, 2010; Zimmermann & Stevens, 2006; Folz, Abdelrazek, & Chung, 2009; Melkers & Willoughby, 2005; Poister & Streib, 1999). My study addresses how agency theory, institutional theory, and the measures of organizational capacity can be used to advance the field of performance measurement in nonprofits.

Fourth, some of the empirical studies examining the use of performance measures are based on a sample of a particular city or state (e.g., MacIndoe & Barman, 2012; Barman & MacIndoe, 2012; Thomson, 2010; Zimmermann & Stevens, 2006). Thus, it requires caution in applying the research findings to nonprofits in the other states or cities. Since my study is not limited to a particular city or state, it is more generalizable to all states or cities than prior studies based on a specific city or state. However, I focus on 3 sub-industries - foster care organizations (NTEE code P32), family violence shelters (NTEE code P43), and family counseling organizations (NTEE code P46) - in the human services field so that I will be looking at similar performance measures. In addition, my study is not applicable to small sized nonprofits. I will elaborate on this in the method section again.

Fifth, Van Dooren (2008) argues that "If we want to study the successes and failures of performance movements, we have to study the use of performance information." (p.22). There are studies which address the use of performance information in the public organizations (e.g., Moynihan & Pandey, 2010; Folz, Abdelrazek, & Chung, 2009; Julnes & Holzer, 2001; Melkers & Willoughby, 2005). Similar empirical studies should be conducted in the nonprofit organizations. Since the main focus of this study is the adoption of outcome measures, I examine the use of outcome information at a preliminary level. However, this study can be a starting point for future studies with regard to the success or failure of the performance measurement in the nonprofit organizations.

Lastly, in this dissertation, I examine the use of outcome measures more in depth than the prior studies. The use of performance measures can be examined by the development and the actual uses of the performance measures. According to Julnes and Holzer (2001), "The utilization of performance-measurement information is a process in which at least two stages can be identified - adoption (development of measures) and implementation (actual use)" (p.703). Some studies, however, examine only either the adoption of performance measurement (e.g., MacIndoe & Barman, 2012; Barman & MacIndoe, 2012; Thomson, 2010) or the use of performance information (e.g., LeRoux & Wright, 2010; Moynihan & Pandey, 2010). In this dissertation, I focus on the adoption of outcome measures, and develop the research with a preliminary analysis of the outcome information use in the nonprofits.

1.6 Summary and Preview of Upcoming Chapters

Chapter 1 introduced the research questions and background of this study. This dissertation has two topics. First, it examines the adoption of outcome measurement. There is a variation in the adoption of outcome measurement. However, the literature has not explained what factors lead to this variation. Thus, I examine what factors influence the adoption of outcome measurement based on agency theory, institutional theory, and the measures of organizational capacity. Second, although public and nonprofit organizations invest a lot of resources for outcome measurement, the use of outcome information has been comparatively neglected (LeRoux & Wright, 2010; Moynihan & Pandey, 2010). In this sense, the second topic of my research is the use of outcome information. I investigate the extent to which nonprofits use the outcome information. Lastly, Chapter 1 describes the contributions of this dissertation to the field of nonprofit studies.

Chapter 2 contains the literature review and the hypotheses of this study. First, I review the current state of performance measurement in the nonprofits and the value of balanced uses of performance measures. Then, I review the main tenets of agency theory, institutional theory, and organizational capacity. Next, I review prior studies related to the main variables of this study. Lastly, I develop the hypotheses of this study.

Chapter 3 provides the methodology of this study. First, I describe the research design of this dissertation. The primary data source of this study is the National Center for Charitable Statistics (NCCS). Through the NCCS Core Data File, I choose 1,039 nonprofit organizations - foster care organizations (NTEE code P32), family violence shelters (NTEE code P43), and family counseling organizations (NTEE code P46) - and get the main data for

this study by survey. Next, I explain how the variables of this study are operationalized. Then, I describe how the use of outcome information will be examined in this study through the survey.

Chapter 4 begins with the interpretation of the survey results. First, I explain how I controlled the data quality, and address which hypotheses are supported or not supported. In this chapter, I describe what analytic models I tried and how I ended up with the logistic regression model as the best model to test the hypotheses. Lastly, I discuss my research findings.

In chapter 5, I examine the theoretical and practical implications of this study. Also, I address the limitations of this study. Lastly, I suggest future research topics based on my research findings.

CHAPTER 2: LITERATURE REVIEW AND HYPOTHESES

2.1 Introduction

In the previous chapter, I described the performance measurement in the nonprofit organizations, logic model as a tool for measuring performance, the adoption of outcome measurement, the use of outcome information, and the contributions of this study. In this chapter, I review the current state of performance measurement in the nonprofits, the value of balanced use of performance measures, the prior studies related to the main variables of this study and develop the hypotheses. Recapping Chapter 1, according to the logic model, nonprofit organizations' performances are generally measured by inputs, processes, outputs, and outcomes (The Urban Institute, 2002). In this study, I investigate what factors affect the adoption of outcome measurement in the nonprofits, using agency theory, institutional theory, and the measures of organizational capacity.

2.2 Background

2.2.1 Current State of Performance Measurement in Nonprofits

According to Moxham (2009), "As the majority of the performance criteria stipulated by the funders focused on short-term measures, such as targets and evidence of attendance, there is little incentive for nonprofits to demonstrate the effects of their services" (p. 6). In addition, Poole et al. (2000) argue that "Over the past twenty-five years, these agencies have been largely held accountable for outputs (i.e., number of units of service and number of people served) rather than outcomes (i.e., changes in participant knowledge, attitudes, values,

skills, behavior, condition, or status). As a result, many of these agencies are having trouble today making the transition to "performance measurement," which includes outputs and outcomes (Newcomer, 1997)" (p.16). In fact, according to Sheehan (1996), only 14% of 101 Franklin County, Ohio nonprofits paid attention to outcome measures. Also, in 1998 survey of 900 nonprofits, Wiener, Kirsch, and McCormack (2002) found that most nonprofit organizations prefer output measures to outcome measures and they have lots of problems in measuring outcomes.

Lindgren (2001) argues that this is attributed to the nature of performance measurement system. "Due to the nature of measures in performance-management systems, the frequency with which data are often collected, and the very strong emphasis on results in today's public sector, the primary focus of performance measures tends to be programme output" (Lindgren, 2001, p.293). Curristine, Lonti, & Joumard (2007) and Moore (2003) argue that measuring outcomes is too expensive and too slow to provide necessary information about how the organization is performing to managers.

On the other hand, there are research findings that considerable percentage of nonprofits uses the outcome measures. For instance, Fine, Thayer, and Coghlan (2000) found that 56% of nonprofits conducted outcome measurement in the past three years. Poole et al. (2000) found that 60.3% of nonprofit agencies in Central Florida had valid outcome indicators. Also, 65% of nonprofits in South Carolina indicated that they used the outcome measures in the performance evaluation process (Zimmerman & Stevens, 2006). Carman and Fredericks (2008) found that 62% of nonprofits in Indiana gathered outcome information of their programs on a regular basis. According to Carman (2009), 60% of nonprofits in New

York reported collecting outcome measures. Lastly, MacIndoe and Barman (2012) found that 72% of nonprofits in the greater Boston area use the outcome measurement.

Thomson (2010) expects that "There is significant room for growth in the percentage of nonprofits measuring outcomes" (p. 614). As described above, the percentage of nonprofits claimed to be measuring the outcomes has been dramatically increased over time. I suggest a couple of interpretations for this increase as follows. First, as nonprofit organizations may learn the value of outcome measurement through professional networks or education, the adoption of outcome measurement might substantively be increased in the nonprofit organizations over time. The value of outcome measurement and its diffusion process are elaborated in the following sections: section 2.2.2 (The value of balanced use of performance measures) and section 2.6.3 (Literature review of normative isomorphism).

Second, some scholars who explain the adoption of organizational practices with mimetic isomorphism argue that "over time, imitation turns more and more into replication" (Roy & Seguine, 2000, p. 455). As early adopters are perceived more successful and nonadopters are perceived less legitimated, there is an "increasing degree of conformity" over time in the adoption process of organizational practices (Westphal, Gulati, & Shortell, 1997; Roy & Seguine, 2000, p. 455). In this sense, some scholars argue that the imitative process of organizational practice has a "self-reinforcing cycle" (Abrahamson & Bartner, 1990; Wezel & Saka-Helmhout, 2006, p.270). I elaborate how the mimetic pressure influences the adoption of outcome measurement in the section 2.6.2 (mimetic isomorphism).

Lastly, the increasing demand for nonprofits' accountability might reinforce the trend of outcome measurement. As I mentioned in the introduction section, nonprofits are under

increasing pressure to demonstrate their performance and accountability due to the financial and competitive pressures within the sector as well as current economic recession. Consequently, performance measurement is increasingly receiving an attention of practitioners as well as scholars (e.g., Cairns et al., 2005; Paton, 2003; Poister, 2008; Moxham, 2009). In addition, Poole et al. (2000) and Newcomer (1997) argue that current performance measurement of nonprofits are undergoing transition from output focused measurement to the mix of output and outcome measures. In this sense, it is reasonable to infer that nonprofits are increasingly adopting the outcome measurement over periods. All these arguments, however, should be to be supported by empirical findings.

2.2.2 The Value of Balanced Use of Performance Measures

The absence of outcome measures can bring about goal displacement. If nonprofits use only output measures, once they achieve the outputs, they may not care about longer term performance, that is, the outcomes. Bohte and Meier (2000) argue that "The major problem with evaluating agency performance based on outputs rather than outcomes is that attaching importance to these measures may lead to goal displacement as agencies focus on generating numbers that please political officials, rather than devoting their energies to achieving more meaningful policy outcomes" (p.174). In addition, Lindgren (2001) insists that focusing only on outputs can bring about some flaws. One of the main problems is poor validity. That is to say, "the measure or set of measures defined does not truly represent program goals and therefore does not measure what they are intended to. Many (if not most) activities in the public sector, by their very nature, are complex and intangible, and cannot be reduced to a

numeric figure. That which is measured thus often bears little resemblance to what is relevant." (p. 295).

Moxham (2009) indicates that although most nonprofits focus on short-term measures, outcomes is a "key to voluntary sector accountability" (p. 6). In this sense, Perrin (1998) suggests the use of multiple indicators as a strategy for the effective use of performance measures; "Use multiple indicators, in order to examine a variety of program aspects, including process as well as outputs and outcomes" (p. 376). In addition, Moore (2003) argues that "Nonprofit managers are probably going to need a mix of outcome, output, process and input measures to allow them to recognize value in what they are doing, and find ways to improve their performance" (p.15).

However, some nonprofits do not use the outcome measures because they are too expensive and hard to get necessary information, while some organizations use the outcome measurement because its absence can lead to goal displacement. That is to say, there is a variation in the adoption of outcome measurement. According to Thomson (2010), a variation in the adoption of outcome measurement has been found across states and metropolitan areas. However, it has not been investigated enough what factors cause this variation. This study seeks to fill this gap in the literature by investigating the factors which influence the adoption of outcome measurement in the nonprofits.

2.2.3 The Use of Outcome Information

Thomson (2010) argues that "Although increasing the extent of OM among nonprofits is a laudable goal, the more important goal for those who seek to use OM to

enhance accountability and program performance is to increase nonprofits' utilization of OM in decision making" (p.624). While nonprofits have invested substantial resources toward measuring their performance, there is a lack of research regarding the use of performance information in the nonprofit organizations (LeRoux & Wright, 2010). This common phenomenon emerges in the public sector as well; "Although governments have devoted a great deal of energy and resources into creating performance information systems, they have largely neglected the question of how to foster information use" (Moynihan & Pandey, 2010, p. 850).

Organizations ostensibly adopt the performance measurement because they assume that they will use the performance information produced by measuring their performance (Moynihan, 2005; Wholey, 1999; LeRoux & Wright, 2010). According to LeRoux and Write (2010), "Performance data can provide managers with valuable insights into organizational strengths and weaknesses, thereby equipping managers with information to guide strategic decision making" (p.575). Only some nonprofits, however, think about performance measurement as a strategic management tool. Others think about performance measurement as an external, promotional tool, or even as "resource drain and distraction" (Carman & Fredericks, 2008, p.52). This finding implies that not all nonprofits make use of the benefits of performance data (LeRoux & Wright, 2010). Propper and Wilson (2003) argue that although performance measurement has been widely adopted, its usefulness is not examined. The first step to examine the usefulness of performance measurement is to investigate to what extent organizations use the performance information. Van Dooren (2008) argues that

"If we want to study the successes and failures of performance movements, we have to study the use of performance information" (p.22).

Moynihan and Pandey (2010) investigate the use of performance information in public organizations. Similar empirical study is needed for nonprofit organizations. Folz, Abdelrazek, and Chung (2009) found that chief executives' reasons for measuring the performance (e.g., to improve management decisions, to appraise personal performance, to support budget decisions, and to comply with government requirements). They did not, however, examine how the outcome information is used for these goals. Hence, in this study, I examine how the nonprofits use the performance data, focusing on outcome information.

2.2.4 Innovation Adoption and Diffusion

My research question is related to the regular adoption literature as I investigate the factors that influence the adoption of an organizational practice – outcome measurement. There are some common themes between the innovation adoption literature and my study. However, I use different theoretical approach from the regular adoption studies. In this section, I describe how the adoption and diffusion of innovative practices has been studied and how they are similar to and different from my study.

Adoption and diffusion of innovation has been studied in a variety of contexts. First, some studies examine what factors influence the adoption or adoption process of innovation. Damanpour and Schneider (2009) examine the association between innovation characteristics, manager characteristics, and innovation adoption in the U.S. local governments. Goldenberg, Han, Lehmann, and Hong (2009) investigate how the hubs, "people with an exceptionally

large number of social ties", influence the adoption and diffusion of innovation (p.1). Ko, Kim, Kim, and Woo (2008) explore how the organizational characteristics influence the adoption of Customer Relationship Management (CRM) in the Korean fashion industry. Kwon, Berry, and Feiock (2009) examine what factors influence the adoption of economic development tools in the U.S. local governments, comparing early and late adopters. The authors investigate how the factors influence the policy adoption differently in the early adoption and late adoption. Zhu, Dong, Xu, and Kraemer (2006) explore how the innovation characteristics and contextual factors influence the post-adoption stages of innovation diffusion, focusing on enterprise digital transformation.

Second, some studies focus on the motivation of innovation adoption and diffusion. For instance, Zattoni and Cuomo (2008) explore what motivations drive the adoption of good governance code by comparing institutional and efficiency perspectives. Jun and Weare (2011) explore what role institutional motivation played in the adoption of E-government in the American municipalities. The authors find how the externally oriented motivation and internally oriented motivation influence differently the innovation adoption process. Teodoro (2010) argues that bureaucratic mobility is a main condition of innovation diffusion. The main finding of their study is that the motivation of innovation is stronger when the executives are promoted from outside than they are promoted from inside.

Third, there are some studies with regard to the mechanism, process, or pattern of innovation diffusion. For example, Zhu, Kraemer, and Xu (2006) explore the assimilation process of internet based E-business innovation by firms. The authors identify three assimilation stages – initiation, adoption, and routinization - and examine how technological,

organizational, and environmental factors influence each stage. Shipan and Volden (2008) identify the patterns of policy diffusion by investigating antismoking policies by the U.S. Cities and how the patterns depend on the internal and external context. Straub (2009) examine the process of technology adoption by incorporating adoption theories: Rogers's innovation diffusion theory, the Concerns-Based Adoption Model, the Technology Acceptance Model (TAM), and the United Theory of Acceptance and Use of Technology. Ansari, Fiss, and Zajac (2010) investigate how the corporate practice varies during diffusion process and the pattern of practice adaptation depends on the "fit between the diffusing practice and the adopting organization" (p.68). Yu and Tao (2009) examine the factors influencing business level innovation technology adoption, identifying each stage of innovation diffusion - pre-decision, in-decision, and post-decision.

Lastly, a lot of innovation studies focus on specific case of innovation adoption or technology acceptance. For instance, the case of E-government adoption has been discussed by several scholars (e.g., Jun & Weare, 2011; Carter & Weerakkody, 2008). In addition, the adoption of innovative practice has been examined in diverse industries. Kennedy and Fiss (2009) investigate TQM adoption among U.S. hospitals. Krause (2011) examines the climate protection initiatives by U.S. cities. Ko, Kim, Kim, and Woo (2008) explore the Customer Relationship Management (CRM) adoption in the Korean fashion industry. Some studies focus on information or communication technology. Park and Chen (2007) examine the adoption of the innovative use of smart phone. Mallat et al. (2009) focus on the case of mobile ticketing. Al-Senaidi, Lin, and Poirot (2009) examine the barriers to adopting information and communication technologies (ICT) in Omani context. Hsu, Lu, and Hsu

(2007) focus on the adoption of mobile internet. Some studies are conducted using a case of specific industry. Dearing and Kreuter (2010) apply the innovation diffusion theory to the diffusion of cancer communication innovations. Lee (2004) examines nurse's adoption of technology. Tung, Chang, and Chou (2008) investigate the adoption of the electronic logistics information system in Hospital Information System (HIS) in the medical industry. Soderholm and Klaassen (2007) analyze the innovation and diffusion in the European wind power sector.

As shown above, the adoption of innovative practices or technologies has been studied in diverse contexts. There are some commonalities and differences between regular adoption studies and my research. First, as other adoption literature does (e.g., Damanpour & Schneider, 2008; Goldenberg, Han, Lehmann, & Hong, 2009; Ko, Kim, Kim, & Woo, 2008; Kwon, Berry, & Feiock, 2009; Zhu, Dong, Xu, & Kraemer, 2006), I investigate what factors influence the adoption of innovative practices – adoption of outcome measurement. Second, as many adoption studies focus on specific industries (e.g., Kennedy & Fiss, 2009; Krause, 2010; Ko, Kim, Kim, & Woo, 2008; Park & Chen, 2007; Mallat et al., 2009), I focus on the US human service nonprofit organizations. Third, some studies examine the stages or process of innovation adoption (Zhu, Kraemer, & Xu, 2006; Straub, 2009; Ansari, Fiss, & Zajac, 2010; Yu & Tao, 2009). In my study, I examine the two different stages of utilization of outcome measurement - adoption and implementation of outcome measurement - although the main focus of this study is the adoption. I explain these separated stages with a concept of "loose coupling" (Meyer & Rowan, 1977; DiMaggio & Powell, 1983) in the literature review section. However, my study is distinguishable from the regular adoption studies in that while

they mainly use Rogers' innovation diffusion theory or TAM theory, I use organizational theories - agency theory and institutional theory - to explain the variation in the adoption of outcome measurement. Since theoretical approaches are different, my study provides differentiated theoretical implications from those of regular adoption literature. Before I review the literature regarding the adoption of outcome measurement, I explain my literature search strategy in the following section.

2.3 Literature Search Strategy

2.3.1 Search Terms and Criteria for Inclusion/Exclusion

The search terms include "performance measurement", "outcome measurement", "performance measures", "nonprofit", and each independent variable (e.g., risk aversion, professionalism). This literature review mainly includes studies about performance measurement in nonprofits, especially the use of performance measures in nonprofits and the composition of performance measurement in nonprofits.

2.3.2 Literature Review Databases

As shown in Table 2.1, I review the literature based on the following four databases.

Table 2.1: Literature Review Databases

Database	Description
Social Science Citation Index	Covers more than 50 disciplines related to the social sciences
Academic Search Premier	Includes indexing and abstracts for more than 8,500 journals, with full text for more than 4,600 of those titles.
Public Administration Abstracts	Contains bibliographic records covering essential areas related to public administration, including public administration theory, and other areas of key relevance to the discipline
JSTOR	Includes the arts, humanities, and social sciences complements research in core disciplines such as political science, sociology, public policy, and administration.

2.4 Defining Outcome Measurement

To measure the extent of outcome measurement at the organizational level, Thomson (2010) developed an index. He considers "the number of expected outcomes identified, whether the organization had a valid means to assess the expected outcomes (by level of outcome), and whether the organization reported actual outcomes that could be supported by the organization's assessment tools" (Thomson, 2010, p.621). Haigh et al. (2001) construct a list of the most common outcome measures by literature review. Then, they ask respondents to indicate how many times they use each outcome measure with their regular assessments. They found a variation in the use of outcome measures in each group. Some studies require the respondent to list the outcome measures they use. For instance, Hatfield and Ogles (2004) ask respondents "Do you use some form of outcome assessment in your practice?" in the

survey (p. 486). Then, the researchers require survey respondents to indicate what outcome measures they are using. Turner-Stokes and Turner-Stokes (1997) require respondents to list the outcome measurement scales if they use standardized outcome measures. Copeland, Taylor, and Dean (2008) also select three most commonly used outcome measures and ask respondents whether they use it. Some studies measure the extent of outcome measurement by time period of outcome data. For example, Botcheva, White, and Huffman (2002) make a list of data, including outcome. They require the respondents check what type of data they collect. The type of data collected includes service outcomes (short term outcomes) and extended outcomes data (long term). Lastly, MacIndoe and Barman (2012) measure the extent of outcome measurement, using a question "In the last fiscal year, what percentage of your organization's programs/services were subject to outcome measurement?" (p.9) Then, they compare high program use and low program use. High program means that the organization measures the outcome of more than 50% of their services. Conversely, low program means the organization measure the outcome of less than 50% of the services.

2.5 Agency Theory

According to agency theory, "Agents and principals are motivated by opportunities for their own personal gain" (Davis, Schoorman, & Donaldson, 1997, p.22). Agency theory has two assumptions: "(1) There is goal conflict between the power (bureaucratic and political) and budget/wealth-maximizing behavior of the principal and the utility-maximizing behavior of the agent and (2) Agents have more information than principals, which agents

can exploit for self-gain rather than for the collective interests of the contracting parties leading to moral hazard problems" (Van Slyke, 2007, p. 162).

Also, Agency theory postulates that individuals and organizations are in principal-agent relationship, "in which one party (the principal) delegates work to another (the agent), who performs that work" (Lambright, 2009; Eisenhardt, 1989, p. 58). In the nonprofit studies using agency theory, nonprofit organizations are usually considered as the agent. The principals include government, donors, funders, or other resource providers (e.g., Guo & Acar, 2005; Lambright, 2009). As an agent, a nonprofit organization may measure its performance to meet the principal(s)'s requirements and to demonstrate its accountability. In this sense, the traits of the relationship between principal(s) and nonprofits may influence the adoption of outcome measures in nonprofits. Agency theory, thus, explains how the adoption of outcome measurement depends on the relationship between nonprofits and principal(s). In this study, I examine how the risk aversion of nonprofits influences the adoption of outcome measurement, using Agency theory. According to Van Slyke (2007), "Agency theory is explicit about redistributing risk through a variety of incentive and sanction mechanisms" (p.176). More details of this relationship are described in the hypothesis section below.

2.5.1 Risk Aversion

Risk aversion has been studied in a variety of contexts. First, some scholars focus on measuring the risk attitude (e.g., Holt & Laury, 2002; Abdellaoui, Driouchi, & L'Haridon, 2011). For instance, Holt and Laury (2002) suggest a method to measure the risk aversion, using the results of lottery choice experiment. Abdellaoui, Driouchi, and L'Haridon (2011)

suggest a simple way to measure risk aversion by generalizing Holt and Laury (2002)'s multiple prices method. In addition, some studies examine the risk aversion in the perspective of international relations, international politics, or public policy choices (e.g., O'Neill, 2001; Vertzberg, 1995; Boettcher, 2004; Ehrlich & Maestas, 2010; Krause, 2003; Kam & Simas, 2010). Third, some scholars compare the risk aversion of public organizations and private organizations (e.g., Bellante & Link, 1980; Bozeman & Kingsley, 1998). Lastly, risk aversion is studied in public choice, public good game, or rent seeking games (e.g., Millner & Pratt, 1991; Teyssier, 2012; Yamazaki, 2009; Treich, 2010; Münster, 2006; Millner & Pratt, 1991).

Some studies regarding risk aversion is conducted at individual level (e.g., Teyssier, 2012; Kam & Simas, 2010; Kam & Simas, 2012; Grund & Sliwka, 2010; Ehrlich & Maestas, 2010; Druckman & McDermott, 2008; Miller & Hoffmann, 1995; Van de Werfhorst & Hofstede, 2007; Bellante & Link, 1981), while others are studied at organizational level (e.g., Sanders & Hambrick, 2007; Vasi & King, 2012; Bozeman & Kingsley, 1998; Krause, 2003; Fiegenbaum & Thomas, 1988; Leyden & Link, 1993; Sanders & Hambrick, 2007). Also, risk attitude is examined both as a dependent variable (e.g., Druckman & McDermott, 2008; Wiseman & Gomez-Mejia, 1998; Boettcher, 2004; Vasi & King, 2012; Sanders & Hambrick, 2007; Bellante & Link, 1981) and as an independent variable (e.g., Miller & Hoffmann, 1995; Sitkin & Pablo, 1992; Van de Werfhorst & Hofstede, 2007; Kam & Simas, 2012).

To summarize, literature has studied the risk aversion with a variety of purposes (e.g., risk attitude measurement, international relations, public policy choice, and organizational comparison). Some researchers found that risk aversion was a significant factor in their

empirical studies (e.g., Miller & Hoffmann, 1995; Van de Werfhorst & Hofstede, 2007; Kam & Simas, 2012). In the next paragraphs, I examine the prior studies which are directly related to the purpose of my research, investigating how risk aversion influences the adoption of outcome measurement.

2.5.1.1 Risk Aversion and Outcome Measurement

Prior studies have not examined how the risk aversion influences the adoption of outcome measurement, but some studies have examined how the risk aversion is related to the performance based contracts. For instance, Eisenhardt (1989) develops propositions regarding how individual factors, including risk aversion, are related to the behavior or outcome based contracts using agency theory. She hypothesizes that risk averse nonprofits are less likely to have outcome based contracts. Literature supports this hypothesis in that outcome based contracts shift the risk to the agent, while behavior based contracts shift the risk to the principal (e.g., Logan, 2000; Bergen, Dutta, & Walker, Jr., 1992; Jiang & Qureshi, 2006).

In addition, Grund and Sliwka (2006) examine how risk aversion influences the incentive contracts. The authors argue that a risk averse agent perceives the benefits of performance pay less than a risk taking agent due to the costs of performance measurement. Thus, a risk averse agent prefers fixed wages to performance pay. The empirical finding of this study supports the authors' argument. This finding implies that the perception of the costs of performance measurement is influenced by the level of risk aversion. Although the

authors do not elaborate on the costs of performance measurement, outcome measurement typically is more costly than the other types of measurements (e.g., Moore, 2003).

In addition, some studies investigate how individuals' risk attitude influences their choices. For instance, Kam and Simas (2010) examine the relationship between individuals' risk orientation and their policy preferences. The authors find that "Those who were more risk acceptant were more likely to prefer policies with probabilistic outcomes, whereas those who were more risk averse were more likely to prefer policies that promised sure things" (Kam & Simas, 2010, p.393). That is to say, risk aversion is negatively related to the uncertainty surrounding policy consequences. Literature argues that outcome measurement involves more uncertainties than output measurement because there are a lot of "outside uncontrolled variables" influencing the outcomes (Bowornwathana, 2000, p.399). In this sense, Sawhill and Williamson (2001) argue that even if the outcomes of the Girl Scouts of the USA (GSUSA) and the Boy Scouts of America (BSA) are measured, it cannot be guaranteed that the Scouting was the only reason for their outcomes.

Based on the literature discussed above, it is concluded that there is a significant relationship between risk aversion and the adoption of outcome measurement, with the effect being in the negative direction. According to agency theory, agents tend to be more risk averse than principals. Eisenhardt (1989) argues that "agents, who are unable to diversify their employment, should be risk averse and principals, who are capable of diversifying their investments, should be risk neutral" (p. 60-61). In this study, however, I relax this assumption. That is to say, I assume that there is a variation in the risk attitude of agents.

Shapiro (2005) argues that "some agents are more risk averse than others" (p. 267). Also, MacCrimmon and Wehrung (1988) found that individuals' attitude toward risk vary widely.

One of the main issues that agency theory is concerned with is the risk sharing between principal and agent. This problem arises "when the principal and agent have different attitudes toward risk" (Eisenhardt, 1989, p. 58). In my study, "risk" means perils of unfavorable results coming from taking the responsibility for the final outcomes of nonprofits' services/programs. If nonprofits make a contract based on outcome and clients are not satisfied with the final results of the services/programs, the nonprofits may lose the grant/funding because they agreed to be responsible for the final outcome in the contracting process. This would bring about financial problems for the nonprofits and lead to negative impacts on their reputation and activities in the field. More risk averse nonprofits might perceive these perils more seriously than less risk averse nonprofits.

Next, literature using agency theory has explored the principal-agent relationship with a concept of contracts, outcome based contracts or behavior based contracts (Jensen & Meckling, 1976; Eisenhardt, 1989). In this study, I use the contracts as a mediating variable to explain the relationship between risk aversion of nonprofits and the adoption of outcome measurement. Attitude toward risk aversion might influence nonprofits' decision to use an outcome based contracts or a behavior based contracts, which might influence the adoption of outcome measurement. This relationship is elaborated in following paragraphs.

Outcome based contracts transfer risk to the agent, that is, nonprofits. Thus, as the nonprofits are more risk averse, it is hard for principals to pass risk to the nonprofits using an outcome based contracts (Eisenhardt, 1989). In this sense, Eisenhardt (1989) hypothesizes

that the risk aversion of nonprofits is negatively related to outcome based contracts and positively related to behavior based contracts.

If nonprofit organizations are responsible for the final outcomes as a result of the outcome based contracts, they would be willing to measure the outcomes because measurement is one of the major means for achieving organizations' final goals (Tornow, 1993). According to Amaratunga and Baldry (2002), "Measurement provides the basis for an organization to assess how well it is progressing toward its predetermined objectives" (p.218). On the contrary, nonprofits contracting based on behavior might be less motivated to adopt the outcome measurement because they are not responsible for the final outcomes in the contracts and outcome measurement is generally too expensive and too slow (Curristine, Lonti, & Joumard, 2007; Moore, 2003).

In brief, given that the other factors influencing the outcome measurement are controlled, if nonprofit organizations are risk averse, they would be less likely to make a contract based on outcomes. This relationship is shown in Figure 2.1. Hence, I hypothesize as following:

Hypothesis 1: The risk aversion of the nonprofit organization is negatively related to the adoption of outcome measurement

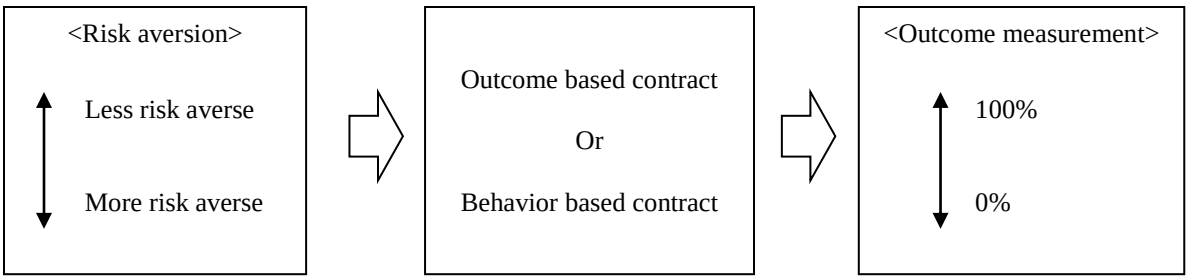

Figure 2.1: Risk and adoption of outcome measurement

2.6 Institutional Theory

"Organizations compete not just for resources and customers, but for political power and institutional legitimacy" (DiMaggio & Powell, 1983, p.150). According to the DiMaggio and Powell (1983), organizations achieve the legitimacy through isomorphic processes: 1) Coercive isomorphism: "Coercive isomorphism results from both formal and informal pressures exerted on organizations by other organizations upon which they are dependent and by cultural expectations in the society within which organizations function" (p. 150); 2) Mimetic isomorphism: "When goals are ambiguous, or when the environment creates symbolic uncertainty, organizations may model themselves on other organizations" (p.151); 3) Normative isomorphism: It stems primarily from professionalization which is defined as "collective struggle of members of an occupation to define the conditions and methods of their work" (p. 152). In this study, I use institutional isomorphism to explain the adoption of outcome measurement in nonprofit organizations. More details on this relationship are described in the hypothesis section below.

As I mentioned in the introduction section, the utilization of outcome measurement is composed of two different stages - the adoption and implementation of outcome measurement. According to the prior studies, although public and nonprofit organizations invest a great deal of resources for adopting outcome measurement, the use of outcome information - substantive implementation of outcome measurement - has been neglected (LeRoux & Wright, 2010; Moynihan & Pandey, 2010). Institutional theory literature explains the gap between the adoption and implementation of outcome measurement with "loose coupling between legitimated external practices and internal organizational behavior" (Meyer & Rowan, 1977; DiMaggio & Powell, 1983, p.155). Meyer and Rowan (1977) argue that "Organizations reflecting institutional rules tend to buffer their formal structures from the uncertainties of technical activities by becoming loosely coupled" (as cited in Collier, 2001, p.468). Organizational behavior to get institutional legitimacy and technical activities actually performed inside the organization are interdependent but different (Orton & Weick, 1990; Fogarty & Rogers, 2005). In this sense, I examine the adoption of outcome measurement and the actual use of outcome information in a separated section.

2.6.1 Coercive Isomorphism

The mandate from public, clients, resource providers, or other stake holders for performance measurement has been discussed in many studies. Some scholars examine the relationship between the mandate from stakeholders and innovative practices (e.g., Guo & Acar, 2005; Aristigueta & Foote, 2009; Carter Jr, Jambulingam, Gupta, & Melone, 2001; Arndt & Bigelow, 2000; Delimas, 2002; Ketokivi & Schroeder, 2004). First, Guo and Acar

(2005) use institutional theory to infer how government mandate in the funding process influences the nonprofits' decision to collaborate. Aristigueta and Foote (2009) examine how the federal mandate promotes performance contracts arrangement as a mechanism for improving the accountability of service providers. Next, Carter Jr, Jambulingam, Gupta, and Melone (2001) investigate how bidding on contracts which mandate the use of the innovation is associated with the innovation adoption. Arndt and Bigelow (2000) examine how coercive pressure is related to the hospitals' structural innovation. Delimas (2002) associates the diffusion of environmental management standards, ISO 14001, in the Europe and United States with coercive isomorphism. Lastly, Ketokivi and Schroeder (2004) discuss how the adoption of innovative manufacturing practices (e.g., ISO 9000) can be explained by examining coercive isomorphism.

To summarize, literature has studied how coercive isomorphism influences organizational behaviors, performance, or practices (e.g., collaboration, performance contracts, structural innovation, innovative manufacturing practices). Some researchers found that coercive pressure was a significant factor in their empirical studies (e.g., Guo & Acar, 2005; Aristigueta & Foote, 2009; Arndt & Bigelow, 2000; Delimas, 2002; Ketokivi & Schroeder, 2004). In next paragraphs, I examine the prior studies which are directly related to my study, investigating how coercive isomorphism influences the adoption of outcome measurement.

2.6.1.1 Coercive Isomorphism and Outcome Measurement

According to Frumkin (2005), philanthropic foundations require recipients to adopt the outcome measurement in a grant making process. The relationship between resource providers' mandate and the adoption of outcome measurement is empirically tested in many studies (e.g., Thomson, 2010; MacIndoe & Barman, 2012; Barman & MacIndoe, 2012). First, Thomson (2010) investigates how funders' outcome reporting mandates affect the extent of outcome measurement among nonprofits. Based on resource dependence theory, the author argues that nonprofits are likely to act in accordance with funders' requirements because of their dependence on the funders' resources. The research findings support the author's hypothesis. That is to say, a performance reporting mandate from funders is positively associated with the extent of outcome measurement. Also, MacIndoe and Barman (2012) examine the role of resource providers in the implementation of outcome measurement. They find that funders and external stakeholders drive the budget allocation for outcome measurement. In other research, the authors argue, "Some resource providers in the nonprofit sector require that nonprofits engage in OM as a condition of their funding" (p.79) and find that organizations supported by philanthropic foundation are more likely to use outcome measurement (Barman & MacIndoe, 2012).

As described in the literature review above, prior studies argue that coercive isomorphism influences organizational behavior and practice, more specifically, the adoption of outcome measurement. Carman (2009) posits that nonprofit organizations adopt outcome evaluation practices because they are required by resource providers (e.g., federal and state government). In the grant making process, the resource providers may require nonprofits to

measure the outcomes in the performance evaluation process. For instance, the national United Way encourages local United Ways to measure program outcomes (MacIndoe & Barman, 2012; Julian & Kombarakaran, 2006; Plantz, Greenway, & Hendricks, 1997). Also, philanthropic foundations require outcome evaluation of their grantees to decide projects to provide financial support (MacIndoe & Barman, 2012; Benjamin, 2010; Carson, 2000; Frumkin, 2005).

As a coercive isomorphic pressure, the mandate from higher authorities may provide organizations with an impetus for their behaviors and practices. That is to say, the resource providers' mandate for the outcome measurement may force nonprofit organizations to measure their outcomes. For instance, organizations might be externally mandated to measure outcomes as a condition for program funding. Hence, I hypothesize as following:

> *Hypothesis 2:* The more program funding that is tied to mandates to collect outcome measurement, the greater the adoption of outcome measurement throughout the organization.

2.6.2 Mimetic Isomorphism

The mimetic behavior has been studied in a variety of contexts. First, some scholars examine how mimetic behavior influences the performance or profitability (e.g., Brouthers, O'Donnell, & Hadjimarcou, 2005; Barreto & Baden-Fuller, 2006). For example, Brouthers, O'Donnell, and Hadjimarcou (2005) hypothesize that emerging market firms improve their performance by imitating generic product strategies in triad nation markets, and the authors find that their hypothesis is supported by empirical testing on a sample of Romanian and

Chinese exporters. On the other hand, Barreto and Baden-Fuller (2006) find that the mimetic behavior is negatively related to the profitability by examining Portuguese bank branching decisions between 1988 and 1996.

In addition, some scholars examine how the market entry is influenced by mimetic pressure (e.g., Haveman, 1993; Greve, 2000). For instance, Haveman (1993) finds that firms' entry into new market is influenced by large and profitable organizations' behavior. According to DiMaggio and Powell (1983), "Organizations imitate other organizations that are or are perceived by organizational decision makers to be successful" (as cited in Haveman, 1993, p.620). Also, Greve (2000) examines the market entry decision in Tokyo Banking. The author finds that the entry decision is influenced by large organizations' actions.

Lastly, there are some studies examining the effect of mimetic isomorphism on diffusion of innovative practices (e.g., Arndt & Bigelow, 2000; Guler, Guillen, & Macpherson, 2002; Renzulli & Roscigno, 2005; Radaelli, 2000; Benders, Batenburg, & Blonk, 2006). Arndt and Bigelow (2000) address the effects of institutional pressures, including mimetic pressure, on the process of hospital's structural innovation. Guler, Guillen, and Macpherson (2002) examine the spread of ISO 9000 Quality Certificates, focusing on coercive, mimetic, and normative effects. Renzulli and Roscigno (2005) examine the adoption of charter school legislation and find that there is a mimetic tendency among neighboring states for adopting the charter school legislation. Radaelli (2000) discusses how the uncertainty leads to the "imitation of the (perceived) most successful national model of monetary policy", which results in a paradigm in the policy process of the Economic and

Monetary Union (p. 32). Lastly, Benders, Batenburg, and Blonk (2006) argue that the mimetic pressure coming from media and trend influences the adoption of Enterprise Resource Planning (ERP) systems.

To summarize, literature has studied how mimetic isomorphism influences organizational behavior, performance, or practice (e.g., market entry, profitability, structural innovation, adoption of innovative system). Some researchers found that mimetic isomorphism was a significant factor in their empirical studies (e.g., Brouthers, O'Donnell, & Hadjimarcou, 2005; Haveman, 1993; Greve, 2000; Renzulli & Roscigno, 2005; Benders, Batenburg, & Blonk, 2006). In next paragraphs, I focus on studies which are directly relevant to my research hypothesis regarding the relationship between mimetic isomorphic pressure and the adoption of outcome measurement.

2.6.2.1 Mimetic Isomorphism and Outcome Measurement

First, Barman and MacIndoe (2012) hypothesize that mimetic isomorphism would be positively related to the adoption of outcome measurement. However, the authors find that the relationship between mimetic pressure and adoption of outcome measurement is insignificant. Authors explain this insignificant result with Tolbert and Zucker (1983)'s argument that mimetic isomorphism is valid only when the practice is dominant in the field. According to their survey result, only 45% of their sample adopted the outcome measurement to evaluate more than 50% of their services. Authors argue that the mimetic isomorphism is

not valid in their empirical results because outcome measurement is not widespread yet in the field.

I expect that my study result might be different from that of Barman and MacIndoe (2012) due to the scope of survey samples. Barman and MacIndoe (2012) collected the survey samples in diverse sub-industries: human service, education, arts/culture, environment/animals, health, and public benefit. On the other hand, I focus only on human service nonprofits. Literature argues that human service nonprofits are increasingly adopting the outcome measurement (Carrilio, Packard, Clapp, 2004; Julian & Clapp, 2000; Horsch, 1996). Also, MacIndoe and Barman (2012) found that human service nonprofits adopt the outcome measurement more than the other sub-industries' nonprofits. This means that outcome measurement is comparatively dominant in human service nonprofits. Focusing only on human service nonprofits might produce a significant result because the outcome measurement is relatively widespread in the human service area, compared to the other sub-industries.

Also, Hussain and Hoque (2002) examine the design and use of the non-financial performance measurement system of Japanese banks, using institutional theory. Authors' findings do not support the proposition that organizations have a tendency to copy best practices from others. According to Hussain and Hoque (2002), "The greater the established link between strategy, management activities and accounting control system, the less the tendency of the organization to copy organizational systems such as a PMS from other organizations" (p. 177).

As described in the literature review above, prior studies argue that mimetic isomorphism influences the organizational behavior and practice although it was not a significant factor in some of the empirical studies (e.g., Barman and MacIndoe, 2012). According to the institutional theory, organizations are influenced by isomorphic pressure in the process of achieving legitimacy. One of the isomorphic processes is mimetic isomorphism "resulting from standard responses to uncertainty", which means that organizations imitate those with practices they admire to be sure that they are doing the right things (DiMaggio & Powell, 1983, p.150).

If organizations perceive that the other organizations experience a success by adopting outcome measures, this might make a mimetic pressure to imitate the organization's successful practices. DiMaggio and Powell (1983) argue that "Organizations tend to model themselves after similar organizations in their field that they perceive to be more legitimate or successful" (p.152). Once organizations perceive the other organizations' adoption of some practice and their success, the perception serves as a "strong impetus for the mimetic process" to the remaining nonadopters (Westphal, Gulati, & Shortell, 1997; Roy & Seguine, 2000, p.454). This makes a mimetic pressure to imitate other organizations' practices because organizations have a concern about the legitimacy. As non adopters are perceived as less legitimated, the imitation is replicated more and more over time and the degree of conformity is increased in the adoption process (Westphal, Gulati, & Shortell, 1997; Roy & Seguine, 2000).

Folz et al (2009) argue that some nonprofits use the outcome measures because they do not want to be seen as less progressive than other organizations. By conforming to the

mimetic pressure, organizations can gain legitimacy. Bolman and Deal (1991) argue, "Even if the change is only *pro forma*", the adoption of the outcome measures make the organizations look proactive, competent, and responsive (as cited in Roy & Seguine, 2000, p.455). Roy and Seguine explain this process in the case of public organizations as following: "The adoption of fads makes the leadership of an organization appear as oriented toward modern issues and methods. This presents public organizations as attentive to emerging issues and highly efficacious in solving new problems with modern management methods" (p.455). Their argument can be applied to the nonprofit organizations as well because public and nonprofit organizations have similar organizational activities and behaviors. Nonprofits can enhance their legitimacy through the mimetic process described above. Hence, I hypothesize as following:

> *Hypothesis 3:* The mimetic pressure from the other organizations in the same field is positively related to the adoption of outcome measurement.

2.6.2.2 Uncertainty

Next, uncertainty should be considered in the relationship between mimetic pressure and the adoption of outcome measurement. DiMaggio and Powell (1983) argue, "Organizations mimic or model each other when they face uncertainty, and they look for answers to their uncertainty in the way in which other organizations in their field have faced similar uncertainties" (As cited in Ellis & Shpielberg, 2003, p. 1246). When there is a high degree of uncertainty, managers try to model themselves after the other organizations in the same field. Mizruchi and FeinSource (1999) argue, "In situations in which a clear course of

action is unavailable, organizational leaders may decide that the best response is to mimic a peer that they perceive to be successful" (p.657). The uncertainty makes organizations to adjust themselves by comparing their practices to other organizations' bevaiors (Festinger, 1954; Haunschild, 1994). That is to say, the uncertainty leads to the mimetic behavior, and organizations take on the legitimated practices by the imitation (Radaelli, 2000; DiMaggio & Powell, 1983; Haunschild, 1994). According to DiMaggio and Powell (1983), "When organizational technologies are poorly understood (March & Olsen, 1976), when goals are ambiguous, or when the environment creates symbolic uncertainty", organizations may model themselves after the others (p. 151). In this sense, literature indicates that the uncertainty causes the mimetic behavior.

Mimetic isomorphism is a result of "standard responses to uncertainty" (DiMaggio & Powell, 1983, p.150). Thus, organizations are more likely to imitate other organizations' behaviors perceived to be successful when the relationship between means and ends is more uncertain (DiMaggio & Powell, 1983). As shown above, I hypothesize a positive relationship between mimetic pressure and the adoption of outcome measurement. Uncertainty might intervene in this relationship. To put it another way, when organizations experience more uncertainty, they might be more motivated to mimic other organizations' practices – the adoption of outcome measurement. Hence, I hypothesize as following:

> *Hypothesis 4:* Uncertainty has a positive interaction effect on the relationship between mimetic isomorphic pressure and the adoption of outcome measurement.

2.6.3 Normative Isomorphism

The normative isomorphism has been studied in a variety of contexts (e.g., Wouters & Wilderom, 2008; Radaelli, 2000; Guler, Guillen, & Macpherson, 2002). First, Wouters and Wilderom (2008) hypothesize that professionalism positively influences the employees' positive perception toward performance measurement system; "If an employee is more inclined to improve her work practice, then performance measures are more likely to be seen as positive, stimulating, challenging, and helpful" (p. 495). The authors' empirical findings support their hypothesis. In addition, Radaelli (2000) examines how the institutional isomorphism influences policy transfer, using three case studies – single currency, tax policy, and media ownership policy. The author discusses how the normative pressures produced by Delors Committee shaped Economic and Monetary Union (EMU) policy according to the shared values and beliefs in the area. Lastly, Guler, Guillen, and Macpherson (2002) hypothesize that two aspects of normative isomorphism - knowledge base and cohesion in network - positively influence the spread of ISO 9000 quality certificates. The authors' empirical test results support the influence of cohesive trade relationship on the number of certificates, while it does not support the influence of scientific and technological knowledge.

To summarize, literature has studied how normative isomorphism influences organizational behaviors or practices (e.g., perception toward performance measurement system, EMU policy, adoption of ISO 9000 certification). Some researchers found that normative isomorphism was a significant factor in their empirical studies (e.g., Wouters & Wilderom, 2008; Radaelli, 2000; Guler, Guillen, & Macpherson, 2002). In next paragraphs, I focus on studies which are directly relevant to my research hypothesis regarding the

relationship between normative isomorphic pressure and the adoption of outcome measurement.

2.6.3.1 Normative Isomorphism and Outcome Measurement

Some literature argues that professionalism provides a prolific foundation for adopting new management practices and ideas (e.g., Roy & Séguin, 2000; Radaelli, 2000; DiMaggio & Powell, 1983; Fligstein, 1985; Mizruchi & Fein, 1999; Powell et al., 2005; Barman & MacIndoe, 2012). According to Roy and Séguin (2000), "As do interorganizational networks, professional networks give salience to emerging phenomena and provide fertile grounds for promoting new successful endeavors, by means such as newsletters, conferences, and so forth" (p. 456). The professional networks socialize the organizations within the networks into "similar worldviews" through a similar education and training process, and diffuse the management ideas among the members of the networks (Radaelli, 2000; DiMaggio & Powell, 1983; Fligstein, 1985; Mizruchi & Fein, 1999, p. 657). Hence, it has been argued that professional networks are "sites of new management ideas" (Powell et al., 2005; Barman & MacIndoe, 2012, p. 87).

In this sense, Barman and MacIndoe (2012) hypothesize that two aspects of professionalism, professional network and accreditation, are positively related to the adoption of outcome measurement. The results of their empirical test, however, do not support this hypothesis. The authors find that the membership in a professional network and the possession of accreditation do not significantly influence the adoption of outcome measurement. The authors explain their contradictory results as follows: while professional

networks mainly focus on latest organizational practices (e.g., Powell et al., 2005), outcome measurement is not a relatively new practice, thus, might not be actively circulated in the professional associations.

I expect that my study result would be different from that of Barman and MacIndoe (2012) due to the different methods for measuring professionalism. I measure professionalism in three ways: organization's level of education, executive directors' level of education, and accreditation. First of all, Barman and MacIndoe (2012)'s argument that outcome measurement is not actively discussed in professional networks does not have empirical evidence. Also, many universities are increasingly providing performance measurement courses in their curriculum (Poister, 2008; Russ-Eft & Preskill, 2009; LaVelle & Donaldson, 2010). In brief, since I measure professionalism in three different ways and many universities are increasingly teaching performance measurement courses, the findings of my study might be different from that of Barman and MacIndoe (2012).

As described in the literature review above, normative isomorphism influences the organizational behavior and practice although there are some insignificant findings (e.g., Barman & MacIndoe, 2012). According to DiMaggio and Powell (1987), normative isomorphism stems primarily from professionalization which is defined as "collective struggle of members of an occupation to define the conditions and methods of their work" (p. 152).

The absence of outcome measures might bring about goal displacement. If nonprofits use only output measures, once they achieve the outputs, they may not care about longer term performance, that is, outcome. In this sense, Bohte and Meier (2000) argue that nonprofits

may focus on "generating numbers that please political officials, rather than devoting their energies to achieving more meaningful policy outcomes" if their performance measurement system excessively emphasizes outputs rather than outcomes (p.174). In addition, Moxham (2009) indicates that outcome measurement is a "key to voluntary sector accountability" (p. 6). By measuring the outcome, nonprofit organizations can make sure whether they are on the right track to achieve the goals that they originally intended, and ultimately, improve their accountability.

DiMaggio and Powell (1987) address two aspects of professionalization as main sources of normative isomorphism. First is "resting of formal education and of legitimation in a cognitive base produced by university specialists" and second is "growth and elaboration of professional networks that span organizations and across which new models diffuse rapidly" (p.152). The norms and best practices of good governance may be diffused through the professional networks (MacIndoe & Barman, 2012; DiMaggio & Powell, 1983; Hwang & Powell, 2009). For instance, professional associations, including Child Welfare League of America, Big Brothers Big Sisters of America, and Girl Scouts of America, have detailed plans for measuring their outcomes and educating their member organizations about the outcome measurement (MacIndoe & Barman, 2012; Plantz et al., 1997; Zimmerman & Stevens, 2006). Through these mechanisms of professionalization, organizations may share the value of outcome measurement I described above. To put it another way, if nonprofit organizations have more staff who are educated about the performance measurement from universities, professional networks, and associations, they would be more aware of the purposes and values of outcome measurement.

Sharing the value of outcome measurement may form a normative isomorphism. According to institutional theory, it provides nonprofit organizations with a normative pressure to adopt some practices (e.g., outcome measurement). Also, organizations might get an accreditation in the process of professionalization. In the accreditation process, some nonprofit organizations might be required to adopt the outcome measurement by the industry or accreditation bodies (MacIndoe & Barman, 2012; Packard, 2010). This is another institutional pressure for nonprofit organizations to adopt the outcome measurement in their practices. Hence, I hypothesize as follows:

Hypothesis 5: Professionalism is positively related to the adoption of outcome measurement.

2.7 Organizational Capacity

Barman and MacIndoe (2012) argue that the variation in the implementation of outcome measurement can be best explained "by the inclusion of the concept of organizational capacity alongside variables drawn from new institutionalism" (p. 71). In addition, Berman and Wang (2000) indicate that examining the organizational capacity is a "useful diagnostic strategy to identify areas of organizational deficiency" for performance measurement (p. 417). In this sense, organizational capacity is often mentioned as an essential part of an effective performance measurement system (Berman and Wang 2000; Franklin 2002; Melkers & Willoughby, 2005).

Conceptualizing and measuring the organizational capacity in the case of performance measurement has been conducted in a variety of ways. For instance, Barman

and MacIndoe (2012) focus on external expectation by defining the capacity as "an organization's ability to implement externally generated expectations from their organizational field" (p. 73). In addition, focusing on the organizational mission, Eisinger (2002) defines the capacity as "a set of attributes that help or enable an organization to fulfill its missions" (p. 117). Sowa, Selden, and Sandfort (2004) explain the capacity by focusing on process and structure; "Capacity refers to how the organization or program operates, the structures in place, and the operating processes that dictate and direct employee action" (Sowa, Selden, & Sandfort, 2004, p.715). Lastly, Eisinger (2002) define the organizational capacity as "a set of attributes that help or enable an organization to fulfill its missions" (p.117).

Organizational capacity is categorized in several ways. Barman and MacIndoe (2012) categorize the capacity into administrative capacity, expert capacity, and impersonal capacity. In addition, Berman and Wang (2000) categorize it into the technical (or infrastructure) capacity and stakeholder (or political) capacity. Sowa, Selden, and Sandfort (2004) measure the capacity in dimensions of management and program. Christensen and Gazley (2008) characterize the capacity as "a function of (1) organizational infrastructure, (2) human resources, (3) financial resources and management systems and (4) political and market characteristics of the external environment" (p.268). Lastly, Eisinger (2002) examines the capacity in terms of staffing, institutionalization, and external help network.

In this study, I focus on how the organizational capacity influences the adoption of outcome measurement. Based on the prior studies, I classify the main elements of the organizational capacity as follows: financial capacity, human resource capacity,

organizational infrastructure – technical capacity, administrative capacity, political capacity, and network capacity (Tuckman & Chang, 1991; MacIndoe & Barman, 2012; Poole et al., 2001; Christensen and Gazley, 2008; Berman &Wang, 2010; Barman & MacIndoe 2012; The Urban Institute, 2002).

First, Ingraham, et al. (2003) and Christensen and Gazley (2008) argue the financial capacity as one dimension of organizational capacity. When organizations have sufficient financial capacity, they can train staff, hire outside experts to collect/analyze the performance data, and develop internal process to consistently implement organizational practices (MacIndoe & Barman, 2012; Kaplan, 2001). These form the organizational capacity.

Second, the human resources are listed as one aspect of organizational capacity by multiple scholars (e.g., Christensen & Gazley, 2008; Eisinger, 2002; Moore, 2003; Ingraham et al., 2003; Berman & MacIndoe, 2012). Poole et al. (2001) define the human resources as staff with skill and knowledge to implement organizational practices. Also, Berman and MacIndoe (2012) incorporate "the presence of professional expertise among organizational members" as one dimension of organizational capacity (p. 81).

Third element is technical capacity. According to Berman and Wang (2000), it refers to organizations' ability to "develop performance goals and measures and to overcome such conceptual barriers as distinguishing outcomes from outputs" (p. 410). In addition, they argue that many organizations do not have the capacity for conducting a survey in a scientific way, while stakeholder surveys are needed for outcome measurement (Berman & Wang, 2000). Technical (or organizational infrastructure) capacity is related to this aspect. In this sense, technical capacity explores whether the organization has adequate information

technology, communication system, and organizational structure for performance measurement system (Berman & Wang, 2010).

Next element is administrative capacity. Based on Weber's (1978) concept of Bureaucracy, Barman and MacIndoe (2012) include the administrative capacity as one dimension of organizational capacity. Administrative capacity is defined as "organization's ability to implement an institutionally sanctioned practice such as OM" (Barman & MacIndoe, 2012, p. 88). The administrative capacity enables the organizations to behave "in a consistent, centralized, and coordinated fashion" (Weber, 1978; Barman & MacIndoe, 2012, p.81).

Fifth, multiple scholars (e.g., Berman & Wang, 2000; Christensen & Gazley, 2008) mention political capacity as one of the aspects of organizational capacity. Héritier (2003) argue that "A political process is considered to have political capacity (a) when a decision can be reached without long negotiations, and (b) when it enjoys the political support of all concerned actors and therefore has a high consensus capacity" (p. 107). In this sense, I classify the political capacity into two dimensions: political negotiation and political support. First, political capacity of organization is based on its ability to negotiate with stakeholders to establish a consensus in a short period of time (Héritier, 2003). It helps ensure compliance among the stakeholders and convince them to adopt and implement organizational practices (Borzel, 2002). Second, support from top management is essential for adopting the performance measurement system because "it forecloses back channels, legitimates reforms and new performance expectations, and helps ensure funding for new efforts" (Berman & Wang, 2010, p. 410). Besides, Berman and Wang (2010) list the support for the performance

measurement of most employees, citizen, and higher organizations as a main element of political capacity.

The last element of the capacity is network capacity (Eisinger, 2002; Christensen & Gazley, 2008; Chaskin, 2001). External networks may provide nonprofits with a variety of assistance for implementing organizational practices. Eisinger (2002) indicates that utilizing this external network is one dimension of organizational capacity. In addition, Christensen and Gazley (2008) list linkage to external knowledge, collaboration of services, and support from external environment as aspects of organizational capacity. In this study, I capture these aspects as network capacity.

The details of the operationalization of the all dimensions of organizational capacity mentioned above are described in the methods section.

2.7.1 Organizational Capacity and Outcome Measurement

Organization capacity has been studied in a variety of contexts. First, there are myriad studies regarding organizational capacity building (e.g., Cosner, 2009; Joffres, et al., 2004; Jones, 2001; Sobeck & Agius, 2007; Griffith et al., 2010; Austin, et al., 2011). Since this is not the main focus of my research, however, I do not discuss the details of the organizational capacity building in this study.

Next, some studies examine the effect of organizational capacity on diffusion, results, or performance of innovative practices (e.g., Gazley, 2010; Andrews & Boyne, 2010; Judge & Elenkov, 2005; McGuire & Silvia, 2010). First, Gazley (2010) examines what dimensions of collaborative capacity support the collaborative performance. The author finds that

intensity of shared goals and the level of investment in the partnership are the most important factors that influence the performance in government-nonprofit partnership. Second, Andrews and Boyne (2010) hypothesize that the management capacity positively influences the public service performance. The empirical findings support their hypothesis. The authors find that management capacity - capital management, financial management, human resource management, and information technology management - is positively related to the local government performance in England and this impact is enhanced by leadership. Also, Judge and Elenkov (2005) address how the organizational capacity for change is associated with environmental performance - "ecological results of an organization-wide commitment (or noncommitment) to preserve and protect the natural environment (Stead &Stead, 1992; Judge & Elenkov, 2005, p.895)". The authors find a positive relationship between organizational capacity for change and environmental performance. Lastly, McGuire and Silvia (2010) examine how the organizational capacity is related to the intergovernmental collaboration. The authors find that managerial capacity and technical capacity positively influence the extent of collaborative practice.

To summarize, literature has studied how organizational capacity influences organizational performance or practices (e.g., collaborative performance, public service performance, environmental performance, intergovernmental collaboration). Some researchers found that organizational capacity was a significant factor in their empirical studies (e.g., Gazley, 2010; Andrews & Boyne, 2010; Judge & Elenkov, 2005; McGuire & Silvia, 2010). In the next paragraphs, I examine prior studies which are directly related to my

study, investigating how organizational capacity influences the adoption of outcome measurement.

First, Barman and MacIndoe (2012) classify organizational capacity into three dimensions - administrative capacity, expert capacity, and impersonal capacity - and examine how each dimension of organizational capacity influences the adoption of outcome measurement. They find that administrative capacity and expert capacity significantly and positively influence the adoption of outcome measurement, while impersonal capacity does not. Second, Berman and Wang (2000) examine the US counties' organizational capacity to implement performance measurement. The authors find that technical capacity and political capacity are positively related to the breadth and depth of performance measurement use.

In addition, although resources (e.g., financial resources, human resources) and organizational capacity are not exactly the same, scholars address organizational resources as aspects of organizational capacity (e.g., Christensen & Gazley, 2008; Eisinger, 2002). MacIndoe and Barman (2012) address how the allocation of resources for the outcome measurement results in the use of outcome measures. The authors hypothesize that nonprofits which allocate more of their budget to outcome measurement would use the outcome measurement more. However, this hypothesis is not supported by the empirical findings. Their findings indicate that some nonprofits with a sufficient budget for outcome measurement do not employ the outcome measurement, while other nonprofits without a specific budget for the outcome measurement adopt it to manage their programs.

As described above, literature argues that organizational capacity influences the organizational behavior, performance, and practice, more specifically, the adoption of

outcome measurement. Outcome measurement usually requires a lot of resources. Moore (2003) argues, "The efforts to measure outcomes are too expensive and too slow to provide comprehensive, fast feedback about how an organization is performing" (p.15). Buckmaster (1999) also indicates that "the greatest impediment to measuring outcomes is one of resource availability" (p. 186). It is reported that a lot of nonprofit organizations do not have sufficient resources for the adoption of outcome measurement (Carman & Fredericks, 2010; Barman & MacIndoe, 2012).

Organizations with insufficient organizational capacity may not have enough resources for measuring the outcome. As I mentioned above, since the outcome measurement requires a lot of resources, the lack of resources would negatively influence the adoption of the outcome measurement. In this sense, the reason why nonprofit organizations do not adopt outcome measurement might be that they do not have a capacity for the outcome measurement: expertise to develop the outcome indicators and set up the measurement system, staff to manage the measurement system, budget to hire experts on performance measurement systems, organizational infrastructure to support the system, administrative capacity to consistently implement the outcome measurement, or political support for outcome measurement (MacIndoe & Barman, 2012; Poole et al., 2001; Christensen & Gazley, 2008; Berman &Wang, 2010; Barman & MacIndoe 2012; The Urban Institute, 2002). Organizations with strong organizational capacity would have sufficient resources (e.g., money and staff), which would provide a favorable foundation for measuring the outcome. I hypothesize the relationship between the adoption of outcome measurement and each

dimension of organizational capacity described in the literature review section in following paragraphs.

First element of organizational capacity is financial capacity. MacIndoe and Barman (2012) argue that nonprofit organizations can train their staff for the appropriate use of outcome measurement, hire performance measurement experts, or develop a performance measurement/management system when they allocate sufficient budgetary recourses for the outcome measurement. To put it another way, organizations with sufficient financial resources can use the resources for the outcome measurement. Hence, I hypothesize as follows:

Hypothesis 6a: Financial capacity is positively related to the adoption of outcome measurement.

Second element of organizational capacity is the human resource capacity. Staff with knowledge about outcome measurement is needed for adopting the outcome measurement. "They can analyze data, put it in an easily understandable form, and do breakouts by key client characteristics to seek explanations for the results. These analyses make the outcome information more understandable and useful for both program staff and organization management" (The Urban Institute, 2002, p.13). Human resources are indispensable for outcome measurement because it requires skills and knowledge. In this sense, I hypothesize as follows:

Hypothesis 6b: Human resource capacity is positively related to the adoption of outcome measurement.

Another element of capacity is technical capacity (organizational infrastructure). It refers to organizational structure, information technology, and communication systems for the performance measurement. Even if nonprofits have human resources and financial resources for adopting the outcome measurement, the organizations would still need the information technology for measuring the outcome and systems/structure supporting the outcome measurement process. Hence, I hypothesize as follows:

Hypothesis 6c: Technical capacity is positively related to the adoption of outcome measurement.

Forth, I include administrative capacity as one element of the organizational capacity. The administrative capacity refers to "presence of written rules and records" (Weber, 1978; Barman & MacIndoe, 2012, 81). The written rules and procedures guiding the practices enable the organizations to "act in a consistent, centralized, and coordinated fashion" (Weber, 1978; Barman & MacIndoe, 2012, p.81). Thus, organizations with administrative capacity better adopt practices such as outcome measurement (Barman & MacIndoe, 2012). Hence, I hypothesize as follows:

Hypothesis 6d: Administrative capacity is positively related to the adoption of outcome measurement.

Fifth element of organizational capacity is political capacity. The importance of political support for adopting new practices has been discussed in the literature (e.g., Daniels Clark-Daniels, 2000; Borins, 2002; Anguelovski & Carmin, 2011; Berman & Wang, 2010; Héritier, 2003). Political capacity is essential for adopting outcome measurement because it helps assign budgets for the new practice, legitimize the practice, establish an organizational

consensus about the outcome measurement, and develop strategies to convince stakeholders and obtain their support (Berman & Wang, 2010; Héritier, 2003). Hence, I hypothesize as follows:

Hypothesis 6e: Political capacity is positively related to the adoption of outcome measurement.

The last element of the organizational capacity is network capacity. Utilizing external network has been discussed in the literature as an aspect of organizational capacity (e.g., Herman & Renz, 2008; Moore, 2003). For instance, Herman and Renz (2008) argue that effectiveness of an organization often depends on the "effectiveness of other organizations and people with which it is interconnected and the ways in which they are interconnected" (p. 409). As I mentioned above, adopting outcome measurement requires a lot of financial and human resources. Since establishing reliable resources for the outcome measurement is difficult in itself, expanding support through an external network would positively influence the adoption of outcome measurement. Hence, I hypothesize as follows:

Hypothesis 6f: Network capacity is positively related to the adoption of outcome measurement.

2.8 Other Factors Impacting Outcome Measurement

Although this study focuses on risk aversion, the resource providers' mandate, mimetic pressure, professionalism, and organizational capacity, it is important to control the other factors which can influence the dependent variable. Two control variables, organizational age and size are identified in the literature on the outcome measurement.

Hence, I control organizational age and size to remove the statistical interference of these variables on this study and to generalize the study results across all ages and sizes of human service nonprofits.

2.8.1 Organizational Age

According to the literature, organizational age and the adoption of outcome measurement can have both a positive and negative relationship. Some researchers (e.g., Poole et al., 2000; Ranger-Moore, 1997; Barman & MacIndoe, 2012) predict that the organizational age negatively influences the adoption of outcome measurement. On the other hand, according to the other studies (e.g., Fichman & Levinthal, 1991; Freeman, Carroll, & Hannan, 1983), it is possible to predict that the organizational age is positively related to the outcome measurement.

On the one hand, nonprofit organizations have focused on outputs rather than outcomes for decades (Poole et al., 2000). Organizations tend to persist the way they were founded. If some organizations were founded in the era when outcome measurement was less emphasized, they may be less likely to use it still. Since organizations usually are path dependent, it is hard to change their current practices (e.g., Garud & Karnoe, 2013). In this sense, it has been reported that lots of nonprofit organizations have problems in making a transition to balanced use of output and outcome measures (Poole et al., 2000; Newcomer, 1997). Also, Ranger-Moore (1997) argues that old organizations have more inertia coming from the time dependent processes than young organizations. To put it another way, old organizations may not use outcome measures because they did not focus on outcome

measurement in the past and they are dependent on their past practices. Lastly, Barman and MacIndoe (2012) hypothesize that age is negatively related to the use of outcome measurement as age negatively influences the organizations' ability to adapt to external expectations. The authors' empirical findings support their hypothesis. According to these studies mentioned above, older organizations might be less likely to adopt outcome measurement.

On the other hand, there is a counter argument that older organizations are more likely to use the outcome measures. Funders value the outcome measurement to verify that the funded programs are completed (Campbell, Lambright & Bronstein, 2012). Young organizations, however, need time to learn their new roles before they produce visible performance (Stinchcombe, 1965; Fichman & Levinthal, 1991; Singh, Tucker, & House, 1986; Freeman, Carroll, & Hannan, 1983; Deeds & Rothaermel, 2003). Fichman and Levinthal (1991) argue, "Both marriages and jobs require that new roles be learned" (p.443). Then, young organizations might not be expected to show results as quickly as older, more established ones. In contrast, as time goes by, the system and structure of organization stabilize. Thus, the organization becomes durable (Freeman, Carroll, & Hannan, 1983). Then, the adolescent nonprofits might have more pressure from funders to verify their outcomes because it is expected that their system is established and they already learned the new roles. Since nonprofits are dependent on the resources that funders provide, this pressure would positively influence the practice of outcome measurement. Eventually, according to these literatures, older organizations might be more likely to adopt outcome measurement. This

contradictory influence of organizational age on the dependent variable is empirically tested, using the survey results in my study (See Chapter 4: Results).

To summarize, some literature (e.g., Poole et al., 2000; Ranger-Moore, 1997; Barman & MacIndoe, 2012) predict that the organizational age negatively influences the adoption of outcome measurement. On the other hand, according to the other literature (e.g., Fichman & Levinthal, 1991; Freeman, Carroll, & Hannan, 1983), it is possible to predict that the organizational age is positively related to the adoption of outcome measurement.

2.8.2 Organizational Size

Literature expects that organizational size positively influences the adoption of outcome measurement (e.g., MacIndoe & Barman, 2012). Hoque and James (2000) argue that large organizations are more likely to rely on performance measurement because they have more needs for the effective communication through the performance measurement than small organizations. Since outcome measurement is a subcategory of performance measurement, nonprofits which rely more on performance measurement might be more likely to use outcome measures than otherwise. In addition, MacIndoe and Barman (2012) find that organizational size is positively related to the use of outcome measurement: "Nonprofits with less than US$50,000 in annual revenue are less likely to report the use of OM than nonprofits with larger revenues" (p.8). In addition, Thomson (2010) finds that the size of budget and the number of staff are positively related to the extent of outcome measurement. In this sense, literature argues that organizational size (e.g., FTEs, total expenses) is positively related to the adoption of outcome measures. To summarize, the organizational size is expected to

positively influence the adoption of outcome measurement in the literature (e.g., Thomson, 2010; MacIndoe & Barman, 2012).

2.10 Conclusion

2.10.1 Summary

In brief, outcome measurement is key to nonprofits' accountability. The absence of outcome measurement can bring about goal displacement (Bohte & Meier, 2000; Moxham 2009). Thus, the importance of outcome measurement has been addressed in the literature. However, there is a variation in the adoption of outcome measurement (Thomson, 2010). In chapter 1, I described the background, research questions, and the contributions of my study. In this chapter, I reviewed how the main factors in my study have been addressed in the literature, and how I developed the hypotheses based on agency theory, institutional theory, and the measures of organizational capacity. Following is the summary of the research hypotheses.

Hypothesis 1: Risk aversion of the nonprofit organization is negatively related to the adoption of outcome measurement

Hypothesis 2: Coercive isomorphism is positively related to the adoption of outcome measurement.

Hypothesis 3: The mimetic pressure from the other organizations in the same field is positively related to the adoption of outcome measurement

Hypothesis 4: Uncertainty has a positive interaction effect on the relationship between mimetic isomorphic pressure and the adoption of outcome measurement.

Hypothesis 5: Professionalism is positively related to the adoption of outcome measurement.

Hypothesis 6a: Financial capacity is positively related to the adoption of outcome measurement.

Hypothesis 6b: Human resource capacity is positively related to the adoption of outcome measurement.

Hypothesis 6c: Technical capacity is positively related to the adoption of outcome measurement.

Hypothesis 6d: Administrative capacity is positively related to the adoption of outcome measurement.

Hypothesis 6e: Political capacity is positively related to the adoption of outcome measurement.

Hypothesis 6f: Network capacity is positively related to the adoption of outcome measurement.

2.10.2 Preview of Chapter 3

In the next chapter, I explain how I derive my sample, operationalize the main variables of the hypotheses, and analyze the survey data.

CHAPTER 3: METHODLOGY

3.1 Introduction

In the previous chapter, I reviewed the current state of performance measurement in nonprofits, the value of outcome measures, prior studies associated with my hypotheses. In addition, I developed a series of hypotheses based on agency theory, institutional theory, and the measures of organizational capacity. In this chapter, I describe the methodology of this study.

3.2 Research Design

3.2.1 Database

The sample for this dissertation is drawn from National Center for Charitable Statistics (NCCS), which is a project of the Center on Nonprofits and Philanthropy (CNP) at the Urban Institute. The mission of NCCS is "to develop and disseminate high quality data on nonprofit organizations and their activities for use in research on the relationships between the nonprofit sector, government, the commercial sector, and the broader civil society" (NCCS, 2013). A major strength of NCCS is a standard classification system of nonprofit organizations. Although IRS has an activity code, it is not a system for categorizing different types of nonprofits in a consistent way. NCCS developed National Taxonomy of Exempt Entities (NTEE). This standard classification system provides a way for coding different types of nonprofits. In addition, NCCS verifies the code of the largest organizations in each NTEE major category every year and NCCS make corrections if they find incorrect

classifications (Lampkin & Boris, 2002). NCCS releases several types of data files on a regular basis.

In this study, I use Core Data files, which contains data on all 501 (c)(3) which are required to file a Form 990 (or Form 990-EZ) (NCCS, 2006). Since nonprofit organizations with gross receipts below $25,000 and congregations are not required to file IRS 990 form, the financial data of these organizations are not available in the Core Data files. Because I exclude small nonprofits with expenses below $500,000, however, this is not problematic in this study.

3.2.2 Sampling Method: Census

In order to maintain comparability among nonprofit organizations in the analysis, this study is restricted to human service nonprofit organizations. Since most human service nonprofits receive funding from the federal and state government but work outside the government structure, assuring accountability of human service nonprofits has received an increasing attention. Thus, human service nonprofits have conducted multiple mechanisms to ensure the accountability, including outcome measurement (Kim, 2005; Dicke, 2002; Dicke & Ott, 1999; Salamon, 1999). In addition, it is increasingly expected that outcomes are measured in human service nonprofits because policy makers look for outcome information when they make funding decisions (Carrilio, Packard, Clapp, 2004; Julian & Clapp, 2000; Horsch, 1996). MacIndoe and Barman (2012) found that human service nonprofits use the outcome measures more than the other types of nonprofits.

Since there are a variety of sub-industries in the human service nonprofits, I focus on 3 sub-industries in the human services field so that I will be looking at similar performance measures. For this purpose, I choose foster care organizations (NTEE code P32), family violence shelters (NTEE code P43), and family counseling organizations (NTEE code P46), among the all human service organizations.

In addition, I get the survey objects from mid and large sized nonprofits because small nonprofits may not have resources to construct the performance measurement system. Also, staff of small nonprofits may not understand the concepts of output and outcome measures. Using Salamon, Geller, and Spence (2009)'s method, I filter small nonprofits with expenses below $500,000.

As I mentioned above, in order to help the survey respondents understand the concepts of output and outcome measures, I provide the definitions and examples of output and outcome measures in the survey. I get these measures from the performance plan of Child and Family Services Agency (CFSA). CFSA is a public child welfare agency located in Washington D.C. whose mission is to enhance the well-being of children and families. CFSA has worked with internal and external stakeholders to develop their own performance measurement indicators and consistently measured their output and outcome with specific targets (CFSA, 2014). The examples of output measures are the percentage of children/youth in foster care receiving at least 2 visits per month from the social worker, the percentage of children discharged to permanency (reunification, guardianship, and adoption), the percentage of children receiving a full medical evaluations within 30 days of entering foster care, and the frequency of family counseling services per month, the reduced rate of family

violence, and the percentage of families achieving successful service provision within the first quarter. Next, the examples of outcome measures are the client satisfaction with respect to the family counseling services, the change in the quality of life of beneficiaries, reduced physical/mental illness, and reduced childhood behavioral and emotional disorders (CFSA, 2012).

A search of the NCCS Core Data file (public charities) finds 452 foster care organizations (P32), 920 family violence shelters (P43), and 800 family counseling organizations (P46). As I mentioned above, I filter small nonprofits with expenses below $500,000. With this screening, I find 246 foster care organizations (P32), 569 family violence shelters (P43), and 224 family counseling organizations (P46) through the NCCS Core Data file. Thus, the size of population group is 1,039. I conduct a census with this population. This study involves a collection from organizational websites of contact information for each organization's executive director. I send the survey questions to the executive directors using Qualtrics, online survey program. In case the person who gets the survey cannot accurately answer the questions, I request them to pass this survey on to the person who can best answer the questions.

3.3 Operationalization of Variables

3.3.1 Dependent Variable: Adoption of Outcome Measurement

The dependent variable is measured by asking "In the last fiscal year, what percentage of your organization's programs/services were subject to outcome measurement?" (MacIndoe & Barman, 2012, p. 9). MacIndoe and Barman (2012) compare nonprofits that

adopted the outcome measurement in 50% or more of their programs/services with organizations that adopted the outcome measurement in less than 50% of their programs/services. In order to more thoroughly measure the variation in the dependent variable, I classify the responses by five choices: 1) 0%, 2) 1% - 25%, 3) 26% - 50%, 4) 51% - 75%, and 5) 76% - 100%.

Even though I explain the output and outcome measures in the survey, some respondents might fool themselves into the socially desirable answer. To reduce this problem, I list specific outcome measures and ask if the respondents track them regularly; "How often do you collect the following information?" In addition, I customize this question by NTEE code because each respondent group has its unique outcome measures. To get the outcome measures, I review CFSA (2012), American Association for Marriage and Family Therapy (AAMFT) (2013), and nonprofit organizations' websites. AAMFT is a professional association which represents more than 25,000 marriage and family therapists in the United States, Canada, and abroad. AAMFT develops standards for professional ethics and the practice of marriage and family therapy. The examples of outcomes of family therapy and treatment are described in their website (AAMFT, 2014). For the nonprofits' websites, I select sample organizations with two criteria. First, I select organizations with considerably large amount of revenues and expenses because small nonprofits might not have enough resources for constructing their performance measurement system. Second, I select organizations with well developed websites displaying their programs, services, and activities.

For NTEE code p32 (Foster Care), I provide following measures: Increased placement stability, client satisfaction with respect to the foster care service, reduced

childhood behavioral and emotional disorders, increased quality of life, and enhanced level of education or life skills of foster child. These measures come from the AAMFT (2013) and nonprofit organizations' website in this field (e.g., Cardinal McCloskey Community Services, 2013; Youthville, 2013).

Next, for NTEE code p43 (Family Violence Shelters), I provide following measures: client satisfaction with respect to the shelter's service, reduced physical/mental illness (e.g., depression and other affective disorders), increased quality of life, increased self-efficacy, reduced childhood behavioral and emotional disorders. These measures come from CFSA 2012 and nonprofit organizations' website in this category (e.g., Safe Horizon, 2013; Sanctuary for Families, 2013).

Lastly, for NTEE code p46 (Family Counseling), I provide following measures: client satisfaction with respect to the family counseling service, increased self-efficacy, reduced physical/mental illnesses (e.g., depression and other affective disorders), increased quality of life, reduced drug and alcohol use, reduced childhood behavioral and emotional disorders, reduced marital and relationship problems, and reduced child abuse and neglect. These measures come from AAMFT (2013) and nonprofit organizations' website in this field (e.g., Family & Children's Center, 2013; Center for Human Development, 2013).

In case respondent organizations cannot find their outcome measures in the list, I add one more question, "Please write down the outcomes that you are measuring but not listed above and the frequency of measurement". This is an open ended question. If some respondent answers "never" to all outcome measures I listed and does not write any outcome measure in the open ended question, I assume the response for the dependent variable -

adoption of outcome measurement - is "1", that is, 0% of the respondent organization's programs/services was subject to outcome measurement in the last fiscal year. If the respondent chooses other options for the dependent variable, I assume that this respondent's answer is incorrect. This process will filter some of the dishonest responses and improve the reliability of the survey results. In spite of these efforts, some survey respondents might fool themselves still. Including these questions, however, may lead the respondents to answer more accurately.

An additional advantage of such questions is that some practitioner readers will be interested in just the count data. For example, the percentage of family counseling agencies tracking client satisfaction can give implications for practitioners in that field. I analyze the measurement frequency of each outcome measure by nonprofit group in the result section.

Lastly, as I mentioned in section 2.2.1 (Current state of performance measurement in nonprofits), I measure the period of outcome measurement to examine how the adoption rate of outcome measurement has been changed over time. The adoption rate of outcome measurement by time is measured by asking "Please check all periods when you were measuring outcomes." This item is measured on a five point scale (1 = Before FY1980, 2 = FY1981 - 1990, 3 = FY1991 –2000, 4 = FY2001– 2010, 5 = Since FY2011, and 0 = Not measuring outcomes (Not Applicable).

3.3.2 Independent Variable: Risk Aversion

Julnes and Holzer (2001) measure the risk taking by the existence of the systems that encourage risk taking, and by evaluating the attitudes of managers and employees toward

innovations and organizational changes. Using their method, I measure the risk aversion by three questionnaire items: 1) “There is a reward/incentive system in place in the organization that encourages risk taking (using innovative ideas with the goal of improving performance)”; 2) “Management is willing to implement appropriate organizational innovation and change”; and 3) “Nonmanagement employees willingly accept organizational innovation and change” (Julnes & Holzer, 2001, p.706). This item is measured on a five point scale (1: “strongly agree” to 5: “strongly disagree”). Next, I create a scale for organizational risk aversion by averaging the responses of these three items. Higher value on this scale means that the organization is more risk averse.

In addition, the mediating variable in the relationship between risk aversion and the outcome of outcome measurement is the type of contracts - output or outcome based (See Figure 2.1). Before I ask this question, I repeat the definitions and examples of output and outcome measures to help respondents answer the survey. This variable is measured by two questions. First question is “Does your organization have any performance based contract?” This item is measured on a dichotomous scale (0: Yes or 1: No). I code “No” 1 because ordinal regression in SPSS takes the highest (last) value as a reference category. If respondents answer “No” in this question, they can skip the second question. The second question, main question for measuring the mediating variable, is “To what extent do your performance based contracts with funders focus on outputs or outcomes?" This item is measured on a scale of 10 (1: “Totally output based” to 10: “Totally outcome based”). However, this item is revised after I conduct the pilot study. I describe the revision of this survey item in section 3.5 Pilot Study.

3.3.3 Independent Variable: Mandate

Resource providers' mandate is measured by asking "What percentage of your organization's program funding is subject to mandatory outcomes measurement?" This item is measured on a five point scale (1 = 76% - 100%, 2 = 51% - 75%, 3 = 26% - 50%, 4 = 1% - 25%, and 5 = 0%). As I mentioned above, I reversed the order because ordinal regression in SPSS takes the last category as a reference. In addition, Hendricks, Plantz, and Pritchard (2008) argue that the United Way emphasizes the outcome measurement as a main way for evaluating programs' value and encourages funding recipients to adopt the outcome measurement. Barman and MacIndoe (2012) also measured the coercive isomorphism by asking about United Way funding. In this sense, an additional question to measure the resource providers' mandate is "Does your organization receive funding from the United Way?"

3.3.4 Independent Variable: Mimetic Pressure

To operationalize the mimetic isomorphism, Fligstein (1985) examines the percentage of organizations in the same industry that adopted the practice to be studied. Mizruchi and Fein (1999) argue that Fligstein (1985)'s method is reasonable because organizations are more likely to adopt the practice when others in the same industry have adopted that. In addition, Barman and MacIndoe (2012) measure the mimetic isomorphism by asking the "use of OM by the largest organizations in the focal organization's nonprofit industry" (p.80). Based on these studies, I measure the mimetic pressure by asking "To what extent do you agree with each of the following statements? The largest organizations in our field measure

their outcomes". Also, not only the adoption of the outcome measures but also the results of the adoption may form a mimetic pressure. That is to say, if nonprofits perceive that other organizations improve their performance by adopting the outcome measurement, this perception may form a mimetic pressure. Thus, I use one more survey question: "To what extent do you agree with each of the following statements? Nonprofits in our field that measure their outcomes improve their performance". These items are measured on a five point scale (5: "Strongly disagree" to 1: "Strongly agree"). Lower score means stronger mimetic pressure.

Next, as I mentioned above, I measure the uncertainty in this study because mimetic isomorphism is a result of "standard responses to uncertainty" (DiMaggio & Powell, 1983, p.150). Using Andrews (2008), I classify the uncertainty into socioeconomic uncertainty and political uncertainty. I modify Andrews (2008)' survey items based on other studies regarding measurement of uncertainty (e.g., Budding, 2004; Hoque, 2004; Guo & Acar, 2005). The major modification is that I use only predictability, while Andrews (2008) measure complexity, predictability, and the degree of change to operationalize the environmental uncertainty. This modification is justified in the prior studies (e.g., Budding, 2004; Hoque, 2004) which only focus on predictability of the dimensions of uncertainty to measure the uncertainty. The socioeconomic uncertainty is measured by asking "Changes in our community's socioeconomic status are predictable" (reversed) (Andrews, 2008; Guo & Acar, 2005; Hoque, 2004). Next, political uncertainty is measured by asking "Changes in our community's political environment are predictable" (reversed) (Andrews, 2008; Budding, 2004; Hoque, 2004). These items are measured on a five point scale (1: "Strongly disagree"

to 5: “Strongly agree”). To generate the measure of overall uncertainty, the measures of socioeconomic uncertainty and political uncertainty are summated together. Because the questions are about certainty rather than uncertainty, lower score means more uncertainty.

3.3.5 Independent Variable: Professionalism

In this study, I measure the professionalism with two variables – level of education and accreditation. First, Thornton (1970) argues, “Higher level of education is probably the primary basis of their professionalism” (p. 419). In this sense, the professionalism can be measured by the level of education. According to prior studies, organizational level variable can be derived by aggregating individual level variables (e.g., Campion, Medsker, & Higgs, 1993; Shamir, Zakay, Breinin, & Popper, 1998; Gavin & Hofmann, 2002; Croon & van Veldhoven, 2007). Thus, I examine the organization’s level of professionalism by examining the individuals’ education level.

LeRoux and Wright (2010) measure the professionalism by asking executive director’s education level in their survey. Also, Knudsen and Roman (2004) use the percentage of counselors that had completed at least a Master’s degree for measuring the professionalism. Based on these prior studies, I measure the professionalism by asking three questions. First, I ask the percentage of employees with graduate or professional degrees; “What percentage of your employees have graduate or professional degrees (e.g., Master’s degree, Ph.D, MD, JD, or other advanced degrees beyond a Bachelor’s degree)?” These items are measured on a seven point scale (1 = 51% or higher, 2 = 31% – 50%, 3 = 21% – 30%, 4 = 11% – 20%, 5 = 6% – 10%, 6 = 1% – 5%, 7 = 0%). As I mentioned above, I

revised the order of scale because the last category is treated as a reference group in the SPSS ordinal regression. Lower value means more percentage of the employees have graduate or professional degrees. Second, executive director' education level is measured by asking "Does the executive director (or top manager) of your organization have a graduate degree in management such as a Master of Business Administration (MBA), a Master of Public Administration (MPA), or a similar management degree?" with three choices (1 = Yes, 2 = No, Other graduate degree, and 3 = No, Doesn't have a graduate degree).

Lastly, an accreditation has been examined as an indicator of certain level of professionalism in the literature (e.g., Hildebrandt & Eom, 2011; Miller et al., 2010; Mikkelson, 2013; Ullian & Schink, 2012; Yang & Taylor, 2013). For instance, Yang and Taylor (2013) argue that "The process of accreditation creates another level to professionalization" (p. 3). Thus, I use one more survey question for measuring the professionalization: "Is your organization accredited by any external bodies?" This question is measured on a dichotomous scale (Yes: 0 or No: 1). If respondents answer "Yes" in this question, they are asked to write down the name of the accrediting body.

3.3.6 Independent Variable: Organizational Capacity

The organizational capacity is measured by six elements. The first element of organizational capacity is financial capacity. I measure the financial capacity by using Bowman (2011)'s long-term financial capacity because the outcome measurement is related to the long term impacts of organizations rather than short term impacts. Bowman (2011) measures the long term financial capacity with Equity Raito (ER) that "monitors the fraction

of an organization's assets it owns free and clear" (p.40). ER is measured by following equation.

$$ER = (\text{Total Assets} - \text{Total Liabilities}) / \text{Total Assets}$$

Second element of capacity is human resource. It is measured by asking "Our organization has staff with the skills and expertise necessary to track program participants" (Poole et al., 2001). This item is measured on a five point scale (1: "Strongly agree" to 5: "Strongly disagree"). The measure of human resource capacity is differentiated from the idea of education level in the professionalism section. While the education level is measured to examine how the professionalism formed by the education influences the institutional (normative) pressure to adopt the performance measurement, human resource capacity is measured to examine the substantive organizational ability to adopt the performance measurement.

Third element is technical capacity. It is measured by asking "Our organization has adequate infrastructure (e.g., information technology, communication system) for daily operations" (Berman & Wang, 2010). This item is measured on a five point scale (1: "strongly agree" to 5: "strongly disagree").

The fourth element of the organizational capacity is administrative capacity. Using Barman and MacIndoe (2012)'s method, I measure the administrative capacity by asking "Our organization has written rules concerning compensation, governance, job descriptions, conflict of interest, and/or volunteers". This item is measured on a five point scale (1: "strongly agree" to 5: "strongly disagree").

Fifth element of the capacity is political capacity (Berman & Wang, 2000; Christensen & Gazley, 2008; Héritier, 2003). I classify political capacity into political support and political negotiation. First, using the studies of Berman and Wang (2000), Christensen and Gazley (2008), and Héritier (2003), I measure the political support by asking three survey questions: 1) "Top management in our organization supports using data to make decisions"; 2) "Most employees in our organization support using data to make decisions"; 3) "Stakeholders of our organization support using data to make decisions"; These items are measured on a five point scale (1: "strongly agree" to 5: "strongly disagree"). The answers of these items are averaged to get an integrated value on the political support. Second, using Héritier (2003), I measure the political negotiation by asking "Our organization frequently discusses operational issues with stakeholders". This item is measured on a five point scale (1: "strongly agree" to 5: "strongly disagree"). The answers of two aspects of political capacity – political support and political negotiation – are averaged to generate the measure of overall political capacity.

The last element of the capacity is network capacity (Eisinger, 2002; Christensen & Gazley, 2008). Using the studies of Eisinger (2002) and Christensen and Gazley (2008), I measure this element by asking one survey question: "Our organization seeks help from other organizations to improve decision making". This item is measured on a five point scale (1: "strongly agree" to 5: "strongly disagree").

3.3.7 Control Variables: Organizational Age and Size

First, organizational age is measured by asking "In which year was your organization founded?" For instance, American Cancer Society was founded in 1922. Then, its organizational age is 92 (2014 - 1922 = 92). Second, organizational size is measured by total expenses and full time equivalents (FTEs) in the last fiscal year. I get the total expenses for fiscal year 2012 from IRS 990 form. Next, full time equivalents (FTEs) is calculated by the following formula:

FTEs = number of full time employees + ½ * number of part time employees

The number of full time and part time employees is obtained by the survey.

3.4 The Extent of Outcome Information Use on Selected Dimensions

This is measured by asking "To what extent does your nonprofit use outcome information for each of the following purposes?" The survey items include (1) Improve management decisions, (2) Improve strategic planning, (3) Appraise employee performance, (4) Support budget decisions, (5) Report to the board, (6) Improve communication between administrators and elected officials, (7) Respond to public demands for accountability, (8) Respond to pressure from community groups, (9) Comply with government requirements, (10) Report to funders, (11) Get new funding, and (12) For outreach and public relations. Each category is asked on a scale of 5 (1: "Not at all" to 5: "Very great extent"). These categories are chosen in this study because they are represented as the common use of outcome information in the literature (e.g., Carman & Frederick, 2008; Folz, Abdelrazek, & Chung, 2001; Julnes & Holzer, 2001).

3.5 Pilot Study

For the pilot study, I discussed the survey items, question by question, with 6 practitioners - 4 executive directors, 1 associate director, and 1 special projects coordinator - of NC based nonprofits by phone on December 9th, 2013 – January 20th, 2014. Following is the summary of the interview results and the changes I made based on the interviewees' feedback. Overall, interviewees said that the survey questions are clear, straight forward, and easy to understand. One interviewee said "I thought it was a good survey, unlike most I am asked to participate in."

Second, some interviewees pointed out that the question wording (e.g., too much academic jargon) and answer choices. One interviewee said that the wording of textbook and practitioners are different from each other. For instance, the interviewee said that practitioners use a term "grant" rather than "performance based contract". Another interviewee said some respondents might not understand what "performance based contract" means. Following this opinion, I changed the question wording from "Does your organization have any performance base contract" to "Do your grants/contracts require performance information?" Accordingly, I changed the following survey question "To what extent do you performance based contracts with funders focus on outputs or outcomes" to "Which category most closely describes the extent to which your grants/contracts focus on outputs or outcomes?" with a five point scale: 1 = 100% Output and 0% Outcome, 2 = 75% Output and 25% Outcome, 3 = 50% Output and 50% Outcome, 4 = 25% Output and 75% Outcome, and 5 = 0% Output and 100% Outcome. In addition, I included one more choice for survey question asking the use of outcome information: "improve services for clients",

following one interviewee's comment. Lastly, some interviewee mentioned that the wording "organizational innovation and change" is not easy to understand. Thus, I changed the wording, using non-jargon language – "new ways to accomplish our work". Also, I changed the answer scale from 10 point to 5 point to make the survey look simpler.

Third, some interviewees said that a few survey items do not measure the concepts intended correctly. For instance, a survey question for measuring the level of professionalism is the percentage of employees who have graduate or professional degrees. One interviewee said I need to consider that some people are well educated and professionalized, but they do not have professional degree. He/she said that the professionalism of these people is not reflected in my current survey question. The education level, however, has been used as a measure of professionalism by scholars (e.g., LeRoux & Wright, 2010; Knudsen & Roman). Thus, I decided to stay with the original survey item. In addition, I measure the mimetic isomorphism by asking the outcome measurement practice of "largest" organizations. One interviewee said it is ambiguous what "largest" means. He/she said it might mean budget size or the other measures of organizational size. Furthermore, the interviewee said that "largest" do not directly mean the organization is effective and there are a lot of effective small organizations. He/she added that these small nonprofits can be influential as much as the largest organizations can. However, I borrowed the survey item from a published article (Barman & MacIndoe, 2012) which has already proved the survey items' reliability and validity through publication process. Thus, I decided to use the original survey question without changes. Lastly, I measure the organizational size by asking the number of full time and part time employees. One interviewee pointed out that most nonprofits have lots of

volunteers and I should consider the number of volunteers for measuring organizational size. However, FTEs has been used as a measure of organizational size in prior studies (e.g., Thomson, 2010). Thus, I decided to use my original survey item - FTEs which is calculated by the number of full time and part time employees, not volunteers.

Fourth, some interviewees pointed out survey respondents may not be able to answer the questions correctly. One interviewee said survey respondents might guess the answer without correct information for some questions (e.g., percentage of organization's program funding which is subject to mandatory outcomes measurement, percentage of employees with graduate or professional degrees). Another interviewee said that I need to include an answer choice "Don't know" for survey questions. Following this feedback, I included "Don't know" option for some questions in which the choice would be appropriate.

Fifth, some interviewees pointed out the survey question concerned with the examples of the outcome measures and frequency of measurement. For instance, one interviewee said I may need to include more specific outcome indicators. He/she said "quality of life" may need to be elaborated with specific examples. However, I decided to use the original survey item because the purpose of this question is to know whether the respondents collect outcome information related to the "quality of life", not how they measure it. Excessively specifying the survey items and listing too many examples may lengthen the survey and reduce the response rate. Next, an interviewee said that the interviewee's agency measures their program (or service)'s outcome just once. The organization uses follow up survey after they implement the service, but they do not collect the outcome information monthly or quarterly. Following this comment, I revised the

frequency question. First, I ask respondents to answer "Yes" or "No" to each example of outcome measures. Then, in the following question, respondents are given a set of outcomes that they answered "Yes" in the previous question and asked how often they collect the information. I use customization function of Qualtrics in this question, thus, respondents do not have to consider all outcome examples. Also, I changed the answer type from a seven point scale (1 = Never, 2 = Weekly, 3 = Monthly, 4 = Quarterly, 5 = Semiannually, 6 = Annually, and 7 = Biennially or less frequently) to open ended question.

Sixth, one interviewee commented about the third survey question – writing down outcomes that the organization is measuring but not listed in the previous question. The interviewee said that writing down all their outcome indicators is too hard. The interviewee's agency has about 12 outcome indicators but it is too hard to write down all outcomes. Following this comment, I changed the question wording from "Please write down the outcomes that you are measuring but not listed above and the frequency of measurement" to "Please list any outcomes that were not included in the previous question that you consider to be one of your top three outcomes. Please provide the frequency of measurement for any items you list"

Lastly, some interviewees pointed out the difficulties of outcome measurement. One interviewee said that the response rate of their outcome measurement survey is only 1%. Another interviewee said it is difficult conducting surveys and other instruments for outcome measurement especially because many of their clients are in crisis and they do not participate in outcome measurement without incentives. These comments will be elaborated again in the last section - implication and future search - of this study.

3.6 Data Analysis

In this dissertation, I try ordinal regression first because the dependent variable is ordered on one dimension (the adoption of outcome measurement). Ordinal regression supports the use of independent variables which include both categorical factors (e.g., level of risk aversion, measures of organizational capacity, resource providers' mandate of the outcome measurement, mimetic isomorphism) and continuous covariates (e.g., total expenditure, insolvency risk, organizational age, FTEs). However, I try other regression models – poisson regression, binomial regression, and logistic regression - as well in case that the survey data does not meet the assumptions of the ordinal regression. In the result section, I will explain how I ended up with the logistic regression model as the best model to test the hypotheses.

3.7 Summary

To summarize, this study uses survey method to test the hypotheses. 1039 human service nonprofits - 246 foster care organizations (P32), 569 family violence shelters (P43), and 224 family counseling organizations (P46) - are selected from the NCCS Core Data file. A pilot study with six nonprofit practitioners was conducted on December 9th, 2013 – January 20th, 2014 and the survey items were revised based on the results of the pilot study.

CHAPTER 4: RESULTS

4.1 Introduction

In the previous chapter, I explained what survey method I use to test the hypotheses, how the survey sample is chosen, and how the main variables of this study are operationalized. In addition, I addressed the results of a pilot study and how they are reflected to the main survey. In this chapter, I describe the survey results and discuss my research findings.

4.2 Survey response rate

As I mentioned above, the population size of this study is 1,039. I collected the contact information of each organization's executive director (or president) by searching their websites. I collected 775 email addresses. I sent my survey link to these 775 email addresses and 4 out of the 775 emails sent were bounced. Also, I found 57 pop up contact windows; some organizations use a pop up contact window in their websites, which allows website visitors write their questions or comments and send them to appropriate person of the agency. I described the main purpose of my study and sent the survey link through the pop up contact window. Survey responses were collected from March, 5th, 2014 to April, 26th, 2014. I gave a phone call to non-respondents to encourage their survey participation from April, 7th to 18th, 2014. Of the 828 organizations contacted, 317 organizations consented to participate to my survey (38.29%), but 54 out of the 317 organizations did not finish the survey. Finally, 263 of the responses were usable in data analysis. In brief, the response rate of my survey is

31.76%. The response rate of P32 (foster care organizations), P43 (family violence shelters), and P46 (family counseling organizations) are 25.27%, 33.40%, and 34.08% respectively. For the survey sent through the pop up contact window, the response rate is somewhat low (15.79%). If I count only the survey links sent through the email addresses, the response rate is 32.94%. The details of the survey response rate are described in Table 4.1.

Table 4.1: Survey response rate

NTEE	Email survey (Qualtrics)			Pop up contact window		Total number of responses
	Number of email sent	Number of email bounced	Number of responses	Number of pop up windows used	Number of responses	
P32	167	1	45	15	1	46
P43	434	2	148	33	8	156
P46	170	1	61	9	0	61
Sum	771	4	254	57	9	263

4.3 Data quality check

To improve the data quality and the reliability of survey results, I examined the cases of respondents who answered "5" in the survey question for the dependent variable, that is, 76 – 100% of their organization's programs/services was subject to outcome measurement in the last fiscal year. 120 out of the 263 respondents answered "5" in the survey question for dependent variable. As I mentioned in the methodology section, I listed specific outcome indicators - 5 indicators for P32, 5 indicators for P43, and 8 indicators for P46 - and asked how often the respondents track their outcomes. I assumed that the 120 respondents who

answered “5” in the dependent variable survey question would check at least one of the items I listed or have added at least one additional item. I found that 1 out of the 120 respondents did not check any outcome indicators I listed or provide the additional outcome indicators.

In addition, I checked whether there are respondents who neither checked all outcome indicators I listed in the survey nor provided any other additional outcome measures in the open ended question. I assumed the outcome measurement of these respondents’ organizations is 0%, that is, 0% of these organization’s programs/services was subject to outcome measurement in the last fiscal year. If the respondents which belong to this case answer the other choices in the survey question for the percentage of outcome measurement, I regarded them as dishonest responses. I found 3 organizations of this case and 1 out of the 3 organizations belonged to the case above that neither check any outcome indicators I listed nor provide the additional outcome indicators. In brief, total 3 responses failed to pass the data quality examination. I decided to drop these 3 responses from my data analysis.

Lastly, 1 respondent did not answer the survey question for the percentage of outcome measurement although they checked the examples of outcome indicators and answered the other survey questions. I decided to exclude this case from my analysis because when comparing the overall percentage of outcome measurement - the main question for the dependent variable - with the count of outcome indicators, I should compare the same group of organizations. Since they did not answer the first question, this is more than measurement error and I chose to exclude this case.

4.4 Representatives of respondent group

Before I start to analyze the data, I have checked the representativeness of respondent group in terms of total expenses and region, compared to the census of organizations I sent the survey links to.

4.4.1 Total expenses

First, one sample t-test examines how closely the mean of total expenses in the respondent group reflects the true mean in the population from which this group (respondents) was drawn. Table 4.2 indicates the descriptive statistics of total expenses of respondents and population group.

Table 4.2: Total expenses of respondents and population group

	Respondent group (n = 259)			Population group (n = 1039)		
NTEE	Mean	Std. Deviation	Median	Mean	Std. Deviation	Median
P32	5,628,972	9,240,969	2,610,428	5,546,129	11,058,934	2,230,346
P43	1,801,505	1,761,159	1,407,851	1,520,326	1,365,224	1,078,989
P46	2,733,555	3,983,736	1,175,844	2,302,833	3,280,467	1,129,914
Total	2,682,428	4,691,514	1,443,297	2,642,202	5,908,616	1,231,704

As shown in Table 4.3, the 95% confidence interval of the difference of mean between respondents and population group contains zero, which means that the respondent

group's average total expenses are not significantly different from the population group's average total expenses with 95% confidence level.

Table 4.3: One sample t-test of all groups

	Test Value = 2,642,202					
					95% Confidence Interval of the Difference	
	t	Df	Sig. (2-tailed)	Mean Difference	Lower	Upper
TotExp2011	.14	258	.89	40,226.16	-533,828.57	614,280.90

As shown in Table 4.4, the 95% confidence interval of the difference of mean between P32 respondents and population group contains zero, which means that the P32 respondent group's average total expenses are not significantly different from population group's average total expenses with 95% confidence level.

Table 4.4: One sample t-test, P32

	Test Value = 5,546,129					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
TotExp2011	.06	44	.95	82,843.29	-2,693,451.21	2,859,137.79

Table 4.5 shows that the 95% confidence interval of the difference of mean between P43 respondents and population group does not contains zero, which means that P43 respondents group's average total expenses are significantly different from population group's average total expenses with 95% confidence level. It indicates that the results may not be generalizable to all P43 organizations with greater than $500,000 in total expenses. This may also indicate that I should control for NTEE group in my analysis. I will discuss this again in the limitations section of this study.

Table 4.5: One sample t-test, P43

	Test Value = 1,520,326					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
TotExp2011	1.98	153	.05	281,179.43	807.08	561,551.78

As shown in Table 4.6, the 95% confidence interval of the difference of mean between P46 respondents and population group contains zero, which means that P46 respondent group's average total expenses are not significantly different from population group's average total expenses with 95% confidence level.

Table 4.6: One sample t-test, P46

	Test Value = 2,302,833					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
TotExp2011	.84	59	.41	430,722.08	-598,386.15	1,459,830.32

4.4.2 Regional distribution

In this section, I examine whether there is a significant difference in regional distribution between respondents and population group. As shown in Table 4.7 and 4.8, the Chi-Square test of all groups (P32, P43, and P46) indicates an insignificant result ($p = .618$), which means there is no statistically significant difference in the regional distribution between population and respondents group.

Table 4.7: Cross tabulation, Region, Population and Respondents

			Respondents or Population		
			Respondents	Population	Total
Region	Midwest	Count	72	260	332
		% within Pop	27.8%	25.0%	25.6%
	Northeast	Count	40	189	229
		% within Pop	15.4%	18.2%	17.6%
	Puerto Rico	Count	0	4	4
		% within Pop	0.0%	0.4%	0.3%
	South	Count	82	317	399
		% within Pop	31.7%	30.5%	30.7%
	West	Count	65	269	334
		% within Pop	25.1%	25.9%	25.7%
Total		Count	259	1039	1298
		% within Pop	100.0%	100.0%	100.0%

Table 4.8: Chi-Square test, All Groups

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2.648[a]	4	.618
Likelihood Ratio	3.448	4	.486
N of Valid Cases	1298		

a. 2 cells (20.0%) have expected count less than 5. The minimum expected count is .80.

As shown in Table 4.9 and 4.10, the Chi-Square test of P32 group shows an insignificant result (p = .266), which means there is no significant difference in the regional distribution between P32 population and respondents group.

Table 4.9: Cross tabulation, Region, Population and Respondents (P32)

			Respondents or Population		
			Respondents	Population	Total
Region	Midwest	Count	17	60	77
		% within Pop	37.8%	24.4%	26.5%
	Northeast	Count	4	35	39
		% within Pop	8.9%	14.2%	13.4%
	South	Count	9	50	59
		% within Pop	20.0%	20.3%	20.3%
	West	Count	15	101	116
		% within Pop	33.3%	41.1%	39.9%
Total		Count	45	246	291
		% within Pop	100.0%	100.0%	100.0%

Table 4.10: Chi-Square test, P32

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.957[a]	3	.266
Likelihood Ratio	3.829	3	.281
N of Valid Cases	291		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.03.

As shown in Table 4.11 and 4.12, the Chi-Square test of P43 group shows an insignificant result (p = .498), which means there is no significant difference in the regional distribution between P43 population and respondents group.

Table 4.11: Cross tabulation, Region, Population and Respondents (P43)

			Respondents or Population		
			Respondents	Population	Total
Region	Midwest	Count	40	136	176
		% within Pop	26.0%	23.9%	24.3%
	Northeast	Count	21	103	124
		% within Pop	13.6%	18.1%	17.2%
	Puerto Rico	Count	0	3	3
		% within Pop	0.0%	0.5%	0.4%
	South	Count	52	199	251
		% within Pop	33.8%	35.0%	34.7%
	West	Count	41	128	169
		% within Pop	26.6%	22.5%	23.4%
Total		Count	154	569	723
		% within Pop	100.0%	100.0%	100.0%

Table 4.12: Chi-Square test, P43

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.369[a]	4	.498
Likelihood Ratio	4.043	4	.400
N of Valid Cases	723		

a. 2 cells (20.0%) have expected count less than 5. The minimum expected count is .64.

As shown in Table 4.13 and 4.14, the Chi-Square test of P46 group shows an insignificant result (p = .889), which means there is no significant difference in the regional distribution between population and respondents group.

Table 4.13: Cross tabulation, Region, Population and Respondents (P46)

			Respondents or Population		
			Respondents	Population	Total
Region	Midwest	Count	15	64	79
		% within Pop	25.0%	28.6%	27.8%
	Northeast	Count	15	51	66
		% within Pop	25.0%	22.8%	23.2%
	Puerto Rico	Count	0	1	1
		% within Pop	0.0%	0.4%	0.4%
	South	Count	21	68	89
		% within Pop	35.0%	30.4%	31.3%
	West	Count	9	40	49
		% within Pop	15.0%	17.9%	17.3%
Total		Count	60	224	284
		% within Pop	100.0%	100.0%	100.0%

Table 4.14: Chi-Square test, P46

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.136[a]	4	.889
Likelihood Ratio	1.345	4	.854
N of Valid Cases	284		

a. 2 cells (20.0%) have expected count less than 5. The minimum expected count is .21.

In this section, I examined whether there is a statistically significant difference between respondents and population group, using total expenses and regional distribution. In general, the survey respondent group is fairly representative of the population group with an exception that the P43 respondents' average total expenses are slightly larger than P43 population group's average total expenses.

4.5 Data analysis

4.5.1 Descriptive statistics

Dependent Variables

Table 4.15 shows the responses of the extent of outcome measurement by NTEE code. Overall, 46% of respondents answered that 76% – 100% of their programs and services was subject to outcome measurement in the last fiscal year and 70% of respondents answered that they measured the outcome of more than 50% of their programs and services. Especially, for P46 (Family counseling organizations), 56% of respondents answered that 76% - 100% of their programs/services was subject to outcome measurement in FY2013.

Table 4.15: Extent of outcome measurement

Survey question: In the last fiscal year, what percentage of your organization's programs/services was subject to outcome measurement?				
	P32	P43	P46	Total
0%	4% (2)	2% (3)	5% (3)	3% (8)
1% - 25%	13% (6)	10% (15)	10% (6)	11% (27)
26% - 50%	11% (5)	19% (29)	14% (8)	16% (42)
51% - 75%	20% (9)	28% (43)	15% (9)	24% (61)
76% - 100%	51% (23)	41% (63)	56% (33)	46% (119)
Total	100% (45)	100% (153)	100% (60)	100% (257)

Note: "n" is reported in parentheses

As mentioned in the method section, I listed specific outcome indicators - 5 indicators for P32, 5 indicators for P43, and 8 indicators for P46 in the survey - and checked whether the organizations use each outcome indicator. As shown in Table 4.16, only 51% of P32 respondents answered that they measure the outcome measure "Reduced childhood behavioral and emotional disorders", while 84% of respondents answered that they measure the outcome indicator "Increased placement stability"

Table 4.16: Outcome indictors of P32 (Foster care organization)

Outcome Indicator	Yes
Increased placement stability	84% (38)
Client satisfaction with respect to foster care service	78% (35)
Reduced childhood behavioral and emotional disorders	51% (23)
Increased quality of life	60% (27)
Enhanced level of education or life skills of foster child	73% (33)

Note: "n" is reported in parentheses

Table 4.17 indicates that only 33% of P43 respondents answered that they measure the outcome measure "Reduced childhood behavioral and emotional disorders", while 92% of the respondents answered that they measure the outcome indicator "Client satisfaction with respect to the shelter's service"

Table 4.17: Outcome indicator of P43 (Family violence shelters)

Outcome Indicator	Yes
Client satisfaction with respect to the shelter's service	92% (140)
Reduced physical/mental illness (e.g., depression and other affective disorders)	40% (61)
Increased quality of life	77% (117)
Increased self-efficacy	81% (122)
Reduced childhood behavioral and emotional disorders	33% (49)

Note: "n" is reported in parentheses

As shown in Table 4.16 and 4.17, the outcome indicator "Reduced childhood behavioral and emotional disorders" was measured by only 51% of P32 organizations and 33% of P43 organizations. However, Table 4.18 shows that 72% of P46 respondents use this indicator. In addition, only 48% of P46 respondents answered that they measure the outcome measure "Reduced child abuse and neglect", while 88% of the respondents answered that they measure the outcome indicator "Client satisfaction with respect to the shelter's service".

Table 4.18: Outcome indicator of P46 (Family counseling organizations)

Outcome Indicator	Yes
Client satisfaction with respect to family counseling service	88% (51)
Increased self-efficacy	71% (39)
Reduced physical/mental illnesses (e.g., depression and other affective disorders)	77% (43)
Increased quality of life	75% (42)
Reduced drug and alcohol use	49% (27)
Reduced childhood behavioral and emotional disorders	72% (41)
Reduced marital and relationship problems	67% (36)
Reduced child abuse and neglect	48% (27)

Note: "n" is reported in parentheses

Table 4.19 shows how many outcome indicators each organization uses by NTEE code. Table 4.19 indicates that a large percentage of respondents answered that they use more than half of the outcome indicators listed in the survey. As shown in Table 4.16, 4.17, and 4.18, most of the outcome measures are checked by more than 50% of the respondents. In addition, according to Table 4.15, 70% of respondents answered that more than 50% of their programs and services was subject to outcome measurement. Thus, to integrate the categories with low frequencies, I recoded each count variable (number of outcome measures) into a dichotomous variable. I counted how many outcome measures each respondent checked (See Table 4.19), then recoded 50% or more of the listed outcome measures "1" and less than 50%

"0". Table 4.20 shows that 74% of respondents checked more than 50% of the outcome measures I presented and 26% of respondents checked less than 50%.

Table 4.19: Number of outcome measures (Count data)

	P32	P43	P46
0	0% (0)	6% (9)	2% (1)
1	13% (6)	7% (10)	8% (5)
2	9% (4)	12% (18)	8% (5)
3	20% (9)	36% (56)	5% (3)
4	33% (15)	20% (30)	10% (6)
5	24% (11)	20% (31)	18% (11)
6	.	.	23% (14)
7	.	.	2% (1)
8	.	.	23% (14)
Total	100% (45)	100% (154)	100% (60)

Note: "n" is reported in parentheses

Table 4.20: Number of outcome measures (Dichotomous)

	P32	P43	P46	Total
Less than 50%	22% (10)	24% (37)	33% (20)	26% (67)
More than 50%	78% (35)	76% (117)	67% (40)	74% (192)
Total	100% (45)	100% (154)	100% (60)	100% (259)

Note: "n" is reported in parentheses

Risk Aversion

As shown in Table 4.21, 4.22, and 4.23, risk aversion is measured by three survey items. Since the items are designed to measure risk taking attitude, the results should be

interpreted reversely. That is to say, responses "Strongly agree" or "agree" mean that the organization is less risk averse, compared to the cases of the other responses. For the first survey item, overall 46% of respondents agreed or strongly with that there is a reward/incentive system in place in the organization that encourages risk taking. Especially, 64% of P32 (Foster care) organizations agreed or strongly agreed with the first survey item. Second, 96% of the respondents agreed or strongly with that management is willing to implement new ways to accomplish our work. Lastly, 74% of respondents agreed or strongly agreed that non-management employees willingly adopt new ways to accomplish our work.

Table 4.21: Risk aversion, survey item1

Survey item1: There is a reward/incentive system in place in the organization that encourages risk taking (using innovative ideas with the goal of improving performance).				
	P32	P43	P46	Total
Strongly Disagree	9% (4)	7% (10)	9% (5)	8% (19)
Disagree	11% (5)	26% (39)	14% (8)	21% (52)
Neutral	16% (7)	25% (37)	38% (22)	26% (66)
Agree	39% (17)	31% (46)	28% (16)	31% (79)
Strongly Agree	25% (11)	13% (19)	12% (7)	15% (37)
Total	100% (44)	100% (151)	100% (58)	100% (253)

Note: "n" is reported in parentheses

Table 4.22: Risk aversion, survey item2

Survey item2: Management is willing to implement new ways to accomplish our work.				
	P32	P43	P46	Total
Strongly Disagree	2% (1)	1% (1)	0% (0)	1% (2)
Disagree	0% (0)	2% (3)	0% (0)	1% (3)
Neutral	2% (1)	3% (4)	2% (1)	2% (6)
Agree	27% (12)	25% (38)	36% (21)	28% (71)
Strongly Agree	69% (31)	69% (104)	63% (37)	68% (172)
Total	100% (45)	100% (150)	100% (59)	100% (254)

Note: "n" is reported in parentheses

Table 4.23: Risk aversion, survey item3

Survey item3: Non-management employees willingly adopt new ways to accomplish our work.				
	P32	P43	P46	Total
Strongly Disagree	0% (0)	3% (4)	2% (1)	2% (5)
Disagree	7% (3)	11% (17)	5% (3)	9% (23)
Neutral	22% (10)	11% (17)	17% (10)	15% (37)
Agree	47% (21)	56% (84)	57% (33)	54% (138)
Strongly Agree	24% (11)	19% (29)	19% (11)	20% (51)
Total	100% (45)	100% (152)	100% (59)	100% (254)

Note: "n" is reported in parentheses

To generate an overall scale of risk aversion, I averaged the responses of three items of risk aversion. In this process, I reversed the scores of the three items because the original survey items are designed to measure risk taking. Thus, higher values on the overall scale of

risk aversion mean more risk aversion. Cronbach's alpha was obtained to assess the internal reliability of the three items of risk aversion. According to Garson (2012a), Cronbach's alpha should be .60 or higher for the "lenient" scale and .70 or higher for the "adequate" scale. As shown in Table 4.24, the Cronbach's alpha of the three items of risk aversion is .582, which is slightly lower than .60. Before I dropped out the 4 cases which failed to pass my data quality examination (See section 4.3), the Cronbach's alpha of risk aversion was .608 in the raw dataset. Deleting the 4 cases slightly reduced the Cronbach's alpha statistics of risk aversion. Deleting any item among the three did not increase the Cronbach's alpha statistics.

Table 4.24: Reliability Statistics of Risk Aversion

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.582	.608	3

While Cronbach's alpha is the most common measure of reliability, there is a similar measure of reliability developed Armor (1974) - Armor's reliability Theta (Garson, 2012a). To confirm the reliability of three items of risk aversion, the Armor's reliability Theta also was obtained. It is calculated from the results of principal components analysis (PCA) as follows:

$$\text{Armor's reliability Theta} = \theta = [p/(p-1)]*[1-(1/\lambda_1)]$$

Where p = the number of items in the scale and where λ1 denotes the first (therefore largest) eigenvalue from the PCA of the items comprising the scale (Garson, 2012a).

When I conducted the PCA, the largest eigenvalue was 1.682. Using the PCA result, the Armor's reliability theta of risk aversion was calculated to 0.61. Since Armor's theta is interpreted similar to Cronbach's alpha (Garson, 2012a), 0.61 indicates an acceptable level. In this sense, even though the Cronbach's alpha of risk version was slightly lower than 0.60, I decided to keep the risk aversion scale in my analysis model.

Performance Information

Next, as shown in Table 4.25, 93% of respondents answered that their grants and contracts require performance information. Table 4.26 shows that respondents' grants and contracts slightly focused more on output rather than outcome. 40% of respondents answered that their grants and contacts use 75-100% output and 0-25% outcome, while only 21% of respondents answered that their grants and contacts use 75-100% outcome and 0-25% output.

Table 4.25: Performance information, survey item1

Survey question: Do your grants/contracts require performance information?				
	P32	P43	P46	Total
Yes	84% (37)	95% (144)	94% (51)	93% (232)
No	16% (7)	5% (7)	6% (3)	7% (17)
Total	100% (44)	100% (151)	100% (54)	100% (249)

Note: "n" is reported in parentheses

Table 4.26: Performance information, survey item2

Survey question: Which category most closely describes the extent to which your grants/contracts focus on outputs or outcomes?				
	P32	P43	P46	Total
100% Output and 0% Outcome	11% (4)	6% (8)	6% (3)	7% (15)
75% Output and 25% Outcome	17% (6)	41% (57)	20% (10)	33% (73)
50% Output and 50% Outcome	42% (15)	35% (48)	55% (27)	40% (90)
25% Output and 75% Outcome	25% (9)	15% (21)	16% (8)	17% (39)
0% Output and 100% Outcome	6% (2)	4% (5)	2% (1)	4% (8)
Total	100% (36)	100% (139)	100% (49)	100% (224)

Note: "n" is reported in parentheses

Coercive Isomorphism

According to Table 4.27, 54% of respondents answered that more than 50% of their program funding is subject to mandatory outcome measurement. On the other hand, 22% of P32 respondents answered that their organizations' program funding is not subject to mandatory outcome measurement. In addition, Table 4.28 shows that only 21% of P32 (Foster Care) respondents answered that they receive funding from the United Way, while 77% of P43 respondents answered that they receive it.

Table 4.27: Coercive isomorphism, survey item1

Survey question: What percentage of your organization's program funding is subject to mandatory outcomes measurement?				
	P32	P43	P46	Total
0%	22% (9)	8% (11)	11% (6)	11% (26)
1% - 25%	15% (6)	15% (21)	22% (12)	16% (39)
26% - 50%	15% (6)	21% (30)	17% (9)	19% (45)
51% - 75%	20% (8)	32% (46)	15% (8)	26% (62)
76% - 100%	29% (12)	26% (37)	35% (19)	28% (68)
Total	100% (41)	100% (145)	100% (54)	100% (240)

Note: "n" is reported in parentheses

Table 4.28: Coercive isomorphism, survey item2

Survey question: Does your organization receive funding from the United Way?				
	P32	P43	P46	Total
Yes	21% (9)	77% (112)	57% (33)	62% (154)
No	80% (35)	22% (32)	42% (24)	37% (91)
Total	100% (44)	100% (144)	100% (57)	100% (245)

Note: "n" is reported in parentheses

Mimetic Isomorphism

Table 4.29 and 4.30 show the responses of survey items for mimetic isomorphism. About 80% of respondents agreed or strongly agreed that the largest organizations in their field measure the outcomes and they improve the performance by measuring the outcome.

Table 4.29: Mimetic isomorphism, survey item1

Survey item: The largest organizations in our field measure their outcomes				
	P32	P43	P46	Total
Strongly Disagree	5% (2)	1% (2)	0% (0)	2% (4)
Disagree	7% (3)	8% (11)	12% (6)	9% (20)
Neutral	7% (3)	11% (15)	16% (8)	11% (26)
Agree	56% (23)	51% (73)	38% (19)	49% (115)
Strongly Agree	24% (10)	29% (41)	34% (17)	29% (68)
Total	100% (41)	100% (142)	100% (50)	100% (233)

Note: "n" is reported in parentheses

Table 4.30: Mimetic isomorphism, survey item2

Survey item: Nonprofits in our field that measure their outcomes improve their performance.				
	P32	P43	P46	Total
Strongly Disagree	2% (1)	0% (0)	0% (0)	0% (1)
Disagree	2% (1)	7% (10)	4% (2)	6% (13)
Neutral	14% (6)	14% (20)	14% (7)	14% (33)
Agree	48% (20)	52% (72)	48% (24)	50% (116)
Strongly Agree	33% (14)	27% (37)	34% (17)	29% (68)
Total	100% (42)	100% (139)	100% (50)	100% (231)

Note: "n" is reported in parentheses

To make an overall scale of mimetic isomorphism, I checked the reliability of two items of mimetic isomorphism. As shown in Table 4.31, however, the Cronbach's alpha of mimetic isomorphism is only .386, which is too low. Thus, I decided to treat the two items of mimetic isomorphism separately without integrating them.

Table 4.31: Reliability Statistics of Mimetic Isomorphism

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.386	.389	2

Uncertainty

Table 4.32 and 4.33 show that responses of the survey items for measuring uncertainty are evenly distributed. About same percentage of respondents agreed or disagreed that the changes in their community's socioeconomic/political environment are predictable. It is notable that only 25% of P46 respondents answered that changes in their community's socioeconomic status are predictable, while about 40% of P32 and P43 respondents answered that. For the changes in their community's political environment, only 21% of P46 respondents answered that the changes in their community are predictable, while 37% of P32 and 38% of P43 respondents answered that.

Table 4.32: Uncertainty, survey item1

Survey item: Changes in our community's socioeconomic status are predictable				
	P32	P43	P46	Total
Strongly Disagree	5% (2)	2% (3)	4% (2)	3% (7)
Disagree	31% (13)	30% (41)	41% (23)	33% (77)
Neutral	24% (10)	29% (39)	30% (17)	28% (66)
Agree	38% (16)	37% (51)	20% (11)	33% (78)
Strongly Agree	2% (1)	2% (3)	5% (3)	3% (7)
Total	100% (42)	100% (137)	100% (56)	100% (235)

Note: "n" is reported in parentheses

Table 4.33: Uncertainty, survey item2

Survey item: Changes in our community's political environment are predictable				
	P32	P43	P46	Total
Strongly Disagree	7% (3)	2% (2)	7% (4)	4% (9)
Disagree	27% (11)	31% (42)	39% (22)	32% (75)
Neutral	29% (12)	29% (39)	32% (18)	30% (69)
Agree	27% (11)	34% (46)	14% (8)	28% (65)
Strongly Agree	10% (4)	4% (6)	7% (4)	6% (14)
Total	100% (41)	100% (135)	100% (56)	100% (232)

Note: "n" is reported in parentheses

As shown in Table 4.34, the Cronbach's alpha of two items of uncertainty is adequate level (Cronbach's Alpha = .766). Thus, I averaged the two items of uncertainty to generate overall scale of uncertainty.

Table 4.34: Reliability Statistics of Uncertainty

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.766	.767	2

Professionalism

According to Table 4.35, 66% of P46 (family counseling organization) respondents answered that more than 50% of their employees have graduate or professional degrees. On the other hand, for P43 family violence shelters, only 17% of respondents answered that more than 50% of their employees have graduate or professional degrees. For the executive

director (or top manager)'s education level, Table 4.36 shows that 90% of P46 respondents answered that their executive directors or top managers have a graduate degree. Overall, 80% of respondents answered that their executive directors or top managers have graduate degree. Lastly, according to Table 4.37, 46% of respondents answered that their organizations have accreditation. Especially, for P46 (family counseling organizations), 60% of respondents answered that their organizations are accredited by external bodies.

Table 4.35: Professionalism, survey item1

Survey question: What percentage of your employees have graduate or professional degrees (e.g., Master's degree, Ph.D, MD, JD, or other advanced degrees beyond a Bachelor's degree)?				
	P32	P43	P46	Total
0%	5% (2)	4% (6)	0% (0)	3% (8)
1 – 5%	11% (5)	27% (39)	2% (1)	18% (45)
6 – 10%	9% (4)	17% (25)	3% (2)	12% (31)
11 – 20%	9% (4)	12% (18)	0% (0)	9% (22)
21 – 30%	14% (6)	9% (13)	10% (6)	10% (25)
31 – 50%	11% (5)	14% (21)	19% (11)	15% (37)
51% or higher	41% (18)	17% (25)	66% (38)	33% (81)
Total	100% (44)	100% (147)	100% (58)	100% (249)

Note: "n" is reported in parentheses

Table 4.36: Professionalism, survey item2

Survey question: Does the executive director (or top manager) of your organization have a graduate degree in management such as a Master of Business Administration (MBA), a Master of Public Administration (MPA), or a similar management degree?				
	P32	P43	P46	Total
Yes	34% (15)	34% (50)	32% (19)	34% (84)
No, Other graduate degree	50% (22)	40% (59)	58% (34)	46% (115)
No, Doesn't have a graduate degree	16% (7)	26% (38)	10% (6)	20% (51)
Total	100% (44)	100% (147)	100% (59)	100% (250)

Note: "n" is reported in parentheses

Table 4.37: Professionalism, survey item3

Survey question: Is your organization accredited by any external bodies?				
	P32	P43	P46	Total
Yes	49% (20)	40% (57)	60% (34)	46% (111)
No	51% (21)	60% (85)	40% (23)	54% (129)
Total	100% (41)	100% (142)	100% (57)	100% (240)

Note: "n" is reported in parentheses

Organizational Capacity

As described in the method section, financial capacity is measured by Equity Ratio (ER) of Bowman (2011). An ER of 1.0 means that the organization does not have debt. On the contrary, an ER of zero indicates that the organization's net assets are zero, that is, the organization's total assets are entirely borrowed one. Lastly, negative ER means the total liabilities – financial obligations – are larger than the total assets of the organization

(Bowman, 2011). Table 4.38 shows the mean of ER is 0.73 and P43 (Family violence shelters) has strongest financial capacity among the three groups.

Table 4.38: Financial capacity (ER), FY2011

NTEE	P32	P43	P46	Total
Mean	0.51	0.81	0.68	0.73
Std. Deviation	0.73	0.20	0.46	0.42
Median	0.75	0.88	0.84	0.84
Minimum	-3.25	0.07	-1.96	-3.25
Maximum	1.00	1.00	1.00	1.00
Range	4.25	0.93	2.96	4.25

For the human resource capacity, as shown in Table 4.39, 88% of respondents agreed or strongly agreed that their organizations have staff with the skills and expertise necessary to track program participants.

Table 4.39: Human Resource Capacity

Survey item: Our organization has staff with the skills and expertise necessary to track program participants				
	P32	P43	P46	Total
Strongly Disagree	2% (1)	0% (0)	0% (0)	0% (1)
Disagree	2% (1)	9% (13)	3% (2)	6% (16)
Neutral	4% (2)	5% (8)	7% (4)	6% (14)
Agree	42% (19)	43% (64)	49% (29)	44% (112)
Strongly Agree	49% (22)	43% (64)	41% (24)	44% (110)
Total	100% (45)	100% (149)	100% (59)	100% (253)

Note: "n" is reported in parentheses

Table 4.40 indicates that 73% of respondents agreed or strongly agreed with that their organizations have adequate infrastructure (e.g., information technology, communication system) for daily operations.

Table 4.40: Infrastructure Capacity

Survey item: Our organization has adequate infrastructure (e.g., information technology, communication system) for daily operations				
	P32	P43	P46	Total
Strongly Disagree	2% (1)	1% (1)	2% (1)	1% (3)
Disagree	11% (5)	21% (32)	13% (8)	18% (45)
Neutral	0% (0)	9% (14)	12% (7)	8% (21)
Agree	38% (17)	48% (73)	47% (28)	46% (118)
Strongly Agree	49% (22)	21% (32)	27% (16)	27% (70)
Total	100% (45)	100% (152)	100% (60)	100% (257)

Note: "n" is reported in parentheses

For administrative capacity, Table 4.41 shows that 98% of respondents agreed or strongly agreed with that their organizations have written rules concerning compensation, governance, job descriptions, conflict of interest, and/or volunteers.

Table 4.41: Administrative Capacity

Survey item: Our organization has written rules concerning compensation, governance, job descriptions, conflict of interest, and/or volunteers				
	P32	P43	P46	Total
Strongly Disagree	0% (0)	0% (0)	2% (1)	0% (1)
Disagree	0% (0)	0% (0)	0% (0)	0% (0)
Neutral	0% (0)	1% (1)	5% (3)	2% (4)
Agree	29% (13)	29% (45)	25% (15)	28% (73)
Strongly Agree	71% (32)	70% (107)	68% (41)	70% (180)
Total	100% (45)	100% (153)	100% (60)	100% (258)

Note: "n" is reported in parentheses

According to Table 4.42 and 4.43, 73% of respondents agreed or strongly agreed with that most employees in their organizations support using data to make decisions, while 96% of respondents agreed or strongly agreed with that top management in their organizations do that. In addition, Table 4.44 shows that 91% of respondents agreed or strongly agreed with that stakeholders of their organizations support using data to make decisions.

Table 4.42: Political support, survey item1

Survey item: Top management in our organization supports using data to make decisions				
	P32	P43	P46	Total
Strongly Disagree	0% (0)	0% (0)	0% (0)	0% (0)
Disagree	0% (0)	1% (2)	3% (2)	2% (4)
Neutral	0% (0)	4% (6)	0% (0)	2% (6)
Agree	34% (15)	38% (58)	32% (19)	36% (92)
Strongly Agree	66% (29)	57% (86)	64% (38)	60% (153)
Total	100% (44)	100% (152)	100% (59)	100% (255)

Note: "n" is reported in parentheses

Table 4.43: Political support, survey item2

Survey item: Most employees in our organization support using data to make decisions				
	P32	P43	P46	Total
Strongly Disagree	0% (0)	0% (0)	0% (0)	0% (0)
Disagree	9% (4)	7% (11)	4% (2)	7% (17)
Neutral	13% (6)	22% (33)	25% (14)	21% (53)
Agree	33% (15)	49% (74)	46% (26)	46% (115)
Strongly Agree	44% (20)	22% (33)	26% (15)	27% (68)
Total	100% (45)	100% (151)	100% (57)	100% (253)

Note: "n" is reported in parentheses

Table 4.44: Political support, survey item3

Survey item: Stakeholders of our organization support using data to make decisions				
	P32	P43	P46	Total
Strongly Disagree	0% (0)	0% (0)	0% (0)	0% (0)
Disagree	0% (0)	1% (2)	3% (2)	2% (4)
Neutral	9% (4)	9% (13)	3% (2)	8% (19)
Agree	51% (22)	51% (74)	45% (26)	50% (122)
Strongly Agree	40% (17)	38% (55)	49% (28)	41% (100)
Total	100% (43)	100% (145)	100% (59)	100% (247)

Note: "n" is reported in parentheses

As shown in Table 4.45, the Cronbach's alpha of three items of the political support is .709, which is an adequate level. Thus, I averaged the responses of three items to generate the overall scale of political support.

Table 4.45: Reliability Statistics of Political Support

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.709	.721	3

Table 4.46 shows that 69% of respondents agreed or strongly agreed with that their organizations frequently discuss operational issues with stakeholders.

Table 4.46: Political Negotiation

Survey item: Our organization frequently discusses operational issues with stakeholders				
	P32	P43	P46	Total
Strongly Disagree	0% (0)	1% (1)	0% (0)	0% (1)
Disagree	2% (1)	16% (23)	7% (4)	11% (28)
Neutral	12% (5)	23% (34)	16% (9)	20% (48)
Agree	35% (15)	40% (59)	34% (19)	38% (93)
Strongly Agree	51% (22)	20% (30)	43% (24)	31% (76)
Total	100% (43)	100% (147)	100% (56)	100% (246)

Note: "n" is reported in parentheses

As shown in Table 4.47, the Cronbach's alpha of two items of political capacity is .529, which is not an acceptable level. Thus, I decided not to integrate two concepts of political capacity: political support and political negotiation.

Table 4.47: Reliability Statistics of Political Capacity

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.529	.586	2

Lastly, for the network capacity, Table 4.38 indicates that 78% of respondents agreed with strongly agreed with their organization seeks help from other organizations to improve decision making.

Table 4.48: Network Capacity

Survey item: Our organization seeks help from other organizations to improve decision making				
	P32	P43	P46	Total
Strongly Disagree	0% (0)	0% (0)	0% (0)	0% (0)
Disagree	7% (3)	5% (7)	3% (2)	5% (12)
Neutral	11% (5)	20% (30)	17% (10)	18% (45)
Agree	48% (21)	53% (79)	43% (25)	50% (125)
Strongly Agree	34% (15)	23% (34)	36% (21)	28% (70)
Total	100% (44)	100% (150)	100% (58)	100% (252)

Note: "n" is reported in parentheses

Control Variables

Table 4.49 shows that the mean and median of organizational age are 37 and 35 years respectively.

Table 4.49: Organizational Age

NTEE	P32	P43	P46	Total
Mean	32	33	52	37
Std. Deviation	23	10	39	24
Median	26	34	41	35
Minimum	8	8	9	8
Maximum	114	115	201	201
Range	106	107	192	193

Next, for the organizational size, FTE and total expenses were measured. Table 4.50 indicates the mean and median of FTE of respondents' organizations are 36 and 22 respectively.

Table 4.50: FTE

NTEE	P32	P43	P46	Total
Mean	49	29	43	36
Std. Deviation	56	26	55	41
Median	26	21	21	22
Minimum	3	6	3	3
Maximum	220	155	301	301
Range	218	150	298	299

Table 4.51 shows that the mean and median of total expenses of respondents' organizations are 2,682,428 and 1,443,297 respectively.

Table 4.51: Total expenses

NTEE	P32	P43	P46	Total
Mean	5,628,972	1,801,505	2,733,555	2,682,428
Std. Deviation	9,240,969	1,761,159	3,983,736	4,691,514
Median	2,610,428	1,407,851	1,175,844	1,443,297
Minimum	545,934	508,178	514,391	508,178
Maximum	46,152,548	13,153,788	22,484,988	46,152,548
Range	45,606,614	12,645,610	21,970,597	45,644,370

Table 4.52 shows that 85% of respondents answered that they have ever received any training program on outcome measurement.

Table 4.52: Outcome Measurement Training

Survey question: Have you ever received any training program on outcome measurement?				
	P32	P43	P46	Total
Yes	84% (38)	85% (128)	85% (51)	85% (217)
No	16% (7)	15% (23)	15% (9)	15% (39)
Total	100% (45)	100% (151)	100% (60)	100% (256)

Note: "n" is reported in parentheses

4.5.2 Outcome measurement rate by time

Table 4.53 indicates the trend of adoption of outcome measurement by period. Before 1980, the adopt rate of outcome measurement was only 6.43%. However, it is increased to 77.99% after 2011.

Table 4.53: The Adoption of Outcome Measurement by Period

Period	Before 1980	1981 - 1990	1991 - 2000	2001 - 2010	After 2011
Number of organizations measuring outcomes	9	28	75	195	202
Total number of organizations in the period	140	206	241	259	259
Percentage	6.43%	13.59%	31.12%	75.29%	77.99%

Figure 4.2 also shows that the rate of outcome measurement adoption has been increased over time. This result supports my expectation I described in section 2.2.1. The outcome measurement has been remarkably increased in1990s and 2000s.

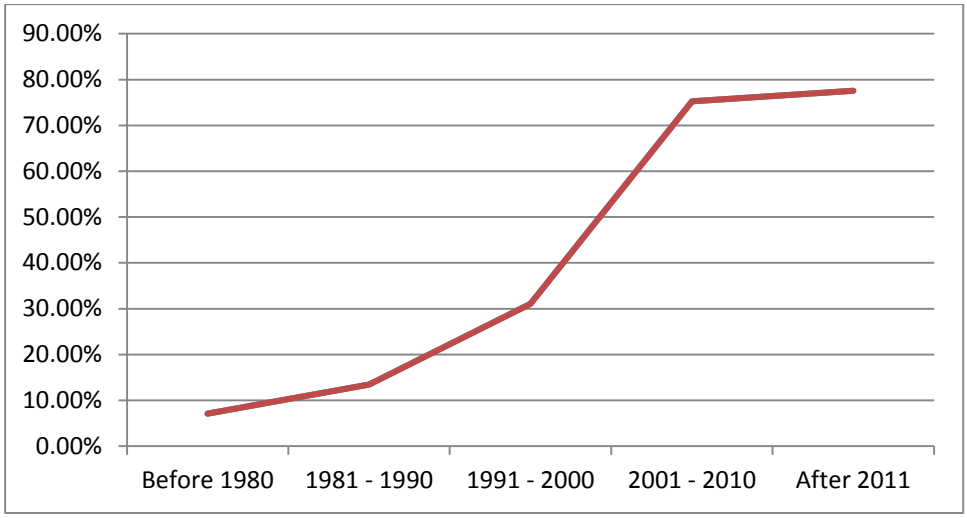

Figure 4.2: Outcome Measurement Rate by Period

4.5.3 Missing Data Imputation

As expected, there are a number of cases with missing information in my dataset due to the lack of knowledge or survey fatigue. The result of analysis may be distorted by improper handling of the missing data. Although there is no simple decision rule for whether to drop the cases with missing values or impute the missing data, if the number of cases with missing values is larger than 5% of the total number of cases, the missing values are not ignorable (Garson, 2012b). As shown in Figure 4.3, more than 40% of cases have missing values. In the following section, I checked whether the missing values of my dataset should

be dropped from the analysis or imputed by comparing the groups with and without missing information and conducting Little's MCAR test.

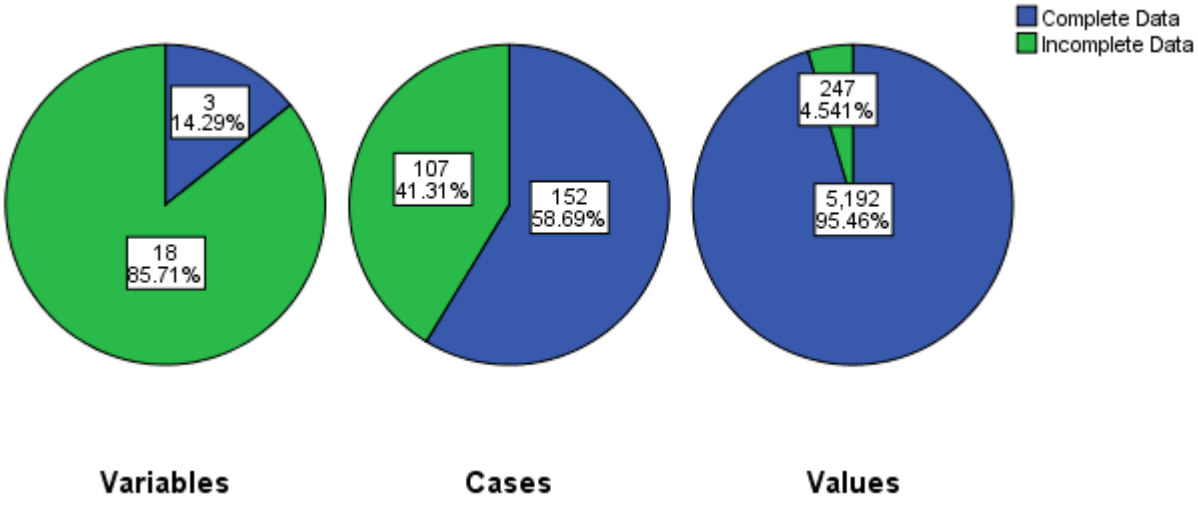

Figure 4.3: Overall summary of missing values

4.5.3.1 Comparing the group with and without missing information

I checked whether the respondents with missing information are systematically different than those that answered all of the questions. Garson (2012b) suggest t-tests of mean differences on key variables to check whether the two groups are significantly different from each other. I conducted independent sample t-test with total expenses, organizational size (FTE), and organizational age. I found insignificant p-value in Table 4.54, 4.55, and 4.56. That is to say, the results of independent samples t-test indicate that the two groups are not significantly different in terms of total expenses, organization size, and organizational age.

Table 4.54: Independent Samples Test (Total Expenses)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% CI of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Total Expenses	Equal variances assumed	1.39	.24	-.52	257	.61	-309,317	599,629	-1,490,130	871,496
	Equal variances not assumed			-.58	256	.57	-309,317	536,840	-1,366,499	747,864

Table 4.55: Independent Samples Test (FTE)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% CI of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
FTE	Equal variances assumed	2.11	.15	-.78	230	.43	-4.52	5.76	-15.87	6.84
	Equal variances not assumed			-.85	172	.40	-4.52	5.31	-14.99	5.96

Table 4.56: Independent Samples Test (Organizational Age)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% CI of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Organizational Age	Equal variances assumed	.91	.34	-1.14	256	.26	-3.49	3.06	-9.52	2.53
	Equal variances not assumed			-1.20	239	.23	-3.49	2.91	-9.23	2.25

4.5.3.2 Multiple Imputation

When missing values are randomly distributed, it is called Missing Completely At Random (MCAR). If the p value for Little's MCAR test is not significant, then the data may be assumed to be MCAR. As shown in Table 4.57, Little's MCAR test is significant (p = .003). This means the missing values in my dataset are not completely at random. If data are not MCAR, missing values should be imputed (Garson, 2012b).

Table 4.57: Little's MCAR test

EM Correlations[a]

	Number of Outcome Measures	Risk Aversion	Uncertainty	Political Capacity	Ln(Organizational Age)	Ln(FTE)	Ln(Total Expenses)
Number of Outcome Measures	1						
Risk Aversion	-.066	1					
Uncertainty	.021	-.125	1				
Political Capacity1	.115	-.357	.097	1			
Ln(Organizational Age)	.188	.057	-.056	-.088	1		
Ln(FTE)	.215	-.028	.052	.034	.309	1	
Ln(Total Expenses)	.228	-.007	.078	.104	.172	.827	1

a. Little's MCAR test: Chi-Square = 95.748, DF = 61, Sig. = .003

The independent t tests in section 4.4.3.1 (Table 4.54, 4.55, and 4.56) shows that the two groups with and without missing data are not significantly different. The Little's MCAR test, however, shows that the missing variable in my dataset is not completely random. In addition, Figure 4.3 shows that 41.31% of the total cases have missing values. Finally, I

decided to impute the missing data by using multiple imputation method to prevent the distortion of results coming from the missing values. According to Yuan (2010), "Multiple Imputation provides a useful strategy for dealing with data sets with missing values. Instead of filling in a single value for each missing value, Rubin's (1987) multiple imputation procedure replaces each missing value with a set of plausible values that represent the uncertainty about the right value to impute" (p.1). I have obtained a new dataset with 1,554 cases because the Multiple Imputation provides 5 imputation models by default.

4.5.4 Analysis Models

4.5.4.1 Ordinal Regression

First of all, I tried the ordinal regression to test my hypotheses because the dependent variable of my study is ordered on one dimension (the adoption of outcome measurement) and ordinal regression supports the use of independent variables which include both categorical factors and continuous covariates. When I ran ordinal regression, however, SPSS displayed a warning message that there are 80% cells with zero frequencies. This is problematic because the rule of thumb to use the ordinal regression is that 80% of cells with count of 5 or more and no cells with zero count (Garson, 2012c). The large percentage of cells with zero frequencies makes the validity of analysis results uncertain. Thus, I decided to use other analytic methods.

4.5.4.2 Poisson Regression

Poisson regression is appropriate for rate data as well as count data. One of my dependent variables - outcome measurement rate - is rate data. I generated this variable by following equation.

Outcome measurement rate = Number of outcome measures

/ Number of total outcome measures listed * 100.

For instance, if some respondent checked 3 out of 5 outcome indicators, I transformed it to 60 (3/5 *100). Even though the data generation process is appropriate for Poisson, however, I need to check whether the data meet the distributional assumptions. First, the main assumption of Poisson distribution is that the mean equals the variance (Institute for Digital Research and Education, 2014). As shown in Table 4.58, there is a huge difference between mean and variance of organization measurement rate. Thus, the distribution of outcome measurement rate does not satisfy the assumption of Poisson distribution.

Table 4.58: Descriptive statistics of outcome measurement rate

	N	Mean	Variance
Outcome Measurement Rate	259	64.5753	759.993
Valid N (listwise)	259		

To confirm the result above, I conducted Kolmogorov-Smirnov test which shows that the assumption of Poisson distribution is met. As shown in Table 4.59, the p-value is

significant, which means I should reject the null hypothesis. The null hypothesis is that entered variable follows the Poisson distribution. That is to say, the significant result means that this variable does not follow the Poisson distribution. Based on the two tests above, I decided to use other analytic methods.

Table 4.59: Kolmogorov-Smirnov Test

Null Hypothesis	Test	Sig	Decision
The distribution of Outcome Measurement Rate is Poisson with mean 64.6	One-Sample Kolmogorov-Smirnov Test	.000	Reject the null hypothesis

4.5.4.3 Binomial Regression

I considered using Generalized Linear Model (GZLM) in this study because it supports not only traditional regression models but also logistic regression models for binary dependents by removing the assumptions of normality, linearity, and content variance (Norusis, 2010; Garson 2012d). For the distribution of dependent variable, I tried binomial distribution. According to Norusis (2010), if the variable has only two values and the observations are independent, then researcher can assume it is the binomial distribution. Also, if the variable is the count of "events" in a fixed number of independent trials, the number of events also follows the binomial distribution.

In my study, I ask 5 outcome measures for P32 and P43 organizations and 8 outcome measures for P46 organizations. The possible answer for each survey item is binary (yes or

no). Also, I count the number of outcome measures (number of "yes"). Hence, the number of outcome measures may follow the binomial distribution.

Next, I checked the three assumptions of binomial distribution. First assumption is that each trial of the process results in one of two possible outcomes. In my dataset, each outcome measure has two options – yes or no. Thus, the first assumption is met in my study. Second assumption is that the trials are independent, that is, a success in one trial does not influence the probability of success in another. In my dataset, the probability of answering "yes" to one outcome measure does not influence the probability of saying "yes" to another outcome measure. Thus, the 2nd assumption also is met. The last assumption is that the probability of success is the same for each trial. In order for the probability of having a particular outcome measure to be equal across measures, the cost and ease of collecting the measure would need to be equal. If I can prove this, I can use the binomial distribution. Measuring the cost and ease of each outcome indicator, however, is beyond the scope of this study. Also, as shown in Table 4.60, the mean of each outcome measure is different from one another. This result implies that the count of outcome measures does not meet the third assumption of binomial distribution. Since the assumption is violated, I decided to use another analytic method.

Table 4.60: Probability of each outcome measure

Outcome Indicator (P32)	Yes
P32 - 1: Increased placement stability	84% (38)
P32 - 2: Client satisfaction with respect to foster care service	78% (35)
P32 - 3: Reduced childhood behavioral and emotional disorders	51% (23)
P32 - 4: Increased quality of life	60% (27)
P32 - 5: Enhanced level of education or life skills of foster child	73% (33)
Outcome Indicator (P43)	**Yes**
P43 - 1: Client satisfaction with respect to the shelter's service	92% (140)
P43 - 2: Reduced physical/mental illness (e.g., depression and other affective disorders)	40% (61)
P43 - 3: Increased quality of life	77% (117)
P43 - 4: Increased self-efficacy	81% (122)
P43 - 5: Reduced childhood behavioral and emotional disorders	33% (49)
Outcome Indicator (P46)	**Yes**
P46 - 1: Client satisfaction with respect to family counseling service	88% (51)
P46 - 2: Increased self-efficacy	71% (39)
P46 - 3: Reduced physical/mental illnesses (e.g., depression and other affective disorders)	77% (43)
P46 - 4: Increased quality of life	75% (42)
P46 - 5: Reduced drug and alcohol use	49% (27)
P46 - 6: Reduced childhood behavioral and emotional disorders	72% (41)
P46 - 7: Reduced marital and relationship problems	67% (36)
P46 - 8: Reduced child abuse and neglect	48% (27)

Note: "n" is reported in parentheses

4.5.4.4 Logistic Regression Model

In this study, I use logistic regression model to test my hypotheses. Binary logistic regression is used when the dependent variable is dichotomous and the independent variables are of any type (Garson, 2012e). For the dependent variable, I used the "number of outcome measures (dichotomous variable)" (See Table 4.20). In the survey, I asked 5 outcome indicators for P32 and P43 organizations and 8 outcome measures for P46 organizations, and

I counted how many outcome indicators each organization uses. I coded 50% or more outcome measures "1" and less than 50% "0". For instance, if some respondent checked 4 out of 5 outcome measures (80%), I coded it 1. Also, I have several categorical variables (e.g., mimetic isomorphism, human resource capacity, political capacity). As shown in Table 4.61, I recoded the five point categorical variables (Strongly Agree, Agree, Neutral, Disagree, and Strongly Disagree) into dichotomous variables.

Table 4.61: Recoding scheme

5 point categorical variable	Dichotomous variable
5: Strongly Agree	1
4: Agree	1
3: Neutral	0
2: Disagree	0
1: Strongly Disagree	0

To find the best model, I examined the models based on the five attributes of good model of Gujarati and Porter (2010): 1) Parsimony; 2) Indentifiability; 3) Goodness of Fit; 4) Theoretical Consistency; and 5) Predictive Power. Table 4.62 shows the best model of this study – logistic regression model with imputed dataset. In this model, I entered only some of the variables I originally planned to use. I have multiple independent variables for measuring same concept. If I include all variables, the model contains more than 30 variables. Considering that I have only 256 cases in my dataset, including too many variables in the

model may produce too high standard errors, which make the model unreliable (Garson, 2012e). Thus, I dropped some of those variables for the sake of parsimony.

In addition, I included three interaction terms. First, as discussed in the literature review section, I included the interaction between mimetic isomorphism (largest organizations' outcome measurement) and uncertainty because when there is a high degree of uncertainty, managers try to model themselves after the other organizations in the same field (Mizruchi & FeinSource, 1999). Second, I included an interaction term between accreditation and risk aversion because risk aversion might negatively influence the relationship between accreditation and the adoption of outcome measurement. For risk averse organizations, the accreditation may less positively influence the adoption of outcome measurement than for risk taking organizations. Third, I included an interaction term between largest organizations' outcome measurement and risk aversion because the positive relationship between largest organizations' outcome measurement and adoption of outcome measurement can be influenced by the risk aversion. Mimetic isomorphism is expected to positively influence the adoption of outcome measurement. If the organization is risk averse, however, the positive influence of the largest organizations' outcome measurement on the adoption of outcome measurement might be attenuated.

As shown in Table 4.62, United Way funding[1], Largest organizations' outcome measurement[2], Accreditation[3], Human Resource Capacity[4], Political Capacity (Political

[1] Hypothesis: The organizations which receive funding from United Way (coercive isomorphism) are more likely to adopt the outcome measurement.
[2] Hypothesis: The mimetic pressure from the largest organizations in the same field is positively related to the adoption of outcome measurement.
[3] Hypothesis: Accreditation (Professionalism) is positively related to the adoption of outcome measurement.

Negotiation)[5], and Total Expenses (control variable) are positively significant in the logistic regression model. In addition, the interaction between risk aversion and accreditation is negatively significant. Although the Risk Aversion is not significant variable in the model, this interaction term reflects the negative influence of Risk Aversion on the positive relationship between Accreditation and the adoption of outcome measurement. The odds ratio column indicates the effect size measures - relative importance - of the independent variables in this model. I discuss the odds ratio in the discussion section.

[4] Hypothesis: Human resource capacity is positively related to the adoption of outcome measurement.
[5] Hypothesis: Political capacity is positively related to the adoption of outcome measurement.

Table 4.62: Logistic Regression Model

Independent Variables	B	S.E.	Sig.	Odds Ratio
NTEE(1) – P32	.986	.592	.096	2.682
NTEE(2) – P43	.724	.437	.097	2.064
Risk Aversion	.970	.555	.081	2.639
United Way Funding	1.179**	.406	.004	3.251
Largest organizations' outcome measurement	4.621*	2.107	.029	101.618
Uncertainty	.517	.411	.209	1.678
Largest organizations' outcome measurement* Uncertainty	-.765	.493	.123	.465
Accreditation	3.361**	1.175	.004	28.821
Ln(Organizational Age)	.432	.382	.257	1.541
Ln(Total Expenses in 2011)	.448*	.221	.043	1.565
Financial Capacity	-.367	.443	.407	.693
Human Resource Capacity	1.046*	.518	.044	2.847
Infrastructure Capacity	-.699	.441	.113	.497
Administrative Capacity	-.704	1.159	.543	.494
Political Capacity (Political Negotiation)	.765*	.380	.044	2.148
Network Capacity	-.173	.411	.674	.841
Accreditation * Risk Aversion	-1.348**	.512	.008	.260
Largest organizations' outcome measurement* Risk Aversion	-.993	.581	.087	.371
Constant	-11.711	3.847	.002	.000

Note: This logistic model was obtained from the imputed data
Dependent variable: Number of outcome measures (dichotomous variable)
P46 is reference category
Number of cases: 259
-2 Log likelihood of original data: 158.03
Nagelkerke R Square of original data: 0.32
Degrees of freedom: 18
**p < .01, *p < .05 (two-tailed tests)

Next, I conducted the significant test of the logistic regression. Hosmer and Lemeshow test is recommended for the overall fit test of logistic regression model. Non-significant result means that the model fits well. Also, an alternative method to the Hosmer and Lemeshow test is Omnibus test which reports significance levels by the traditional chi-

square method. For the Omnibus test, significant finding means that at least one of the independent variables is significant (Garson, 2012e).

The SPSS logistic regression does not produce significant test statistics with the pooled data. It only reports the statistics with original data and imputed data 1 through 5. Thus, I check the results from the regression without the imputed missing data and the results from the pooled data. If the results are similar, then I conclude that the model fit statistics from the model with the missing data would probably be a good indicator for the model with the pooled data.

As shown in Table 4.63, I compared the results from the model with original data and that with pooled data to judge whether I can use the significance test results of the model with original data. Table 4.63 indicates that the two results are somewhat similar. Thus, I decided to report the significant test results of the model with original dataset.

Table 4.63: Similarity between results from the regression with and without missing data

Similarity	Only in the pooled dataset	Only in the original dataset
- United Way funding is positively significant in both models. - Largest organizations' outcome measurement is positively significant in both models. - Accreditation is positively significant in both models. - Interaction between risk aversion and accreditation is negatively significant in both models. - Total expenses is positively significant in both models	- Human resource capacity is positively significant only in the pooled dataset. - Political capacity is positively significant only in the pooled dataset.	- Risk aversion is positively significant only in the original dataset. - Interaction between risk aversion and largest organizations' outcome measurement is negatively significant only in the original dataset.

As shown in Table 4.64, the Hosmer and Lemeshow test indicates a non-significant result in the original dataset. A finding of non-significance corresponds to the conclusion that the overall model adequately fits.

Table 4.64: Hosmer and Lemeshow Test

Imputation Number	Step	Chi-square	df	Sig.
Original data	1	11.962	8	.153
1	1	10.305	8	.244
2	1	11.048	8	.199
3	1	5.441	8	.710
4	1	7.659	8	.467
5	1	10.376	8	.240

As shown in Table 4.65, the Omnibus test indicates a significant result in the original dataset. A finding of significance of Omnibus test corresponds to the conclusion that the overall model adequately fits. That is to say, at least one of the predictors is significantly related to the dependent variable.

Table 4.65: Omnibus Tests of Model Coefficients

Imputation Number			Chi-square	df	Sig.
Original data	Step 1	Step	43.638	18	.001
		Block	43.638	18	.001
		Model	43.638	18	.001
1	Step 1	Step	51.539	18	.000
		Block	51.539	18	.000
		Model	51.539	18	.000
2	Step 1	Step	52.938	18	.000
		Block	52.938	18	.000
		Model	52.938	18	.000
3	Step 1	Step	57.134	18	.000
		Block	57.134	18	.000
		Model	57.134	18	.000
4	Step 1	Step	55.481	18	.000
		Block	55.481	18	.000
		Model	55.481	18	.000
	Step 1	Step	51.052	18	.000
		Block	51.052	18	.000
		Model	51.052	18	.000

4.5.5 Hypothesis testing and discussion

Prior studies have found a variation in the adoption of outcome measurement, and this dissertation investigated what factors explain this variation in the adoption of outcome measures, focusing on human service nonprofits. I suggested eleven hypotheses using agency

theory, institutional theory, and the measures of organizational capacity. In the following paragraphs, I describe what hypotheses are supported (or not supported) by my research findings.

Hypothsis1 (Agency Theory)

Hypothesis 1: Risk aversion of the nonprofit organization is negatively related to the adoption of outcome measurement.

First, I expected risk aversion of the nonprofits to negatively influence the adoption of outcome measurement. According to the logistic regression model, the risk aversion is not significant at a .05 level ($p = 0.081$). While this result is unexpected, the interaction term between risk aversion and accreditation indicates the significant influence of risk aversion on the adoption of outcome measurement. The interaction term between risk aversion and accreditation is negatively significant, while accreditation is positively significant. That means the positive relationship between accreditation and the adoption of outcome measurement is negatively influenced by the risk aversion. In other words, accredited organizations are more likely to adopt outcome measurement. If the organization is risk averse, however, the positive influence of accreditation on the adoption of outcome measurement is decreased. In this sense, although the risk aversion is not significant at a .05 level, the interaction term between risk aversion and accreditation supports the hypothesis1. I discuss the odds ratio of this interaction term in the accreditation section.

Hypothesis 2 (Institutional Theory – Coercive Isomorphism)

Hypothesis 2: Coercive isomorphism is positively related to the adoption of outcome measurement.

The second hypothesis is that coercive isomorphism positively influences the adoption of outcome measurement. In the logistic regression model, "United Way funding" is positive and significant (p = .004). Thus, hypothesis 2 is supported by the research finding. This finding is reasonable in that United Way emphasizes outcome measurement as a main way for evaluating their funded programs and encourages funding recipients to adopt the outcome measurement (Hendricks, Plantz, & Pritchard, 2008). The odds ratio of "United Way funding" indicates that organizations which receive funding from the United Way are 3.251 times more likely to adopt outcome measurement than the organizations which do not receive the United Way funding (See Table 4.62).

Hypothesis 3 (Institutional Theory – Mimetic Isomorphism)

Hypothesis 3: The mimetic pressure from the other organizations in the same field is positively related to the adoption of outcome measurement.

It was hypothesized that the mimetic isomorphism is positively related to the adoption of outcome measurement. In the logistic regression model, "largest organizations' outcome measurement" is positive and significant (p = .029). Thus, the research finding supports the hypothesis3. Barman and MacIndoe (2012) found that the relationship between mimetic pressure and the adoption of outcome measurement to be insignificant. They explained the insignificant result in that outcome measurement is not widespread yet, while mimetic

isomorphism is valid only when the practice is dominant in the field (Tolbert & Zucker, 1983; Barman & MacIndoe, 2012). As I expected, my study result is different from that of Barman and MacIndoe (2012). I focused only on human service nonprofits, while Barman and MacIndoe (2012) collected the survey samples in diverse sub-industries. Since human service nonprofits are increasingly adopting the outcome measurement (Carrilio, Packard, Clapp, 2004; Julian & Clapp, 2000; Horsch, 1996), it is comparatively dominant in human service nonprofits. In this sense, focusing only on human service nonprofits produced a significant result. As shown in the odds ratio column of Table 4.62, organizations more influenced by mimetic isomorphism are 101.618 times more likely to adopt outcome measurement than the organizations which less influenced by mimetic isomorphism.

Hypothesis 4(Institutional Theory - Uncertainty)

Hypothesis 4: Uncertainty has a positive interaction effect on the relationship between mimetic isomorphic pressure and the adoption of outcome measurement.

I expected uncertainty to positively influence the relationship between mimetic isomorphism and the adoption of outcome measurement because "the more uncertain the relationship between means and ends the greater the extent to which an organization will model itself after organizations it perceives to be successful" (DiMaggio & Powell, 1983, p. 154). Also, the mimetic isomorphism is a result of "standard responses to uncertainty" (DiMaggio & Powell, 1983, p.150).

As shown in the logistic regression model, however, the interaction term between largest organizations' outcome measurement and uncertainty is not significant ($p = .123$).

This means that uncertainty is not significantly related to the relationship between mimetic isomorphic pressure and the adoption of outcome measurement. In the following paragraph, I explain the lack of support for the hypothesis4.

Agency theorists argue that environmental uncertainty makes it hard for principals to correctly assess the outcome of agents because the uncertainty may create an unexpected outcome (e.g., Celly & Frazier 1996). Bergen, Dutta, and Walker Jr, (1992) and Hendry (2002) describe that it is costly for principals to transfer the risk to agents using outcome based contracts when the uncertainty is high due to uncontrollable factors coming from the environment. In addition, it is difficult to control the outcomes in an uncertain environment because the outcomes fluctuate more widely, compared to a stable environment (Kumar, Scheer, & Steenkamp, 1995). In this sense, prior studies based on agency theory suggest that environmental uncertainty is positively related to behavior based contracts and negatively related to outcome based contracts (Celly & Frazier, 1996; Bergen, Dutta, & Walker Jr, 1992; Hendry, 2002).

Institutional theorists argue that organizations imitate other organizations' behaviors in a response to uncertainty. Thus, I expected to see that the uncertainty is positively related to the relationship between mimetic isomorphism and the adoption of outcome measurement. On the other hand, agency theorists argue that the environmental uncertainty is negatively related to outcome based contracts because the environmental uncertainty may create an unexpected outcome and the excuse making of agents can be prevented by using behavior based contracts. Less use of outcome based contacts may lead to less adoption of outcome measurement. Potentially, these conflicting effects of environmental uncertainty might be

leading to an insignificant effect of uncertainty on the adoption of outcome measurement in the logistic regression model.

Hypothesis 5 (Institutional Theory – Normative Isomorphism)

Hypothesis 5: Professionalism is positively related to the adoption of outcome measurement.

Hypothesis 5 predicts that organizations with professionalism are more likely to adopt the outcome measurement. In the logistic regression model, "accreditation" is positive and significant (p = .004). Hence, the research finding supports Hypothesis 5. Since the accreditation is involved in an interaction effect, the b coefficient of accreditation must be adjusted before exponentiation as follows.

Adjusted b coefficient of accreditation = 3.361 – 1.348*risk aversion.

I interpret the interaction term between risk aversion and accreditation based on Garson (2012e) and Li and Barry (2014). When the risk aversion is 1, the b coefficient and odds ratio of accreditation are 2.013 and 7.486 respectively (See Table 4.66). That is to say, when the risk aversion is 1 (the lowest level), organizations with accreditation are 7.486 times more likely to adopt the outcome measurement than organizations without accreditation. However, the positive influence of accreditation on the adoption of outcome measurement is decreased when the risk aversion is increased. When the risk aversion is increased to 2 from 1, the b coefficient and odds ratio of accreditation are 0.665 and 1.944 respectively. In other words, when the risk aversion is 2, organizations with accreditation are 1.944 times more likely to adopt outcome measurement than organizations without

accreditation. The B coefficient of accreditation is positive at low levels of risk aversion and negative at high levels of risk aversion. This indicates that the positive effect of professionalism is decreased when the organization is more risk averse. When the risk aversion is more than 3, the negative effect of risk aversion on the adoption of outcome measurement outweighs the positive effect of accreditation. Thus, the sign of accreditation becomes negative when the risk aversion level is 3 or higher.

Table 4.66: Adjusted B coefficient and odds ratio of accreditation

Risk aversion	Adj B	Odds Ratio
1 (The least risk averse)	2.013	7.486
2	0.665	1.944
3	-0.683	0.505
4	-2.031	0.131
5 (The most risk averse)	-3.379	0.034

Hypothesis 6a (Organizational Capacity: Financial Capacity)

Hypothesis 6a: Financial capacity is positively related to the adoption of outcome measurement.

I hypothesized that organizations with financial capacity are more likely to adopt the outcome measurement. As shown in logistic regression model, however, financial capacity is not significant (p = .407). This means the financial capacity does not significantly influence the adoption of outcome measurement. One plausible explanation for this contradictory finding is that I filtered small nonprofits with expenses below $500,000 and focused on mid

and large nonprofits, thus the respondent group might not capture full range of variation in the financial capacity of population group. As shown in Table 4.67, there are differences in financial capacity (ER) between respondents and population group. Overall, the ER of respondent group is higher than that of population group.

Table 4.67: ER of respondents and population group

	Respondent group (n = 259)			Population group (n = 1039)		
NTEE	Mean	Std. Deviation	Median	Mean	Std. Deviation	Median
P32	0.51	0.73	0.75	0.31	2.34	0.67
P43	0.81	0.20	0.88	0.77	0.49	0.88
P46	0.68	0.46	0.84	0.62	0.48	0.77
Total	0.73	0.42	0.84	0.63	1.24	0.82

As I did in section 4.4 (Representativeness of respondent group), I examined how different the mean of ER of the respondent group is from the true mean of ER of the population, using one sample t-test. As shown in Table 4.68, the 95% confidence interval of the difference of mean between respondents and population group does not contains zero, which means that the mean of ER of respondent group is significantly different from that of the population group with 95% confidence level.

Table 4.68: One sample t-test, ER (all groups)

	Test Value = 0.63					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
ER	3.68	258	.00	.10	.04	.15

Table 4.69 indicates that the 95% confidence interval of the difference of mean between P32 respondents and population group contains zero, which means that the mean of ER of P32 respondent group is not significantly different from that of the population group with 95% confidence level.

Table 4.69: One sample t-test, ER (P32)

	Test Value = 0.31					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
ER	1.85	44	.07	.20	-.02	.42

As shown in Table 4.70, the 95% confidence interval of the difference of mean between P43 respondents and population group does not contains zero, which means that the mean of ER of P43 respondent group is significantly different from that of the population group with 95% confidence level.

Table 4.70: One sample t-test, ER (P43)

	Test Value = 0.77					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
ER	2.21	153	.03	.04	.004	.07

Table 4.71 indicates that the 95% confidence interval of the difference of mean between P46 respondents and population group contains zero, which means that the mean of ER of P46 respondent group is not significantly different from that of the population group with 95% confidence level.

Table 4.71: One sample t-test, ER (P46)

	Test Value = 0.62					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
ER	1.03	59	.31	.06	-.06	.18

The one sample t-test shows that the ER of all respondent group is significantly higher than that of population group although the one sample t-test results were not significant in P32 and P46 group. This indicates that the respondent group's financial capacity is significantly higher than population's one. As I mentioned in the section 4.4, there

is a bias in the P43 respondent group toward larger nonprofits in terms of total expenses. Table 4.70 shows that P43 respondent group is financially healthier than P43 population group. This might be the reason why the influence of financial capacity on the adoption of outcome measurement was not significant in the logistic regression model. Future research is needed to examine whether considering the whole variation in the ER produces a different finding.

Hypothesis 6b (Organizational Capacity: Human Resource Capacity)

Hypothesis 6b: Human resource capacity is positively related to the adoption of outcome measurement.

I expected that organizations with human resource capacity are more likely to adopt the outcome measurement. Staff with knowledge about outcome measurement is needed for adopting the outcome measurement. The logistic regression model indicates that human resource capacity is positive and significant ($p = .044$). Thus, hypothesis 6b is supported by the research finding. As shown in the odds ratio column of logistic regression model, organizations with more human resource capacity are 2.847 times more likely to adopt the outcome measurement than organizations with less human resource capacity.

Hypothesis 6c (Organizational Capacity: Technical Capacity)

Hypothesis 6c: Technical capacity is positively related to the adoption of outcome measurement.

It was hypothesized that organizations with technical capacity are more likely to adopt the outcome measurement. Interestingly, the logistic regression result indicates that technical capacity is not significant ($p = .113$). It is notable that the adequate technical capacity (e.g., information technology, communication system) for daily operations is not a significant factor which influences the adoption of outcome measurement. This contradictory finding might be explained by the prior studies. I measure technical capacity by asking participants to evaluate the following survey item: "Our organization has adequate infrastructure (e.g., information technology, communication system) for daily operations". More than 73% of respondents agreed or strongly agreed with this statement in my survey.

However, prior studies suggest a different point. While information technology requires an intensive training for effective use (Corder, 2001), the training of information technology has received insufficient funding in nonprofit organizations (Hackler & Saxton, 2007). According to an interview conducted by Saigal (2008), one executive director of a nonprofit organization said, "Had we more time for more training and getting everybody on the same page to learn it, we could be even more efficient. But who has time for that?" (p.35). In this sense, Saigal (2008) points out the lack of staff training as one of the key barriers to the effective use of information technology in nonprofits. While I found that the human service nonprofits have adequate infrastructure, prior studies address that nonprofits do not have sufficient time and funding for the training of the information technology use. If nonprofits cannot utilize their infrastructure for the outcome measurement due to the lack of training, the infrastructure might not be significantly related to the adoption of outcome measurement. In this sense, the influence of technical capacity on the adoption of outcome

measurement might not be supported by the empirical results in my study because I did not capture the organization's ability to use the infrastructure.

Hypothesis 6d (Organizational Capacity: Administrative Capacity)

Hypothesis 6d: Administrative capacity is positively related to the adoption of outcome measurement.

Hypothesis 6d predicts that organizations with administrative capacity are more likely to adopt outcome measurement. The logistic regression model, however, shows that the administrative capacity is not significant ($p = .543$). This might be explained by the insufficient variation in the administrative capacity. More than 98% of respondents agreed or strongly agreed that their organization has written rules concerning compensation, governance, job descriptions, conflict of interest, and/or volunteers. Since the majority of the respondents answered that they have adequate administrative capacity, the effect of administrative capacity on the adoption of outcome measurement might not be significant in the logistic regression model.

Hypothesis 6e (Organizational Capacity: Political Capacity)

Hypothesis 6e: Political capacity is positively related to the adoption of outcome measurement.

I expected a positive relationship between political capacity and the adoption of outcome measurement. My research findings support the Hypothesis 6e. The logistic regression model indicates that the political capacity (political negotiation) is positively

significant at a .05 level (p = .044). When organizations adopt outcome measurement, political capacity is needed because it helps the organizations establish a consensus, develop strategies to convince stakeholders, and obtain their support (Berman & Wang, 2010; Héritier, 2003). The political capacity of organizations is based on their ability to negotiate with stakeholders to establish a consensus in a short period of time (Héritier, 2003). The political negotiation helps ensure compliance among the stakeholders and convince them to adopt and implement organizational practices (Borzel, 2002). As shown in the odds ratio column of the logistic regression model, organizations with more political capacity are 2.148 times more likely to adopt outcome measurement than the organizations with less political capacity.

Hypothesis 6f (Organizational Capacity: Network Capacity)

Hypothesis 6f: Network capacity is positively related to the adoption of outcome measurement.

It was hypothesized that organizations with more network capacity are more likely to adopt outcome measurement. Surprisingly, network capacity is not significant (p = .674) in the logistic regression model. This result indicates that seeking help from the other organizations does not significantly influence the adoption of outcome measurement. This contradictory finding might be explained by examining the characteristics of nonprofits. In the human service sector, nonprofits have a relatively diverse network, while public organizations have "centralized and homogeneous network of technical expertise" (Corder, 2001, p.202). Nonprofit agencies provide more various human services and have more diverse resources than public organizations. Furthermore, board members and staff have a

diverse background across nonprofit agencies. Corder (2001) argues, "The diverse demands of donors, agency leaders, and state government coupled with the discretion afforded the nonprofit agency suggest that there will be considerably more variation in technology in the nonprofit sector than the public sector" (p.203). In this sense, the resources coming from the network might be more diverse in nonprofits, compared to public organizations. Another capacity to coordinate, manage and facilitate the diverse resources in the network might be needed (Cohen & Levinthal, 1990). Without considering this capacity, the influence of network capacity on the adoption of outcome measurement might not be significant in the empirical research.

Control Variables (Organizational Age and Total Expenses)

In this study, I included organizational age and size as control variables. First, the organizational age was not significant in the logistic regression model. As I described in the literature review section, I expected that the organizational age can have both a positive and negative effect. Combined, these opposing effects may cancel each other out, thus resulting in a non-significant result. This canceling out effect of the contradictory relationship has been examined in prior studies (e.g., MacKinnon, Krull, & Lockwood, 2000; MacKinnon, Fairchild, & Fritz, 2007). Second, I predicted that organizational size would positively influence the adoption of outcome measurement. As expected, the organizational size measured by total expenses was a significant factor in the logistic regression model ($p = .043$).

CHAPTER 5: IMPLICATION AND CONCLUSION

5.1 Introduction

In the previous chapter, I described the survey response rate and data quality check. Second, I discussed the descriptive statistics of the main variables and the missing data imputation. Third, I explained what analytic models I tried and how I ended up with the logistic regression model as the best model to test the hypotheses. Lastly, I discussed my research findings. In this section, I describe the results of preliminary study of outcome information use, implications and limitations of this study and the future research questions.

5.2 Use of outcome information

In this dissertation, I mainly focused on the adoption of outcome measurement. However, I examined the use of outcome information at a preliminary level as well because the performance measurement is composed of two different stages: "adoption (development of measures) and implementation (actual use)" (Julnes & Holzer, 2001, p.703), and the success of performance measurement can be evaluated by examining the use of performance information (Van Dooren, 2008). I listed 13 purposes of outcome information uses in the survey, which were measured by a five point scale: 1 "Strongly disagree", 2 "Disagree", 3 "Neutral", 4 "Agree", and 5 "Strongly agree". The means of 13 uses of outcome information are described in Table 5.1. Also, as shown in the one sample t-test results (Table 5.2), the means of some outcome information uses are significantly higher or lower than the average level. For instance, the mean of use 8 (Respond to public demand for accountability) is

significantly lower than the average. On the other hand, the means of use 10 (Comply with government requirements), use 11 (Report to funders), and use 13 (Outreach and public relations) are significantly higher than the average.

Table 5.1: Use of outcome information

Use of outcome Information	Valid	Missing	Mean
Use 1: Improve management decisions	252	7	3.44
Use 2: Improve strategic planning	253	6	3.63
Use 3: Improve services for clients	252	7	3.18
Use 4: Appraise employee performance	251	8	3.57
Use 5: Support budget decisions	255	4	3.76
Use 6: Report to the board	241	18	3.05
Use 7: Improve communication between administrators and elected officials	242	17	3.30
Use 8: Respond to public demands for accountability	234	25	2.74
Use 9: Respond to pressure from community groups	246	13	3.87
Use 10: Comply with government requirements	253	6	4.13
Use 11: Report to funders	252	7	4.03
Use 12: Get new funding	252	7	3.63
Use 13: Outreach and public relations	252	7	4.00
Average of 13 uses	214	45	3.57

Table 5.2: One sample t-test of 13 outcome information uses

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Use_1	55.66	251	0	3.44	3.32	3.57
Use_2	58.38	252	0	3.63	3.51	3.75
Use_3	45.93	251	0	3.18	3.04	3.31
Use_4	55.72	250	0	3.57	3.44	3.70
Use_5	63.06	254	0	3.76	3.64	3.88
Use_6	40.41	240	0	3.05	2.90	3.19
Use_7	45.22	241	0	3.30	3.16	3.45
Use_8	35.70	233	0	2.74	2.59	2.89
Use_9	54.93	245	0	3.87	3.73	4.00
Use_10	68.71	252	0	4.13	4.01	4.24
Use_11	61.79	251	0	4.03	3.90	4.16
Use_12	56.22	251	0	3.63	3.51	3.76
Use_13	71.86	251	0	4.00	3.89	4.11
Average Use	69.60	214	0	3.57	3.47	3.67

As shown in Table 5.3, the 13 uses of outcome information can be categorized into three dimensions, using factor analysis. Kearns (1994) categorizes accountability into four types based on two dimensions: "(1) a set of performance standards - explicit or implicit - generated by the organization's strategic environment, and (2) a response - reactive or proactive - from inside the organization" (p. 187). Using Kearns (1994)' framework, I name the three components of outcome information use as follows.

1) Component 1 (outcome information use 1 - 5 and 13): Professional/discretionary uses of outcome information. The uses of outcome information in component 1 are proactive and implicit.

2) Component 2 (outcome information use 9 - 12): Compliance uses of outcome information. This category contains the uses of outcome information which are reactive and explicit.

3) Component 3 (outcome information use 6 - 8): Negotiated uses of outcome information. The uses of outcome information in this category are reactive and implicit.

As shown in Table 5.4, the means of component 1, 2, and 3 are 3.61, 3.91, and 3.03 respectively. In addition, the one sample t-test results in Table 5.4 indicate that the mean of component 2 (compliance use of outcome information) is significantly higher than the average of outcome information use and that of component 3 (negotiated use of outcome information) is significantly lower than the average. It is a fruitful area for future research to investigate which organizations or what factors are associated with how the outcome information is used. Although it is beyond the scope of this study, I examined the relationship between the 3 dimensions of outcome information use, and risk aversion, institutional factors, and the measures of organizational capacity by correlation.
Since correlation does not provide evidence of causation (Garson, 2012f), the uses of outcome information cannot be investigated only with the correlation results. However, the correlation matrix provides a starting point to study the outcome information use by measuring the degree of relationship between variables – three dimensions of outcome

information use and organizational factors (risk aversion, institutional factors, and measures of organizational capacity)

Table 5.3: Rotated Component Matrix, Outcome Information Use

	Component		
	1	2	3
Use_1	.822	.159	.256
Use_2	.826	.332	.131
Use_3	.725	.057	.358
Use_4	.795	.164	.299
Use_5	.726	.392	.154
Use_6	.389	.231	.691
Use_7	.228	.336	.780
Use_8	.227	.133	.856
Use_9	.163	.649	.375
Use_10	.245	.861	.080
Use_11	.194	.837	.176
Use_12	.391	.618	.303
Use_13	.734	.413	.174

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 6 iterations.

Table 5.4: One sample t-test of outcome information use (3 components)

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Component1	69.59	244	0	3.61	3.51	3.72
Component2	72.68	241	0	3.91	3.81	4.02
Component3	45.18	224	0	3.03	2.89	3.16
Average Use	69.60	214	0	3.57	3.47	3.67

As shown in Table 5.5, Professional/discretionary uses of outcome information have a significantly negative correlation with the Risk Aversion, while the Compliance uses and Negotiated uses of outcome information are not significantly correlated with the Risk Aversion. The correlation coefficient (-.380) indicates the relationship is a moderate level. This negative correlation indicates that more risk aversion is associated with less Professional/discretionary uses of outcome information. Using the outcome information for management decisions and strategic planning (Professional/discretionary uses of outcome information) is inherently risky because it implies making changes in the organization and changes involve risk (Greve, 1996). Conversely, since low level of Risk Aversion means risk taking, this negative correlation indicates that the Professional/discretionary uses of outcome information are positively correlated with risk taking attitude. Literature argues that risk taking organizations are likely to have proactive behaviors (e.g., Naldi, Nordqvist, Sjöberg, & Wiklund, 2007; Covin & Slevin, 1999). The Professional/discretionary uses of outcome information include more proactive uses than the other dimensions of outcome information

uses. In this sense, risk taking organizations may proactively try to use the outcome information to improve management decision, strategic planning, and services for their clients (Professional/discretionary uses). Although the correlation matrix does not indicate a causal relationship, Table 5.5 implies a negative relationship between risk aversion and the Professional/discretionary uses of outcome Information (positive relationship between risk taking and the Professional/discretionary uses). This relationship is a potential area for future research.

Table 5.5: Correlation between outcome information use and risk aversion

	Professional/ discretionary uses	Compliance uses	Negotiated uses	Risk Aversion
Professional/ discretionary uses	1			
Compliance uses	.060	1		
Negotiated uses	-.045	-.056	1	
Risk Aversion	-.380**	-.151	-.181	1

Note: This correlation matrix was obtained from the imputed data

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed

Table 5.6 shows how institutional factors are correlated with the outcome information use. United Way funding is positively correlated with the Compliance uses of outcome information. Since the United Way funding is a measure of coercive isomorphism, Table 5.6 indicates that Compliance uses of outcome information are positively correlated with the

coercive isomorphism. It does not necessarily mean that coercive isomorphism leads to Compliance uses because correlation does not imply causation. Nevertheless, this relationship makes sense in that Compliance uses include responding to pressure from community groups and complying with government requirements. Pressure from community group and government requirements may form a coercive pressure. Also, Largest organization's outcome measurement – mimetic isomorphism - is positively correlated with the Compliance uses of outcome information as well. Since institutional theory addresses organizational behaviors reacting to or complying with institutional pressures (e.g., Kondra & Hinings, 1998; Tolbert & Zucker, 1996), the positive correlation between Compliance uses of outcome information and the institutional factors are theoretically reasonable. On the other hand, Table 5.6 also shows that Largest organizations' outcome measurement is negatively correlated with the Professional/discretionary uses of outcome information. Although this does not necessarily mean mimetic isomorphism is negatively related to the Professional/discretionary uses of outcome information, the correlation matrix implies a negative relationship between them. For some organizations influenced by institutional isomorphism, adoption of organizational practices, including performance measurement, might be symbolic (e.g., Roy & Seguin, 2000; Modell, 2001& 2009; Yang, 2009). Once organizations achieve the symbolic benefit by adopting the outcome measurement, they may not try to proactively implement the outcome measurement, that is, use the outcome information to improve the strategic planning and management decisions (Professional/discretionary uses). This prediction cannot be supported only by the correlation

result below. Further literature review and empirical studies are needed to address this relationship.

Table 5.6: Correlation between outcome information use and institutional factors

	Professional/ discretionary uses	Compliance uses	Negotiated uses	United Way funding	Largest Organization's outcome measurement	Accreditation
Professional/ discretionary uses	1					
Compliance uses	.060	1				
Negotiated uses	-.045	-.056	1			
United Way funding	-.073	.257**	.025	1		
Largest organization's outcome measurement	-.196**	.141*	.182	.194**	1	
Accreditation	.096	-.034	.077	.122	.061	1

Note: This correlation matrix was obtained from the imputed data

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Lastly, Table 5.7 shows how the uses of outcome information are correlated with organizational capacity. Overall, the dimensions of organizational capacity are positively correlated with the Professional/discretionary uses of outcome information. They have a positive correlation with Human Resource Capacity, Technical Capacity, Administrative Capacity, and Network Capacity. As shown in Table 5.1, Professional/discretionary uses include appraising employee performance, supporting budget decision, and improving management decisions, strategic planning, and services for clients. In my dissertation, measures of organizational capacity were not significant factors to explain the adoption of outcome measurement except for political and human resource capacity (See Table 4.62). Although correlation does not indicate a causal relationship, the positive correlation between organizational capacity and Professional/discretional uses of outcome information implies the organizational capacity may be needed in the actual uses of outcome information, particularly for the Professional/discretionary uses. Future researchers need to study how the dimensions of organizational capacity are related to the process of outcome measurement implementation. Lastly, Table 5.7 shows that Negotiated uses of outcome information are positively correlated with Political Capacity (negotiation) and Network Capacity, while the Negotiated uses were not correlated with risk aversion and institutional factors in Table 5.5 and 5.6. These positive correlations are reasonable in that the negotiated uses involve "some form of negotiation between the agency and its environment" (Kearns, 1994, p.188). Political capacity to negotiate with stakeholders and organizational ability to utilize organizational network may be associated with the negotiated uses of outcome information (e.g., reporting

the outcome information to board, politicians, and general public). These relationships also require further research.

Overall, Table 5.6 and 5.7 indicate that the Compliance uses of outcome information are correlated with institutional factors rather than organizational capacity and risk taking attitude and Professional/discretionary uses of outcome information are correlated with organizational capacity and risk taking attitude rather than institutional factors.

Table 5.7: Correlation between outcome information use and organizational capacity

	Professional/ discretionary uses	Compliance uses	Negotiated uses	Financial Capacity	Human Resource Capacity	Technical Capacity	Administrative Capacity	Political Capacity (Negotiation)	Network Capacity
Professional/ discretionary uses	1								
Compliance uses	.060	1							
Negotiated uses	-.045	-.056	1						
Financial Capacity	-.056	.140	-.063	1					
Human Resource Capacity	.194**	.142	.062	-.029	1				
Technical Capacity	.226**	.094	-.027	-.011	.342**	1			
Administrative Capacity	.195	.235	-.013	.120	.127	.056	1		
Political Capacity (Negotiation)	.310**	.103	.224**	.062	.133*	.134*	.098	1	
Network Capacity	.178*	.055	.149*	.140*	.057	.096	-.016	.330**	1

Note: This correlation matrix was obtained from the imputed data

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

5.3 Implication

To explain the variation in the adoption of outcome measurement, I empirically tested how risk aversion, institutional factors, and the measures of organizational capacity influence the adoption of outcome measurement. I found that coercive, mimetic, and normative isomorphic pressures are positively related to the adoption of outcome measurement. In addition, I found that risk aversion is negatively related to the relationship between normative isomorphism and the adoption of outcome measurement. Lastly, human resource capacity and political capacity out of the six dimensions of organizational capacity were positively significant in the logistic regression model. These findings indicate that institutional theory fits my study explaining the variation in the adoption of outcome measurement well, compared to the other theories. However, this does not imply that institutional theory is the most important theory in the study of outcome measurement because there was no equal weight (e.g., number of variables) among the theories in this dissertation. Although I considered only risk aversion among the main components of the agency theory in this dissertation, the other variables (e.g., motivation, information asymmetry) of the agency theory may be included in the future studies. Future studies considering the equal weight among the theories may examine the exploratory power of the theories against one another in the study of outcome measurement. In the following paragraphs, I describe the implications of the findings of this study.

Institutional factors

My dissertation addressed how institutional theory is adequately applied to the study of performance measurement. Overall, my research findings suggest that organizational behavior to achieve the institutional legitimacy can influence the outcome measurement practice. In the prior studies based on operational improvement and strategic management tradition, performance measurement has been viewed as a management tool to improve work process and ultimately achieve organizational goals (e.g., LeRoux & Write, 2010; Kaplan & Norton, 1996; Pettigrew, Woodman, & Cameron, 2001; Russ-Eft & Preskill, 2009; Skinner, 2004; Poister, 2010; Lee & Nowell, 2014). LeRoux and Write (2010) argue that "performance data can provide managers with valuable insights into organizational strengths and weaknesses, thereby equipping managers with information to guide strategic decision making" (p.575).

On the other hand, organizations' main purposes might be institutional legitimacy (DiMaggio & Powell, 1983). According to the institutional theory approach, performance measurement is a symbolic practice to achieve the intuitional legitimacy from the perspective of organizational stakeholders such as funders, donors, credentialing providers, and potential clients (e.g., Roy & Seguin, 2000; Modell, 2001& 2009; Yang, 2009; Lee & Nowell, 2014). The significant results of institutional factors in my study support this idea. The funding from United Way, the other largest organizations' practices, and the accreditation were main factors which significantly influence the adoption of outcome measurement.

As I mentioned in the literature review section, "even if the change is only *pro forma*", the adoption of the outcome measures make the organizations look proactive, competent, and

responsive (Roy & Seguine, 2000; Bolman & Deal, 1991, p. 278). That is to say, only looking "proactive, competent, and responsive" might be the main purpose of outcome measurement of nonprofits. My survey results support this idea by finding that the biggest purposes of outcome information uses are complying with government requirements and reporting to funders (See Table 5.1). In addition, Lee and Nowell (2014) describe that "performance measurement information is frequently used as part of the public narrative of a nonprofit, helping the organization to tell its story in reports and other media" (p.20). In this sense, my research findings based on institutional theory indicate that performance measurement is shaped strongly by external pressures and motivation to get institutional legitimacy. With the perspective of institutional theory, my research findings provide insightful ideas to address the performance measurement practices of human service nonprofits.

Risk Aversion

Although risk aversion was not significant in the logistic regression model, the interaction term between risk aversion and professionalism indicates the negative effect of risk aversion on the adoption of outcome measurement. That is to say, more risk averse organizations are less likely to adopt outcome measurement. This finding provides important theoretical implications in that the risk aversion has rarely been considered in the outcome measurement studies of nonprofits before. First, agency theory assumes that agents tend to be more risk averse than principals. This can be extended by relaxing the assumption of risk aversion (Eisenhardt, 1989). I assumed that there might be a variation in the risk attitude of

nonprofits. My research findings provide empirical support for this assumption by examining how the risk attitude varies across nonprofit agencies (See Table 4.21, 4.22, and 4.23).

In addition, agency theory is concerned about redistributing risk through various mechanisms (Van Slyke, 2006). The risk implication is one of the main contributions of agency theory as well as one of the main differences between agency theory and the other organizational theories (Eisenhardt, 1989). My research findings extend the performance measurement study by considering the risk implication of agency theory. By examining how the risk attitude of nonprofit organizations influences the adoption of outcome measurement through the performance based contracts, my study established the importance of considering the risk attitude in understanding the outcome measurement practice of nonprofits.

Organizational capacity

Human resource capacity was a significant factor in the logistic regression model. This indicates how important the human resources are in the process of the adoption of outcome measurement. This finding suggests that nonprofit organizations need to invest their resources to develop human resource capacity. Nonprofit agencies should assign more budgets to employees' education and training. Nonprofits heavily rely on volunteers, who might be less motivated for the training and learning, compared to the paid employees in for-profit firms (Corder, 2001). Hence, nonprofits should design education and training programs to develop the human resource capacity as well as self-motivation of their volunteers. My research findings indicate that 85% of the respondents of my survey have received training

on outcome measurement. Since my survey respondents were mainly executive directors and presidents of the nonprofits, this finding indicates that the leaders of human service nonprofits have an interest in the outcome measurement. Therefore, human service nonprofits can potentially develop their human resource capacity through the support from organizational leaders.

I examined the extent to which nonprofit practitioners use the outcome information (Table 5.1). However, more use of outcome information does not guarantee better use of outcome information. That is to say, not only the quantity of outcome information use but also the quality of outcome information use needs to be examined. In the training and education programs of outcome measurement, it should be considered how to appropriately use the outcome information in the implementation process of outcome measurement.

In addition to the human resource capacity, I found that political capacity is positively related to the adoption of outcome measurement. The importance of political capacity for adopting new practices has been discussed in the literature (e.g., Daniels Clark-Daniels, 2000; Borins, 2002; Anguelovski & Carmin, 2011; Berman & Wang, 2010; Héritier, 2003). Political capacity is essential for adopting the outcome measurement because it helps assigns budgets for the new practice, legitimizes it, establishes a consensus, and develops strategies to convince stakeholders and obtain their support (Berman & Wang, 2010; Héritier, 2003). However, literature argues that nonprofits have legal restrictions on political capacity development, compared to for-profit firms. According to Frumkin and Andre-Clark (2000), nonprofits have "competitive disadvantage in comparison to business when it comes to shaping the political environment" (p.150). Human service nonprofits' contributions for

lobbying are not tax deductible. Also, if nonprofits lobby more than permitted, they may lose the tax-exempt status (Frumkin & Andre-Clark, 2000). Although ability to lobby and political capacity are not perfectly the same concept, the restrictions on lobbying may negatively influence the nonprofits' political capacity development. My findings indicate the importance of political capacity in the process of outcome measurement adoption. Hence, this study suggests that human service nonprofits need to construct strategies to overcome the competitive disadvantages and develop their political capacity.

On the other hand, my research findings also indicate that the other dimensions of organizational capacity are not main factors leading to the adoption of outcome measurement. As MacIndoe and Barman (2012) found that more than 50% of organizations measured the outcomes of their programs and services without specific budgets for the outcome measurement, I found that the financial capacity did not significantly influence the adoption of outcome measurement. In addition, the other dimensions of organizational capacity – technical capacity, administrative capacity, and network capacity were not significant as well. This might indicate that the adoption of outcome measurement is influenced by external institutional forces rather than organizations' internal capacities. In this sense, organizational capacity was less predictive in the adoption of outcome measurement, compared to the institutional factors. The implementation of outcome measurement, however, might be different. I will elaborate this idea in the future research section.

5.4 Limitations

While this dissertation provides meaningful implications to the performance measurement of nonprofits, several limitations of this study should be considered. The first limitation of this study is related to the generalizability of my research findings. In this study, I excluded small nonprofits with total expenses below $500,000 because small nonprofits might not set up the performance measurement system. While this filtering makes sense in that the focus of my study is mid and large sized nonprofits, the exclusion of small nonprofits limits the generalizability of my research findings to all sizes of nonprofit organizations. Also, as I mentioned in the result section (See Table 4.5 and 4.6), for the total expenses of P43 organizations, the respondent group's mean "1,800,142.41" falls outside the 95% confidence interval of true population mean. This means that there is a statistically significant difference between the total expenses of P43 respondent group and population group with a 95% confidence level, which indicates that the results of this study may not be generalizable to all P43 organizations with more than $500,000 in total expenses.

The second limitation of this study stems from the response rate of the survey. While considerable efforts were made to increase the response rate, I was not able to get a higher response rate due to the long questionnaire and the limited period of data collection. According to Millar and Dillman (2011), offering a choice of responding by either mail or Web and using email augmentation approach has the highest response rate: average 46.5%. Next, the average response rate of a mail survey is 43.9%. The survey using only email has the lowest response rate: 20.5%. This result indicates that the response rate of my survey (31.8%) using only email is not a low level, compared to the average level. Nevertheless, I

could have attained a higher response rate with the combination of email and mail. The limited number of cases caused some difficulties in the data analysis. As I mentioned in the result section, I could not test all independent variables originally planned due to the limited number of cases. In addition, my study found that only two dimensions of organizational capacity - human resource capacity and political capacity- were significant in the logistic regression model. Including more interaction terms between the dimensions of organizational capacity and the other factors (e.g., professionalism and risk aversion) in the model might explain the effect of organizational capacity on the adoption of outcome measurement. However, I could not include more interaction terms in my research model because of the limited size of the dataset.

The third limitation of this study is related to the analytical method of logistic regression model. I originally planned the ordinal regression, but there was a great percentage of cells with zero frequency which makes the validity of the model uncertain. Thus, after I checked the assumptions of several regression models, I finally decided to use the binary logistic regression. To use this regression model, I used a dichotomous dependent variable instead of one with five categories. In this process, the variation in the dependent variable decreased and I was not able to fully use the information contained in the distribution of the original dependent variable.

The last limitation of the study has to do with the internal reliability of some survey items. Some of my variables are measured by multiple survey items. To generate overall scales of the variables, I checked the Cronbach's alpha level first. As I mentioned in the result section, Cronbach's alpha of the risk aversion, mimetic isomorphism, uncertainty, and

political support were .582, .386, .766, and .709 respectively. The Cronbach's alpha of mimetic isomorphism was unacceptable, while the Cronbach's alpha of uncertainty and political support showed an adequate level. Thus, I treated the two variables of mimetic isomorphism separately. For the risk aversion, the Cronbach's alpha was slightly less than .6 which is a cut off in the exploratory research (Garson, 2012g). However, I averaged the three items of risk aversion for the following reasons: 1) the Cronbach's alpha level was close to the cut off; 2) deleting any item among the three did not increase the Cronbach's alpha statistics; and 3) Cronbach's alpha of the three risk aversion items was .608 in the raw dataset before I dropped out 4 cases which failed to pass my data quality examination. Also, another measure of reliability, Armor's reliability Theta, indicated the reliability of risk aversion has an acceptable level at 0.61. In spite of this justification, the reliability of integrated variables might be the limitation of this study. Some survey items in my questionnaire were designed to measure the same concept, accordingly, high Cronbach's alpha level of the items was expected. The low Cronbach's alpha level might be a limitation of the survey instrument used in this study. In addition, it indicates that more reliable measures of risk aversion should be developed in future research.

5.5 Future Research

This study suggests important areas of future research. First, the study of performance measurement requires two stages: adoption and implementation (Julnes & Holzer, 2001). In this study, I focused on the adoption of the outcome measurement and examined the actual use of outcome information, that is, implementation of outcome measurement only at a

preliminary level. More in-depth discussion on the use of outcome information is needed. For instance, future researchers may examine how practitioners perceive the adoption and implementation of outcome measurement differently, to what extent nonprofit practitioners use the outcome information, and what factors (e.g., institutional factors, organizational characteristics, and environmental factors) influence the use of outcome information. In Table 5.3, 5.4, and 5.5, I examined the correlation between the uses of outcome information and risk aversion, institutional factors, and organizational capacity. Although it is a preliminary level, it provides a starting point for future research on the use of outcome information. By examining the use of outcome information, the substantive usefulness of outcome measurement can be discussed (Propper & Wilson, 2003; Van Dooren, 2008).

My research found that organizational capacities are not the main factors for the adoption of outcome measurement. Institutional theory literature, however, describes that there is a gap between adoption and implementation of organizational practices with a concept of "loose coupling between legitimated external practices and internal organizational behavior" (Meyer & Rowan, 1977; DiMaggio & Powell, 1983, p.155). This indicates that the substantive implementation (internal organizational behavior) might be different from the adoption (external practices) of outcome measurement. Unlike the adoption of outcome measurement, nonprofits may substantively need the organizational capacity when they implement the outcome measurement and use the outcome information. This is one of the fruitful fields of future study. My preliminary study of the outcome information use supports this idea. I found that one of the main purposes of outcome information use is outreach and public relations: Professional/discretionary uses of outcome information. Table 5.5 indicates

a positive correlation between professional/discretionary uses of outcome information and human resource capacity, technical capacity, political capacity, and network capacity. Although the correlation matrix does not suggest a causal relationship between these variables, nonprofits may need the four organizational capacities mentioned above in the process of analyzing the outcome information, improving management decisions and strategic planning, applying the outcome information into the programs and services, and disseminating the outcome data to the public.

In addition, I focused on human service nonprofit organizations in my study. Given that nonprofit, public, and for-profit organizations have different characteristics, comparing the different organizational sectors would provide meaningful findings. Corder (2001) examines the similarities and dissimilarities of nonprofit and public human service agencies in the acquisition of new technology. Building on Corder (2001) and my study, future researchers may discuss how the characteristics of nonprofit and public agencies (e.g., funding type, workforce, centralization, organizational culture, and attitude of personnel) differently shape the practice of outcome measurement.

Third, the costs and benefits of outcome measurement is one meaningful area of future research. How practitioners perceive the costs and benefits of outcome measurement would shape their performance measurement practice. First, reviewing the literature to examine the costs and benefits of outcome measurement, and then asking nonprofit practitioners about the costs and benefits by survey would indicate how scholars and practitioners differently perceive the costs and benefits of outcome measures. Different organizational characteristics (e.g., industry, organizational age, organizational capacity, and

funding streams) may lead to different perceptions of the costs and benefits of outcome measurement. As I mentioned above, I used 5 outcome indicators for P32 and P43 organizations and 8 indicators for P46 organizations to measure the extent of outcome measurement. Each outcome measure might have different costs and benefits from one another. Some outcome indicators might be unrealistic, considering the costs of measurement. For some outcome measures, the costs of measurement might outweigh the benefits. Studying these questions would explain the different outcome measurement practices across nonprofit agencies.

Fourth, as I mentioned in the discussion section, four out of six dimensions of organizational capacity - financial, infrastructure, administrative, and network capacity - were not supported by my research findings. One explanation for this contradictory finding is that I excluded small nonprofits with expenses below $500,000. Only medium and large nonprofits participated in my survey and the majority of survey respondents answered that they have adequate organizational capacity. The small variation in the organizational capacity might be the reason why the four dimensions of organizational capacity were not significant in my study. Future research considering the full range of nonprofits may produce different results.

Lastly, as my pilot study indicated, human service nonprofits have difficulties to measure their outcomes because many of their clients are in crisis and they are not willing to participate in the outcome measurement without incentives. An interviewee said that the response rate of their outcome measurement survey was only 1%. My dissertation, which relied on quantitative methods, did not elaborate on the substantive difficulties of outcome

measurement in human service nonprofits. Qualitative research based on in-depth interviews may identify what difficulties of outcome measurement the human service nonprofits have and how their organizational characteristics or environmental factors are related to them.

In conclusion, as mentioned in the beginning, performance measurement has received an increasing attention in the nonprofit sector as well as the public sector. The outcome measurement, however, is relatively slow and expensive (Curristine, Lonti, & Joumard, 2007; Moore, 2003). Thus, the performance measurement of the nonprofits has focused on output rather than outcome for decades. However, the absence of outcome measurement may lead to goal displacement. This study opens a door to fruitful future research questions which can improve our knowledge of the performance measurement practice of nonprofits. By better understanding the performance measurement of nonprofits, the field will be better positioned to advance the practice as well as the theory of performance measurement.

REFERENCES

AAMFT. (2013). AAMFT Therapy Topics. Retrieved from http://www.aamft.org/imis15/content/consumer_updates/Marriage_and_Family_Therapists.aspx

AAMFT. (2014). About AAMFT. Retrieved from http://www.aamft.org/imis15/content/about_aamft/AAMFT.aspx

Abdellaoui, M., Driouchi, A., & l'Haridon, O. (2011). Risk aversion elicitation: Reconciling tractability and bias minimization. *Theory and Decision, 71*(1), 63-80.

Abrahamson, E., & Bartner, L. R. (1990). When do bandwagon diffusions roll? how far do they go? and when do they roll backwards?: A computer simulation. Paper presented at the *Academy of Management Proceedings*, , *1990*(1) 155-159.

Al-Senaidi, S., Lin, L., & Poirot, J. (2009). Barriers to adopting technology for teaching and learning in oman. *Computers & Education, 53*(3), 575-590.

Amaratunga, D., & Baldry, D. (2002). Moving from performance measurement to performance management. *Facilities, 20*(5/6), 217-223.

Andrews, R. (2008). Perceived environmental uncertainty in public organizations: An empirical exploration. *Public Performance & Management Review, 32*(1), 25-50.

Andrews, R., & Boyne, G. A. (2010). Capacity, leadership, and organizational performance: Testing the black box model of public management. *Public Administration Review, 70*(3), 443-454.

Anguelovski, I., & Carmin, J. (2011). Something borrowed, everything new: Innovation and institutionalization in urban climate governance. *Current Opinion in Environmental Sustainability, 3*(3), 169-175.

Ansari, S. M., Fiss, P. C., & Zajac, E. J. (2010). Made to fit: How practices vary as they diffuse. *Academy of Management Review, 35*(1), 67-92.

Aristigueta, M. P., & Foote, L. R. (2009). The use of performance contracts for delivery of social services in the united states. *Theoretical and Applied Economics, 3*(3), 3.

Arndt, M., & Bigelow, B. (2000). Presenting structural innovation in an institutional environment: Hospitals' use of impression management. *Administrative Science Quarterly, 45*(3), 494-522.

Austin, M. J., Regan, K., Samples, M. W., Schwartz, S. L., & Carnochan, S. (2011). Building managerial and organizational capacity in nonprofit human service organizations through a leadership development program. *Administration in Social Work, 35*(3), 258-281.

Barman, E., & MacIndoe, H. (2012). Institutional pressures and organizational capacity: The case of outcome Measurement1. Paper presented at the *Sociological Forum, , 27*(1) 70-93.

Barreto, I., & Baden-Fuller, C. (2006). To conform or to perform? mimetic behaviour, Legitimacy-Based groups and performance consequences*. *Journal of Management Studies, 43*(7), 1559-1581.

Bellante, D., & Link, A. N. (1980). Are public sector workers more risk averse than private sector workers. *Indus.& Lab.Rel.Rev., 34*, 408.

Benders, J., Batenburg, R., & van der Blonk, H. (2006). Sticking to standards; technical and other isomorphic pressures in deploying ERP-systems. *Information & Management, 43*(2), 194-203.

Benjamin, L. M. (2010). Funders as principals. *Nonprofit Management and Leadership, 20*(4), 383-403.

Bergen, M., Dutta, S., & Walker Jr, O. C. (1992). Agency relationships in marketing: A review of the implications and applications of agency and related theories. *The Journal of Marketing, ,* 1-24.

Berman, E., & Wang, X. (2000). Performance measurement in US counties: Capacity for reform. *Public Administration Review, 60*(5), 409-420.

Boettcher, W. A. (2004). Military intervention decisions regarding humanitarian crises framing induced risk behavior. *Journal of Conflict Resolution, 48*(3), 331-355.

Bohte, J., & Meier, K. J. (2000). Goal displacement: Assessing the motivation for organizational cheating. *Public Administration Review, 60*(2), 173-182.

Bolman, L. G., & Deal, T. E. (1991). *Reframing organizations: Artistry, choice and leadership* Jossey-Bass.

Borins, S. (2002). Leadership and innovation in the public sector. *Leadership & Organization Development Journal, 23*(8), 467-476.

Boston University School of Public Health. (2014). The Binomial Distribution: A Probability Model for a Discrete Outcome. Retrieved from http://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_probability/BS704_Probability7.html

Botcheva, L., White, C. R., & Huffman, L. C. (2002). Learning culture and outcomes measurement practices in community agencies. *American Journal of Evaluation, 23*(4), 421-434.

Bowman, W. (2011). Financial capacity and sustainability of ordinary nonprofits.*Nonprofit Management and Leadership*, *22*(1), 37-51.

Bowornwathana, B. (2000). Governance reform in thailand: Questionable assumptions, uncertain outcomes. *Governance, 13*(3), 393-408.

Bozeman, B., & Kingsley, G. (1998). Risk culture in public and private organizations. *Public Administration Review, ,* 109-118.

Brouthers, L. E., O'Donnell, E., & Hadjimarcou, J. (2005). Generic product strategies for emerging market exports into triad nation markets: A mimetic isomorphism approach. *Journal of Management Studies, 42*(1), 225-245.

Buckmaster, N. (1999). Associations between outcome measurement, accountability and learning for non-profit organisations. *International Journal of Public Sector Management, 12*(2), 186-197.

Budding, G. T. (2004). Accountability, environmental uncertainty and government performance: Evidence from dutch municipalities. *Management Accounting Research, 15*(3), 285-304.

Cairns, B., Harris, M., Hutchison, R., & Tricker, M. (2005). Improving performance? the adoption and implementation of quality systems in UK nonprofits. *Nonprofit Management and Leadership, 16*(2), 135-151.

Campbell, D. A., Lambright, K. T., & Bronstein, L. R. (2012). In the eyes of the beholders. *Public Performance & Management Review, 36*(1), 7-30.

Campion, M. A., Medsker, G. J., & Higgs, A. C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel psychology, 46*(4), 823-847.

Cardinal McCloskey Community Services. (2013). Family Foster Care. Retrieved from http://www.cmcs.org/our-services/foster-care-services/family-foster-care

Carman, J. G. (2009). Nonprofits, funders, and evaluation accountability in action. *The American Review of Public Administration, 39*(4), 374-390.

Carman, J. G., & Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. *New Directions for Evaluation, 2008*(119), 51-71.

Carrilio, T. E., Packard, T., & Clapp, J. D. (2004). Nothing in—nothing out: Barriers to the use of performance data in social service programs. *Administration in Social Work, 27*(4), 61-75.

Carson, E. D. (2000). On foundations and outcome evaluation. *Nonprofit and Voluntary Sector Quarterly, 29*(3), 479-481.

Carter Jr, F. J., Jambulingam, T., Gupta, V. K., & Melone, N. (2001). Technological innovations: A framework for communicating diffusion effects. *Information & Management, 38*(5), 277-287.

Carter, L., & Weerakkody, V. (2008). E-government adoption: A cultural comparison. *Information Systems Frontiers, 10*(4), 473-482.

Celly, K. S., & Frazier, G. L. (1996). Outcome-based and behavior-based coordination efforts in channel relationships. *Journal of marketing research*, 200-210.

Center for Human Development. (2013). Family Support and Stabilization. Retried from http://www.chd.org/index.php/family-support-and-stabilization.html

CFSA. (2012). FY 2012 Performance Plan.

CFSA. (2014). Four Pillars. Retrieved from http://cfsa.dc.gov/page/four-pillars

Chaskin, R. J. (2001). Building community capacity a definitional framework and case studies from a comprehensive community initiative. *Urban Affairs Review, 36*(3), 291-323.

Christensen, R. K., & Gazley, B. (2008). Capacity for public administration: Analysis of meaning and measurement. *Public Administration and Development, 28*(4), 265-279.

Cohen,W.,& Levinthal, D. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly, 35*, 128-152.

Collier, P. M. (2001). The power of accounting: A field study of local financial management in a police force. *Management Accounting Research, 12*(4), 465-486.

Copeland, J. M., Taylor, W. J., & Dean, S. G. (2008). Factors influencing the use of outcome measures for patients with low back pain: A survey of new zealand physical therapists. *Physical Therapy, 88*(12), 1492-1505.

Corder, K. (2001). Acquiring new technology comparing nonprofit and public sector agencies. *Administration & Society, 33*(2), 194-219.

Cosner, S. (2009). Building organizational capacity through trust. *Educational Administration Quarterly, 45*(2), 248-291.

Covin, J. G., & Slevin, D. P. (1991). A conceptual model of entrepreneurship as firm behavior. *Entrepreneurship theory and practice*, *16*(1), 7-25.

Croon, M. A., & van Veldhoven, M. J. (2007). Predicting group-level outcome variables from variables measured at the individual level: a latent variable multilevel model. *Psychological methods*, *12*(1), 45.

Curristine, T., Lonti, Z., & Joumard, I. (2007). Improving public sector efficiency: Challenges and opportunities. *OECD Journal on Budgeting, 7*(1), 161.

Damanpour, F., & Schneider, M. (2009). Characteristics of innovation and innovation adoption in public organizations: Assessing the role of managers. *Journal of Public Administration Research and Theory, 19*(3), 495-522.

Daniels, R. S., & Clark-Daniels, C. L. (2000). *Transforming government: The renewal and revitalization of the federal emergency management agency* PricewaterhouseCoopers Endowment for the Business of Government.

Davis, J. H., Schoorman, F. D., & Donaldson, L. (1997). Toward a stewardship theory of management. *Academy of Management Review, 22*(1), 20-47.

Dearing, J. W., & Kreuter, M. W. (2010). Designing for diffusion: How can we increase uptake of cancer communication innovations? *Patient Education and Counseling, 81*, S100-S110.

Deeds, D. L., & Rothaermel, F. T. (2003). Honeymoons and liabilities: The relationship between age and performance in research and development alliances. *Journal of Product Innovation Management, 20*(6), 468-484.

Delmas, M. A. (2002). The diffusion of environmental management standards in europe and in the united states: An institutional perspective. *Policy Sciences, 35*(1), 91-119.

Dicke, L. A. (2002). Ensuring accountability in human services contracting can stewardship theory fill the bill? *The American Review of Public Administration, 32*(4), 455-470.

Dicke, L. A., & Ott, J. S. (1999). Public agency accountability in human services contracting. *Public Productivity & Management Review, 22*(4), 502-516.

DiMaggio, P., & Powell, W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. American Sociological Review, 48, 147-160.

Druckman, J. N., & McDermott, R. (2008). Emotion and the framing of risky choice. *Political Behavior, 30*(3), 297-321.

Ehrlich, S., & Maestas, C. (2010). Risk orientation, risk exposure, and policy opinions: The case of free trade. *Political Psychology, 31*(5), 657-684.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review, 14*(1), 57-74.

Eisinger, P. (2002). Organizational capacity and organizational effectiveness among street-level food assistance programs. *Nonprofit and Voluntary Sector Quarterly, 31*(1), 115-130.

Ellis, S., & Shpielberg, N. (2003). Organizational learning mechanisms and managers' perceived uncertainty. *Human Relations, 56*(10), 1233-1254.

Family & Children's Center. (2013). Family & Children's Center at a glance. Retried from http://www.fcconline.org/AboutUs.aspx

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2), 117-140.

Fichman, M., & Levinthal, D. A. (1991). Honeymoons and the liability of adolescence: A new perspective on duration dependence in social and organizational relationships. *Academy of Management Review, 16*(2), 442-468.

Fiegenbaum, A., & Thomas, H. (1988). Attitudes toward risk and the risk–return paradox: Prospect theory explanations. *Academy of Management Journal, 31*(1), 85-106.

Fine, A. H., Thayer, C. E., & Coghlan, A. (2000). Program evaluation practice in the nonprofit sector. *Nonprofit Management and Leadership, 10*(3), 331-339.

Fligstein, N. (1985). The spread of the multidivisional form among large firms, 1919-1979. *American Sociological Review, ,* 377-391.

Fogarty, T. J., & Rogers, R. K. (2005). Financial analysts' reports: An extended institutional theory evaluation. *Accounting, Organizations and Society, 30*(4), 331-356.

Folz, D. H., Abdelrazek, R., & Chung, Y. (2009). The adoption, use, and impacts of performance measures in medium-size cities. *Public Performance & Management Review, 33*(1), 63-87.

Franklin, A. L. (2002). An examination of the impact of budget reform on arizona and oklahoma appropriations. *Public Budgeting & Finance, 22*(3), 26-45.

Freeman, J., Carroll, G. R., & Hannan, M. T. (1983). The liability of newness: Age dependence in organizational death rates. *American Sociological Review, ,* 692-710.

Frumkin, P. (2005). Trouble in foundation land: Looking back. *Looking Ahead, ,* 23.

Frumkin, P., & Andre-Clark, A. (2000). When missions, markets, and politics collide: Values and strategy in the nonprofit human services. *Nonprofit and voluntary sector quarterly, 29*(suppl 1), 141-163.

Garson,G.David. (2012a). Reliability Analysis. Blue Book Series. Asheboro, NC: Statistical Associates Publishing.

Garson,G.David. (2012b). Data Imputation for Missing Values. Blue Book Series. Asheboro, NC: Statistical Associates Publishing.

Garson,G.David. (2012c). Ordinal Regression. Blue Book Series. Asheboro, NC: Statistical Associates Publishing.

Garson,G.David. (2012d). Generalized Linear Models and Generalized Estimating Equations. Blue Book Series. Asheboro, NC: Statistical Associates Publishing.

Garson,G.David. (2012e). Logistic Regression. Blue Book Series. Asheboro, NC: Statistical Associates Publishing.

Garson,G.David. (2012f). Correlation. Blue Book Series. Asheboro, NC: Statistical Associates Publishing.

Garson,G.David. (2012g). Factor Analysis. Blue Book Series. Asheboro, NC: Statistical Associates Publishing.

Garud, R., & Karnoe, P. (2013). *Path dependence and creation* Psychology Press.

Gavin, M. B., & Hofmann, D. A. (2002). Using hierarchical linear modeling to investigate the moderating influence of leadership climate. *The Leadership Quarterly*, *13*(1), 15-33.

Gazley, B. (2010). Linking collaborative capacity to performance measurement in Government—Nonprofit partnerships. *Nonprofit and Voluntary Sector Quarterly, 39*(4), 653-673.

Goldenberg, J., Han, S., Lehmann, D., & Hong, J. (2009). The role of hubs in the adoption processes. *Journal of Marketing, 73*(2)

Greve, H. R. (1996, August). Performance, Aspirations and Risky Organizational Change. In Academy of Management Proceedings (Vol. 1996, No. 1, pp. 224-228). Academy of Management.

Greve, H. R. (2000). Marketing niche entry decisions: Competition, learning, and strategy in tokyo banking, 1894–1936. *Academy of Management Journal, 43*(5), 816-836.

Griffith, D. M., Allen, J. O., DeLoney, E. H., Robinson, K., Lewis, E. Y., Campbell, B., . . . Reischl, T. (2010). Community-based organizational capacity building as a strategy to reduce racial health disparities. *The Journal of Primary Prevention, 31*(1-2), 31-39.

Grund, C., & Sliwka, D. (2010). Evidence on performance pay and risk aversion. *Economics Letters, 106*(1), 8-11.

Gujarati, D., & Porter, D. Essentials of Econometrics, 2010.

Guler, I., Guillén, M. F., & Macpherson, J. M. (2002). Global competition, institutions, and the diffusion of organizational practices: The international spread of ISO 9000 quality certificates. *Administrative Science Quarterly, 47*(2), 207-232.

Guo, C., & Acar, M. (2005). Understanding collaboration among nonprofit organizations: Combining resource dependency, institutional, and network perspectives. *Nonprofit and Voluntary Sector Quarterly, 34*(3), 340-361.

Hackler, D., & Saxton, G. D. (2007). The strategic use of information technology by nonprofit organizations: Increasing capacity and untapped potential. *Public Administration Review, 67*(3), 474-487.

Haigh, R., Tennant, A., Biering-Sorensen, F., Grimby, G., Marincek, C., Phillips, S., . . . Thonnard, J. (2001). The use of outcome measures in physical medicine and rehabilitation within europe. *Journal of Rehabilitation Medicine, 33*(6), 273-278.

Hatfield, D. R., & Ogles, B. M. (2004). The use of outcome measures by psychologists in clinical practice. *Professional Psychology: Research and Practice, 35*(5), 485.

Haunschild, P. R. (1994). How much is that company worth?: Interorganizational relationships, uncertainty, and acquisition premiums. *Administrative Science Quarterly*, , 391-411.

Haveman, H. A. (1993). Follow the leader: Mimetic isomorphism and entry into new markets. *Administrative Science Quarterly*, , 593-627.

Hendricks, M., Plantz, M. C., & Pritchard, K. J. (2008). Measuring outcomes of United Way–funded programs: Expectations and reality. *New Directions for Evaluation, 2008*(119), 13-35.

Hendry, J. (2002). The principal's other problems: Honest incompetence and the specification of objectives. *Academy of Management Review, 27*(1), 98-113.

Héritier, A. (2003). New modes of governance in europe: Increasing political capacity and policy effectiveness. *The State of the European Union, 6*, 105-126.

Herman, R. D., & Renz, D. O. (2008). Advancing nonprofit organizational effectiveness research and theory: Nine theses. *Nonprofit Management and Leadership, 18*(4), 399-415.

Hildebrandt, S. A., & Eom, M. (2011). Teacher professionalization: Motivational factors and the influence of age. *Teaching and Teacher Education, 27*(2), 416-423.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review, 92*(5), 1644-1655.

Hoque, Z. (2004). A contingency model of the association between strategy, environmental uncertainty and performance measurement: Impact on organizational performance. *International Business Review, 13*(4), 485-502.

Hoque, Z., & James, W. (2000). Linking balanced scorecard measures to size and market factors: Impact on organizational performance. *Journal of Management Accounting Research, 12*(1), 1-17.

Horsch, K. (1996). Results-based accountability systems: Opportunities and challenges. *The Evaluation Exchange II (1)* *Http://www.Gse.Harvard.edu/hfrp/eval/issue3/theory1.Html*,

Hsu, C., Lu, H., & Hsu, H. (2007). Adoption of the mobile internet: An empirical study of multimedia message service (MMS). *Omega, 35*(6), 715-726.

Hussain, M. M., & Hoque, Z. (2002). Understanding non-financial performance measurement practices in japanese banks: A new institutional sociology perspective. *Accounting, Auditing & Accountability Journal, 15*(2), 162-183.

Hwang, H., & Powell, W. W. (2009). The rationalization of charity: The influences of professionalism in the nonprofit sector. *Administrative Science Quarterly, 54*(2), 268-298.

Ingraham, P.W., Joyce P.G., & Donahue A.K. (2003). Government Performance: Why Management Matters. Johns Hopkins University Press: Baltimore.

Institute for Digital Research and Education. (2014). SPSS Data Analysis Examples Poisson Regression. Retrieved from http://www.ats.ucla.edu/stat/spss/dae/poissonreg.htm

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*(4), 305-360.

Jiang, B., & Qureshi, A. (2006). Research on outsourcing results: Current literature and future opportunities. *Management Decision, 44*(1), 44-55.

Joffres, C., Heath, S., Farquharson, J., Barkhouse, K., Latter, C., & MacLean, D. R. (2004). Facilitators and challenges to organizational capacity building in heart health promotion. *Qualitative Health Research, 14*(1), 39-60.

Jones, M. L. (2001). Sustainable organizational capacity building: Is organizational learning a key? *International Journal of Human Resource Management, 12*(1), 91-98.

Judge, W. Q., & Elenkov, D. (2005). Organizational capacity for change and environmental performance: An empirical assessment of bulgarian firms. *Journal of Business Research, 58*(7), 893-901.

Julian, D. A., & Kombarakaran, F. (2006). Assessment of quality of outcomes within a local united way organization: Implications for sustaining system level change. *American Journal of Community Psychology, 38*(3-4), 175-181.

Julian, D., & Clapp, J. (2000). Outcomes based funding and the logic model: Mechanisms for coordinating human services at the local level. *Evaluation and Program Planning, 23*(2), 231-240.

Julnes, P. d. L., & Holzer, M. (2001). Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation. *Public Administration Review, 61*(6), 693-708.

Jun, K., & Weare, C. (2011). Institutional motivations in the adoption of innovations: The case of e-government. *Journal of Public Administration Research and Theory, 21*(3), 495-519.

Kam, C. D., & Simas, E. N. (2010). Risk orientations and policy frames. *Journal of Politics, 72*(2), 381-396.

Kam, C. D., & Simas, E. N. (2012). Risk attitudes, candidate characteristics, and vote choice. *Public Opinion Quarterly, 76*(4), 747-760.

Kaplan, R. S. (2001). Strategic performance measurement and management in nonprofit organizations. *Nonprofit Management and Leadership, 11*(3), 353-370.

Kaplan, R. S., & Norton, D. P. (1996). The Balanced Scorecard: Translating Strategy into Action. Boston: Harvard Business School Press.

Kennedy, M. T., & Fiss, P. C. (2009). Institutionalization, framing, and diffusion: The logic of TQM adoption and implementation decisions among US hospitals. *Academy of Management Journal, 52*(5), 897-918.

Ketokivi, M. A., & Schroeder, R. G. (2004). Strategic, structural contingency and institutional explanations in the adoption of innovative manufacturing practices. *Journal of Operations Management, 22*(1), 63-89.

Kim, S. (2005). Balancing competing accountability requirements: Challenges in performance improvement of the nonprofit human services agency. *Public Performance & Management Review, 29*(2), 145-163.

Knudsen, H. K., & Roman, P. M. (2004). Modeling the use of innovations in private treatment organizations: The role of absorptive capacity. *Journal of Substance Abuse Treatment, 26*(1), 51-59.

Ko, E., Kim, S. H., Kim, M., & Woo, J. Y. (2008). Organizational characteristics and the CRM adoption process. *Journal of Business Research, 61*(1), 65-74.

Kondra, A. Z., & Hinings, C. R. (1998). Organizational diversity and change in institutional theory. *Organization studies, 19*(5), 743-767.

Krause, G. A. (2003). Coping with uncertainty: Analyzing risk propensities of SEC budgetary decisions, 1949-97. *American Political Science Review, 97*(1), 171-188.

Krause, R. M. (2011). Policy innovation, intergovernmental relations, and the adoption of climate protection initiatives by US cities. *Journal of Urban Affairs, 33*(1), 45-60.

Kumar, N., Scheer, L. K., & Steenkamp, J. B. E. (1995). The effects of supplier fairness on vulnerable resellers. *Journal of marketing research*, 54-65.

Kwon, M., Berry, F. S., & Feiock, R. C. (2009). Understanding the adoption and timing of economic development strategies in US cities using innovation and institutional analysis. *Journal of Public Administration Research and Theory*, *19*(4), 967-988.

Lambright, K. T. (2009). Agency theory and beyond: Contracted providers' motivations to properly use service monitoring tools. *Journal of Public Administration Research and Theory*, *19*(2), 207-227.

Lampkin, L. M., & Boris, E. T. (2002). Nonprofit organization data what we have and what we need. *American Behavioral Scientist*, *45*(11), 1675-1715.

LaVelle, J. M., & Donaldson, S. I. (2010). University-based evaluation training programs in the united states 1980—2008: An empirical examination. *American Journal of Evaluation*, *31*(1), 9-23.

Lee, C & Nowell, B. (in press). A Framework for Assessing the Performance of Nonprofit Organizations. *American Journal of Evaluation.*

Lee, T. (2004). Nurses' adoption of technology: Application of rogers' innovation-diffusion model. *Applied Nursing Research*, *17*(4), 231-238.

LeRoux, K., & Wright, N. S. (2010). Does performance measurement improve strategic decision making? findings from a national survey of nonprofit social service agencies. *Nonprofit and Voluntary Sector Quarterly, 39*(4), 571-587.

Leyden, D. P., & Link, A. N. (1993). Privatization, bureaucracy, and risk aversion. *Public Choice, 76*(3), 199-213.

Li, H., & Barry, J. (2012). StatNews #84: Interpreting Interactions in Logistic Regression. Retrieved from http://www.cscu.cornell.edu/news/statnews/stnews84.pdf

Lindgren, L. (2001). The non-profit sector meets the performance-management movement A programme-theory approach. *Evaluation, 7*(3), 285-303.

Logan, M. S. (2000). Using agency theory to design successful outsourcing relationships. *International Journal of Logistics Management, the, 11*(2), 21-32.

MacCrimmon, K. R., & Wehrung, D. (1988). *Taking risks* Simon and Schuster.

MacIndoe, H., & Barman, E. (2012). How organizational stakeholders shape performance measurement in nonprofits: Exploring a multidimensional measure. *Nonprofit and Voluntary Sector Quarterly,*

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual review of psychology, 58*, 593.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*(4), 173-181.

Maggio, D.P.; POWELL, WW (1983).«The iron cage revisited: Institutional isomorphism and collective rationality in organizational frields». *American Sociological Review, 48*(2), 147-160.

Mallat, N., Rossi, M., Tuunainen, V. K., & Öörni, A. (2009). The impact of use context on mobile services acceptance: The case of mobile ticketing. *Information & Management, 46*(3), 190-195.

March, J. G., Olsen, J. P., & Christensen, S. (1976). Ambiguity and choice in organizations.

Martikke, S. (2008). Commissioning: Possible–Greater Manchester VCS Organizations' Experiences in Public Sector Commissioning. Manchester: Greater Manchester Centre for Voluntary Organization.

McGuire, M., & Silvia, C. (2010). The effect of problem severity, managerial and organizational capacity, and agency structure on intergovernmental collaboration: Evidence from local emergency management. *Public Administration Review, 70*(2), 279-288.

McLean, C., & Brouwer, C. (2010). The Effect of the Economy on the Nonprofit Sector: A June 2010 Survey. Retrieved from http://www.guidestar.org/ViewCmsFile.aspx?ContentID=2963.

Melkers, J., & Willoughby, K. (2005). Models of Performance-Measurement use in local governments: Understanding budgeting, communication, and lasting effects. *Public Administration Review, 65*(2), 180-190.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology, ,* 340-363.

Mikkelson, H. (2013). Universities and interpreter certification. *Translation & Interpreting, 5*(1), 66-78.

Millar, M. M., & Dillman, D. A. (2011). Improving response to web and mixed-mode surveys. *Public Opinion Quarterly*, nfr003.

Miller, A. S., & Hoffmann, J. P. (1995). Risk and religion: An explanation of gender differences in religiosity. *Journal for the Scientific Study of Religion, ,* 63-75.

Miller, G., Scarborough, J., Clark, C., Leonard, J. C., & Keziah, T. B. (2010). The need for national credentialing standards for addiction counselors. *Journal of Addictions & Offender Counseling, 30*(2), 50-57.

Millner, E. L., & Pratt, M. D. (1991). Risk aversion and rent-seeking: An extension and some experimental evidence. *Public Choice, 69*(1), 81-92.

Mizruchi, M. S., & Fein, L. C. (1999). The social construction of organizational knowledge: A study of the uses of coercive, mimetic, and normative isomorphism. *Administrative Science Quarterly, 44*(4), 653-683.

Modell, S. (2001). Performance measurement and institutional processes: a study of managerial responses to public sector reform. *Management Accounting Research, 12*(4), 437-464.

Modell, S. (2009). Institutional research on performance measurement and management in the public sector accounting literature: a review and assessment. *Financial Accountability & Management*, *25*(3), 277-303.

Moore, M. H. (2003). *The public value scorecard: A rejoinder and an alternative to" strategic performance measurement and management in non-profit organizations" by Robert Kaplan* Hauser Center for Nonprofit Organizations, Harvard University.

Moxham, C. (2009). Performance measurement: Examining the applicability of the existing body of knowledge to nonprofit organisations. *International Journal of Operations & Production Management, 29*(7), 740-763.

Moynihan, D. P. (2005). Why and how do state governments adopt and implement "Managing for results" reforms? *Journal of Public Administration Research and Theory, 15*(2), 219-243.

Moynihan, D. P., & Pandey, S. K. (2010). The big question for performance management: Why do managers use performance information? *Journal of Public Administration Research and Theory, 20*(4), 849-866.

Münster, J. (2006). Contests with an unknown number of contestants. *Public Choice, 129*(3-4), 353-368.

Naldi, L., Nordqvist, M., Sjöberg, K., & Wiklund, J. (2007). Entrepreneurial orientation, risk taking, and performance in family firms. *Family Business Review*, *20*(1), 33-47.

NCCS. (2006). Guide to Using NCCS Data.

NCCS. (2013). About NCCS. Retrieved from http://nccs.urban.org/about/index.cfm

Newcomer, K. E. (1997). Using performance measurement to improve programs. *New Directions for Evaluation, 1997*(75), 5-14.

Norušis, M. J. (2010). *PASW Statistics 18 Advanced Statistical Procedures Companion*. Prentice Hall.

O'Neill, B. (2001). Risk aversion in international relations theory. *International Studies Quarterly, 45*(4), 617-640.

Orton, J. D., & Weick, K. E. (1990). Loosely coupled systems: A reconceptualization. *Academy of Management Review, 15*(2), 203-223.

Packard, T. (2010). Staff perceptions of variables affecting performance in human service organizations. *Nonprofit and Voluntary Sector Quarterly, 39*(6), 971-990.

Park, Y., & Chen, J. V. (2007). Acceptance and adoption of the innovative use of smartphone. *Industrial Management & Data Systems, 107*(9), 1349-1365.

Paton, R. (2003). *Managing and measuring social enterprises* Sage.

Perrin, B. (1998). Effective use and misuse of performance measurement. *American Journal of Evaluation, 19*(3), 367-379.

Pettigrew, A. M., Woodman, R. W., & Cameron, K. S. (2001). Studying organizational change and development: Challenges for future research. *Academy of management journal, 44*(4), 697-713.

Plantz, M. C., Greenway, M. T., & Hendricks, M. (1997). Outcome measurement: Showing results in the nonprofit sector. *New Directions for Evaluation, 1997*(75), 15-30.

Poister, T. H. (2008). *Measuring performance in public and nonprofit organizations* Wiley. com.

Poister, T. H. (2010). The future of strategic planning in the public sector: linking strategic management and performance. *Public Administration Review, 70*(s1), s246-s254.

Poister, T. H., & Streib, G. (1999). Performance measurement in municipal government: Assessing the state of the practice. *Public Administration Review, ,* 325-335.

Poole, D. L., Davis, J. K., Reisman, J., & Nelson, J. E. (2001). Improving the quality of outcome evaluation plans. *Nonprofit Management and Leadership, 11*(4), 405-421.

Poole, D. L., Nelson, J., Carnahan, S., Chepenik, N. G., & Tubiak, C. (2000). Evaluating performance measurement systems in nonprofit agencies: The program accountability quality scale (PAQS). *American Journal of Evaluation, 21*(1), 15-26.

Powell, W. W., Gammal, D. L., & Simard, C. (2005). *Close encounters: The circulation and reception of managerial practices in the San Francisco bay area nonprofit community*

Propper, C., & Wilson, D. (2003). The use and usefulness of performance measures in the public sector. *Oxford Review of Economic Policy, 19*(2), 250-267.

Radaelli, C. M. (2000). Policy transfer in the european union: Institutional isomorphism as a source of legitimacy. *Governance, 13*(1), 25-43.

Ranger-Moore, J. (1997). Bigger may be better, but is older wiser? Organizational age and size in the New York life insurance industry. *American Sociological Review, 62*(6), 903-920.

Renzulli, L. A., & Roscigno, V. J. (2005). Charter school policy, implementation, and diffusion across the united states. *Sociology of Education, 78*(4), 344-366.

Roy, C., & Seguin, F. (2000). The institutionalization of efficiency-oriented approaches for public service improvement. *Public Productivity & Management Review, 23*(4), 449-468.

Rubin, D.B. (1987), Multiple Imputation for Nonresponse in Surveys, New York: John Wiley & Sons, Inc.

Russ-Eft, D., & Preskill, H. (2009). Evaluation in organizations: A systematic approach to enhancing learning, performance, and change (2nd ed.). Cambridge, MA: Perseus Press.

Safe Horizen. (2013). Domestic Violence & Abuse: Shelters. Retrieved from http://www.safehorizon.org/index/what-we-do-2/domestic-violence--abuse-53/domestic-violence--abuse-shelters-340.html

Saigal, A. (2008). A Study on the Impact of Information Technology Use on Nonprofit Organizations.

Salamon, L. M. (1999). America's nonprofit sector: A primer. *New York: Foundation Center,*

Salamon, L., Geller, S., & Spence, K. (2009). Impact of the 2007-2009 economic recession on nonprofit organizations. *The John Hopkins Listening Post Project,* (14)

Sanctuary for Families. (2013). Children's Program. Retrieved from http://www.sanctuaryforfamilies.org/index.php?option=com_content&task=view&id=40&Itemid=69

Sanders, W. G., & Hambrick, D. C. (2007). Swinging for the fences: The effects of CEO stock options on company risk taking and performance. *Academy of Management Journal*, *50*(5), 1055-1078.

Sawhill, J. C., & Williamson, D. (2001). Mission impossible?: Measuring success in nonprofit organizations. *Nonprofit Management and Leadership*, *11*(3), 371-386.

Shamir, B., Zakay, E., Breinin, E., & Popper, M. (1998). Correlates of charismatic leader behavior in military units: Subordinates' attitudes, unit characteristics, and superiors' appraisals of leader performance. *Academy of Management Journal*, *41*(4), 387-409.

Shapiro, S. P. (2005). Agency theory. *Annual Review of Sociology*, , 263-284.

Sheehan, R. M. (1996). Mission accomplishment as philanthropic organization effectiveness: Key findings from the excellence in philanthropy project. *Nonprofit and Voluntary Sector Quarterly*, *25*(1), 110-123.

Shipan, C. R., & Volden, C. (2008). The mechanisms of policy diffusion. *American Journal of Political Science*, *52*(4), 840-857.

Singh, J. V., Tucker, D. J., & House, R. J. (1986). Organizational legitimacy and the liability of newness. *Administrative Science Quarterly,31*(2) , 171-193.

Sitkin, S. B., & Pablo, A. L. (1992). Reconceptualizing the determinants of risk behavior. *Academy of Management Review, 17*(1), 9-38.

Skinner, D. (2004). Evaluation and change management: rhetoric and reality. *Human Resource Management Journal*, *14*(3), 5-19.

Sobeck, J., & Agius, E. (2007). Organizational capacity building: Addressing a research and practice gap. *Evaluation and Program Planning*, *30*(3), 237-246.

Söderholm, P., & Klaassen, G. (2007). Wind power in europe: A simultaneous innovation–diffusion model. *Environmental and Resource Economics*, *36*(2), 163-190.

Sowa, J. E., Selden, S. C., & Sandfort, J. R. (2004). No longer unmeasurable? A multidimensional integrated model of nonprofit organizational effectiveness. *Nonprofit and Voluntary Sector Quarterly*, *33*(4), 711-728.

Stead, W., & Stead, J. (1992). Management for a small planet: Strategic decision making and environmental.

Stinchcombe, A. L. (1965). Organizations and social structure. *Handbook of Organizations*, *44*(2), 142-193.

Straub, E. T. (2009). Understanding technology adoption: Theory and future directions for informal learning. *Review of Educational Research, 79*(2), 625-649.

Teodoro, M. P. (2010). Contingent professionalism: Bureaucratic mobility and the adoption of water conservation rates. *Journal of Public Administration Research and Theory, 20*(2), 437-459.

Teyssier, S. (2012). Inequity and risk aversion in sequential public good games. *Public Choice, 151*(1-2), 91-119.

The Urban Institute. (2002). How and Why Nonprofits Use Outcome Information.

Thomson, D. E. (2010). Exploring the role of funders' performance reporting mandates in nonprofit performance measurement. *Nonprofit and Voluntary Sector Quarterly, 39*(4), 611-629.

Thornton, R. (1970). Organizational involvement and commitment to organization and profession. *Administrative Science Quarterly, ,* 417-426.

Tolbert, P. S. & Zucker, L. G. (1996). The institutionalization of institutional theory [Electronic version]. In S. Clegg, C. Hardy and W. Nord (Eds.), Handbook of organization studies (pp. 175-190). London: SAGE.

Tolbert, P. S., & Zucker, L. G. (1983). Institutional sources of change in the formal structure of organizations: The diffusion of civil service reform, 1880-1935. *Administrative Science Quarterly, ,* 22-39.

Tornow, W. W. (1993). Perceptions or reality: Is multi-perspective measurement a means or an end? *Human Resource Management, 32*(2-3), 221-229.

Treich, N. (2010). Risk-aversion and prudence in rent-seeking games. *Public Choice, 145*(3-4), 339-349.

Tuckman, H. P., & Chang, C. F. (1991). A methodology for measuring the financial vulnerability of charitable nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly, 20*(4), 445-460.

Tung, F., Chang, S., & Chou, C. (2008). An extension of trust and TAM model with IDT in the adoption of the electronic logistics information system in HIS in the medical industry. *International Journal of Medical Informatics, 77*(5), 324-335.

Turner-Stokes, L., & Turner-Stokes, T. (1997). The use of standardized outcome measures in rehabilitation centres in the UK. *Clinical Rehabilitation, 11*(4), 306-313.

Ullian, L., & Schink, A. B. (2012). Digital M: Professionalism and credentialing in the field of volunteer management. *The Volunteer Management Handbook: Leadership Strategies for Success, Second Edition, ,* M. 1-M. 19.

Van de Werfhorst, Herman G, & Hofstede, S. (2007). Cultural capital or relative risk aversion? two mechanisms for educational inequality compared1. *The British Journal of Sociology, 58*(3), 391-415.

Van Dooren, W. (2008). Nothing new under the sun? change and continuity in twentieth century performance movements. *Performance Information in the Public Sector: How it is used*, , 11-23.

Van Slyke, D. M. (2007). Agents or stewards: Using theory to understand the government-nonprofit social service contracting relationship. *Journal of Public Administration Research and Theory, 17*(2), 157-187.

Vasi, I. B., & King, B. G. (2012). Social movements, risk perceptions, and economic outcomes the effect of primary and secondary stakeholder activism on firms' perceived environmental risk and financial performance. *American Sociological Review, 77*(4), 573-596.

Vertzberger, Y. Y. (1995). Rethinking and reconceptualizing risk in foreign policy decision-making: A sociocognitive approach. *Political Psychology*, , 347-380.

Weber, M. (1978). *Economy and society: An outline of interpretative sociology* Univ of California Press.

Westphal, J. D., Gulati, R., & Shortell, S. M. (1997). Customization or conformity? an institutional and network perspective on the content and consequences of TQM adoption. *Administrative Science Quarterly*, , 366-394.

Wezel, F. C., & Saka-Helmhout, A. (2006). Antecedents and consequences of organizational change:'Institutionalizing'the behavioral theory of the firm. *Organization Studies, 27*(2), 265-286.

Wholey, J. S. (1999). Performance-based management: Responding to the challenges. *Public Productivity & Management Review*, , 288-307.

Wiener, S. J., Kirsch, A. D., & McCormack, M. T. (2002). *Balancing the scales: Measuring the roles and contributions of nonprofit organizations and religious congregations* Independent Sector.

Wiseman, R. M., & Gomez-Mejia, L. R. (1998). A behavioral agency model of managerial risk taking. *Academy of Management Review, 23*(1), 133-153.

Wouters, M., & Wilderom, C. (2008). Developing performance-measurement systems as enabling formalization: A longitudinal field study of a logistics department. *Accounting, Organizations and Society, 33*(4), 488-516.

Yamazaki, T. (2009). The uniqueness of pure-strategy nash equilibrium in rent-seeking games with risk-averse players. *Public Choice, 139*(3-4), 335-342.

Yang, A., & Taylor, M. (2013). The relationship between the professionalization of public relations, societal social capital and democracy: Evidence from a cross-national study. *Public Relations Review, 39*(4), 257-270.

Yang, K. (2009). Examining perceived honest performance reporting by public organizations: Bureaucratic politics and organizational practice. *Journal of Public Administration Research and Theory*, *19*(1), 81-105.

Youthville. (2013). Foster Care. Retrieved from http://www.youthville.org/FamilyFosterCare/Default.aspx

Yu, C., & Tao, Y. (2009). Understanding business-level innovation technology adoption. *Technovation, 29*(2), 92-109.

Yuan, Y. C. (2010). Multiple imputation for missing data: Concepts and new development (Version 9.0). *SAS Institute Inc, Rockville, MD*.

Zattoni, A., & Cuomo, F. (2008). Why adopt codes of good governance? A comparison of institutional and efficiency perspectives. *Corporate Governance: An International Review, 16*(1), 1-15.

Zhu, K., Dong, S., Xu, S. X., & Kraemer, K. L. (2006). Innovation diffusion in global contexts: Determinants of post-adoption digital transformation of european companies. *European Journal of Information Systems, 15*(6), 601-616.

Zhu, K., Kraemer, K. L., & Xu, S. (2006). The process of innovation assimilation by firms in different countries: A technology diffusion perspective on e-business. *Management Science, 52*(10), 1557-1576.

Zimmermann, J. A. M., & Stevens, B. W. (2006). The use of performance measurement in south carolina nonprofits. *Nonprofit Management and Leadership, 16*(3), 315-327.

APPENDICES

APPENDIX A: DOCUMENTS USED IN SURVEY

A-1: Institution Review Board Approval

North Carolina State University is a land-grant university and a constituent institution of the University of North Carolina.

Office of Research and Innovation
Division of Research Administration

NC STATE UNIVERSITY

Campus Box 7514
Raleigh, North Carolina 27695-7514

919.515.8754 (phone)
919.515.7721 (fax)

From: Jennifer Ofstein, IRB Coordinator
North Carolina State University
Institutional Review Board

Date: November 19, 2013

Title: The Adoption of Outcome Measures in Nonprofit Organizations: Empirical Analysis of Human Service Nonprofits

IRB#: 3625

Dear Chongmyoung Lee,

The research proposal named above has received administrative review and has been approved as exempt from the policy as outlined in the Code of Federal Regulations (Exemption: 46.101. b.2). Provided that the only participation of the subjects is as described in the proposal narrative, this project is exempt from further review. This approval does not expire, but any changes must be approved by the IRB prior to implementation.

NOTE:

1. This committee complies with requirements found in Title 45 part 46 of The Code of Federal Regulations. For NCSU projects, the Assurance Number is: FWA00003429.

2. Any changes to the research must be submitted and approved by the IRB prior to implementation.

3. If any unanticipated problems occur, they must be reported to the IRB office within 5 business days.

Please forward a copy of this letter to your faculty sponsor, if applicable.
Thank you.

Sincerely,

Jennifer Ofstein
NC State IRB

A-2: Informed Consent Form

North Carolina State University
INFORMED CONSENT FORM for RESEARCH

Title of Study: The Adoption of Outcome Measures in Nonprofit Organizations: Empirical Analysis of Human Service Nonprofits

Principal Investigator: Chongmyoung Lee Faculty Sponsor: Richard Clerkin

---

What are some general things you should know about research studies?
You are being asked to take part in a research study. Your participation in this study is voluntary. You have the right to be a part of this study, to choose not to participate or to stop participating at any time without penalty. The purpose of research studies is to gain a better understanding of a certain topic or issue. You are not guaranteed any personal benefits from being in a study. Research studies also may pose risks to those that participate. In this consent form you will find specific details about the research in which you are being asked to participate. If you do not understand something in this form it is your right to ask the researcher for clarification or more information. A copy of this consent form will be provided to you. If at any time you have questions about your participation, do not hesitate to contact the researcher(s) named above.

What is the purpose of this study?
The purpose of this study is to gain a better understanding of the use of outcome measures in nonprofit organizations.

What will happen if you take part in the study?
If you agree to participate in this study, you will be asked to complete several questions regarding the factors that influence the adoption of outcome measures and extent of outcome information use. Total time for your participation is estimated to be 20 minutes. If you cannot accurately answer the questions, please pass this survey on to the person who can best answer them.

Risks
There should be no risk to you from the questions. The participation in this study is voluntary and refusal to participate involves no penalty or loss of benefits to which you are entitled. You may discontinue participation at any time without penalty or loss of benefits.

Benefits
This study will deepen our understanding of the outcome measurement in nonprofit organizations. If you would like to know the results of this study, I will share my findings with you.

Confidentiality
The information in the study records will be kept confidential to the full extent allowed by law. Data will be stored securely in electronic format. No reference will be made in oral or written reports which could link you to the study. Results will only be reported in the aggregate. If the data are made available to anyone beside the investigator, all identifying information will be removed from the database.

Compensation
I will randomly select 7 respondents and send them 1 $100 visa gift card, 1 $50 visa gift card, and 5 $10 starbucks gift cards as a thank you for participating in the survey.

What if you have questions about this study?

If you have questions at any time about the study or the procedures, you may contact the researcher, Chongmyoung Lee, at 1600 Baity Hill Dr. #323, NC, 27514, or 919-332-7510

What if you have questions about your rights as a research participant?
If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Deb Paxton, Regulatory Compliance Administrator, Box 7514, NCSU Campus (919/515-4514).

Consent To Participate
"I have read and understand the above information. I have received a copy of this form. I agree to participate in this study with the understanding that I may choose not to participate or to stop participating at any time without penalty or loss of benefits to which I am otherwise entitled."

"I have read the above informed consent and agree to continue with this survey." _____
"I do not wish to participate in this survey." _____

A-3: Survey Instrument

I. The first set of questions asks about outcomes your organization may measure.

A nonprofit organization's performance is generally measured by its outputs and outcomes. In this study, I am interested in outcome measures. Following are definitions and examples of output and outcome measures:

- Output: A product or service directly produced.
  - Examples: The frequency of family counseling services per month.
- Outcome: Substantial changes or benefits resulting from outputs.
  - Examples: The change in the quality of life of beneficiaries or reduced childhood behavioral and emotional disorders.

1. In the last fiscal year, what percentage of your organization's programs/services was subject to outcome measurement? _____

1 = 0%, 2 = 1% - 25%, 3 = 26% - 50%, 4 = 51% - 75%, and 5 = 76% - 100%

2 (*For NTEE code p32, Foster Care*) Do you collect the following information?

Outcome measures	Yes	No
Increased placement stability		
Client satisfaction with respect to foster care service		
Reduced childhood behavioral and emotional disorders		
Increased quality of life		
Enhanced level of education or life skills of foster child		

3 (*For NTEE code p32, Foster Care*) How often do you collect the following information? Please answer the frequency of outcome measurement (e.g., once per each program/service, monthly, quarterly, yearly, etc). This is free-response question.

Outcome measures	Frequency
Increased placement stability	
Client satisfaction with respect to foster care service	
Reduced childhood behavioral and emotional disorders	
Increased quality of life	
Enhanced level of education or life skills of foster child	

4 (*For NTEE code p32, Foster Care*) Please list any outcomes that were not included in the previous question that you consider to be one of your top three outcomes. Please provide the frequency of measurement for any items you list.

_____

_____

_____

2-1 (*for NTEE code p43, Family Violence Shelters*) Do you collect the following information?

Outcome measures	Yes	No
Client satisfaction with respect to the shelter's service		
Reduced physical/mental illness (e.g., depression and other affective disorders)		
Increased quality of life		
Increased self-efficacy		
Reduced childhood behavioral and emotional disorders		

3-1 (*for NTEE code p43, Family Violence Shelters*) How often do you collect the following information? Please answer the frequency of outcome measurement (e.g., once per each program/service, monthly, quarterly, yearly, etc). This is free-response question.

Outcome measures	Frequency
Client satisfaction with respect to the shelter's service	
Reduced physical/mental illness (e.g., depression and other affective disorders)	
Increased quality of life	
Increased self-efficacy	
Reduced childhood behavioral and emotional disorders	

4-1 (*for NTEE code p43, Family Violence Shelters*) Please list any outcomes that were not included in the previous question that you consider to be one of your top three outcomes. Please provide the frequency of measurement for any items you list.

_____

_____

_____

2-2 (*for NTEE code p46, Family Counseling*) Do you collect the following information?

Outcome measures	Yes	No
Client satisfaction with respect to family counseling service		
Increased self-efficacy		
Reduced physical/mental illnesses (e.g., depression and other affective disorders)		
Increased quality of life		
Reduced drug and alcohol use		
Reduced childhood behavioral and emotional disorders		
Reduced marital and relationship problems		
Reduced child abuse and neglect		

3-2 (*for NTEE code p46, Family Counseling*) How often do you collect the following information? Please answer the frequency of outcome measurement (e.g., once per each program/service, monthly, quarterly, yearly, etc). This is free-response question.

Outcome measures	Frequency
Client satisfaction with respect to family counseling service	
Increased self-efficacy	
Reduced physical/mental illnesses (e.g., depression and other affective disorders)	
Increased quality of life	
Reduced drug and alcohol use	
Reduced childhood behavioral and emotional disorders	
Reduced marital and relationship problems	
Reduced child abuse and neglect	

4-2 (*for NTEE code p46, Family Counseling*) Please list any outcomes that were not included in the previous question that you consider to be one of your top three outcomes. Please provide the frequency of measurement for any items you list.

_____

_____

_____

5. Please check all periods when you were measuring outcomes. _____

1 = Before FY1980, 2 = FY1981 - 1990, 3 = FY1991 –2000, 4 = FY2001– 2010, 5 = Since FY2011, and 0 = Not measuring outcomes (Not Applicable)

II. In this section, I want to learn more about how your nonprofit operates.

1. To what extent do you agree with the following statements?

1 to 5 with 1=strongly disagree, and 5 =strongly agree (0 = Don't know)

1) There is a reward/incentive system in place in the organization that encourages risk taking (using innovative ideas with the goal of improving performance).

2) Management is willing to implement new ways to accomplish our work.

3) Non-management employees willingly adopt new ways to accomplish our work.

2. Do your grants/contracts require performance information? _____

1 = Yes and 2 = No (0 = Don't know)

If your answer is "No" or "Don't know", you can skip following question#3.

Just as a reminder, this survey distinguishes between outputs and outcomes.

- An output is defined as a product or service directly produced. An example of an output is the percentage of children in foster care receiving at least 2 visits per month from the social worker.
- An outcome is defined as substantial changes or benefits resulting from outputs. An example of an outcome is reduced physical/mental illness of children.

3. Which category most closely describes the extent to which your grants/contracts focus on outputs or outcomes? _____

1 = 100% Output and 0% Outcome, 2 = 75% Output and 25% Outcome, 3 = 50% Output and 50% Outcome, 4 = 25% Output and 75% Outcome, and 5 = 0% Output and 100% Outcome (0 = Don't know)

4. What percentage of your organization's program funding is subject to mandatory outcomes measurement? _____

1 = 0%, 2 = 1% - 25%, 3 = 26% - 50%, 4 = 51% - 75%, and 5 = 76% - 100% (0 = Don't know)

5. Does your organization receive funding from the United Way? _____

1 = Yes and 2 = No (0 = Don't know)

6. To what extent do you agree with each of the following statements?

1 to 5 with 1=strongly disagree, and 5=strongly agree (0 = Don't know)

1) The largest organizations in our field measure their outcomes _____

2) Nonprofits in our field that measure their outcomes improve their performance _____

3) Changes in our community's socioeconomic status are predictable _____

4) Changes in our community's political environment are predictable _____

7. What percentage of your employees have graduate or professional degrees (e.g., Master's degree, Ph.D, MD, JD, or other advanced degrees beyond a Bachelor's degree)? _____

1 = 0%, 2 = 1% – 5%, 3 = 6% – 10%, 4 = 11% – 20%, 5 = 21% – 30%, 6 = 31% – 50%, and 7 = 51% or higher (0 = Don't know)

8. Does the executive director (or top manager) of your organization have a graduate degree in management such as a Master of Business Administration (MBA), a Master of Public Administration (MPA), or a similar management degree? _____

1 = Yes, 2 = No, Other graduate degree, 3 = No, Doesn't have a graduate degree (0 = Don't know)

9. Is your organization accredited by any external bodies? _____

1 = Yes and 2 = No (0 = Don't know)

If your answer is "No" or "Don't know", you can skip following question#10.

10. Please write down the name of the accrediting body.

_____

11. To what extent do you agree with the following statements?

1 to 5 with 1=strongly disagree, and 5=strongly agree (0 = Don't know)

1) Our organization has staff with the skills and expertise necessary to track program participants _____

2) Our organization has adequate infrastructure (e.g., information technology, communication system) for daily operations _____

3) Our organization has written rules concerning compensation, governance, job descriptions, conflict of interest, and/or volunteers _____

4) Top management in our organization supports using data to make decisions _____

5) Most employees in our organization support using data to make decisions _____

6) Stakeholders of our organization support using data to make decisions _____

7) Our organization frequently discusses operational issues with stakeholders ____

8) Our organization seeks help from other organizations to improve decision making _____

III. In this section, I want to learn about the use of outcome information in your organization.

1. To what extent does your nonprofit use outcome information for each of the following purposes?

1 to 5 with 1=Not at all, and 5=Very great extent (0 = Don't know)

1) Improve management decisions _____

2) Improve strategic planning _____

3) Improve services for clients _____

4) Appraise employee performance _____

5) Support budget decisions _____

6) Report to the board _____

7) Improve communication between administrators and elected officials _____

8) Respond to public demands for accountability _____

9) Respond to pressure from community groups _____

10) Comply with government requirements _____

11) Report to funders _____

12) Get new funding _____

13) Outreach and public relations _____

IV. Now, I would like to know more about your overall opinion about outcomes measurement.

1. Please include any comments you have about outcome measures or outcome measurement activities of your organization.

_____

2. Please indicate your job title (e.g., President, Executive Director, Associate Director, etc)

_____

3. Have you ever received any training program on outcome measurement? _____

1 = Yes and 2 = No

4. In which year was your organization founded? _____

5. How many full time employees do you have? _____

6. How many part time employees do you have? _____

Thank you so much. Your opinions are very valuable to my study.

If you are willing to be contacted at a later date to provide follow-up or additional information, please enter your name, email address, and phone number

_____

If you would like to be entered into a random drawing for a gift card, please enter your email address here.

_____

If you would like to know the results of this study, please enter your email address here.

_____