INTEGRATING THE BALANCED SCORECARD (BSC) AND THE DATA ENVELOPMENT ANALYSYS (DEA) APPROACHES FOR AN ENHANCED POLICE PERFORMANCE MEASUREMENT SYSTEM

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Myungwoo Lee

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Program:
Criminal Justice

July 2021

Fargo, North Dakota

North Dakota State University
Graduate School

Title

INTEGRATING THE BALANCED SCORECARD (BSC) AND THE DATA ENVELOPMENT ANALYSIS (DEA) APPROACHES FOR AN ENHANCED POLICE PERFORMANCE MEASUREMENT SYSTEM

By

Myungwoo Lee

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Dr. Carol Archbold
Chair

Dr. Jeffrey Bumgarner

Dr. Andrew Myer

Dr. Tim O. Peterson

Approved:

July 13, 2021
Date

Dr. Carol Archbold
Department Chair

ABSTRACT

An effective performance measurement system is an integral part of modern police management. Police agencies have measured their performance against a very restricted set of crime-focused indicators, such as crime rates, arrests, response times, and clearance rates. Police performance should be measured across multiple dimensions to capture public values produced by modern police agencies.

This study set out to present an enhanced performance measurement system for police agencies by integrating the Balanced Scorecard (BSC) and Data Envelopment Analysis (DEA) approaches. The BSC provides the theoretical foundation for building a comprehensive performance measurement framework, while the DEA provides the analytical tool to test the theoretical framework. Integrating the DEA and the BSC approaches can create many synergy effects because they are complementary to each other.

A case-study approach was used to assess the feasibility of the integrated performance measurement system; to critically examine the ways in which performance information can be used for performance management in police agencies; and to put forward some recommendations regarding its successful application in practice. Police stations under the Seoul Metropolitan Police Agency (SMPA) were chosen for conducting this case study.

The Dynamic-Network (DN) DEA, with assumptions of input-orientation, variable returns-to-scale (VRS), and slack-based measure (SBM), was run to estimate the proposed police performance measurement model. The DN DEA presented the overall performance over the entire observed period as well as dynamic changes of the perspective-period performance. The DN DEA also presents the practical ways in which inefficient police stations become more efficient by reporting the specific benchmarking objects and the target input and output levels for

the inefficient police stations. When network and dynamic dimensions, derived from the BSC, are incorporated in a DNDEA model, a more comprehensive information can be obtained and thus enables accurate estimate of organizational performance as well as identify potential improvements in more detail.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIX TABLES

INTRODUCTION

Police agencies have been confronted with the dual pressures of increased demands for services and resources (Mastrofski, 2006; Mastrofski & Willis, 2010). The number of police officers in South Korea in 2016 reached 114,658, representing 18 consecutive years of growth and the highest police officer strength since Korean National Police Agency begun to collect national-level crime and policing data. Total spending on policing was over $ 9.8 billion in 2016, more than double its 2001 level (Korean National Police Agency [KNPA], 2016). Continuous efforts have been made to strengthen the police force for the last few decades. However, the reported crime rate in South Korea has slightly fluctuated from 2001 to 2016 but overall has been steady. In addition, there has been a marked increase in the number of calls to police which involve a wide range of non-criminal (or non-emergency) calls for service, including social and mental health incidents (KNPA, 2016).

In an environment with limited resources, growing policing costs are increasingly recognized as a serious concern for various policing stakeholders. Citizens and taxpayers ask the police to achieve their goals and demonstrate results. Police agencies should compete with other public organizations for public funds by providing accountability and value for money. Police managers are measuring the performance of their organization to obtain adequate information and to provide the needed evidence. Measuring the performance of policing services can play a critical role in addressing the issues of demonstrating accountability and stewardship of taxpayer resources. Performance measurement is fast becoming a critical component in modern police management.

There is increasing concern over a traditional way of measuring police performance. The performance of policing services has been measured and evaluated since the institution of modern policing was established in the mid-nineteenth century (Maguire & Uchida, 2000). The initial measures of police performance involve a very restricted set of crime-related indicators, such as crime rates, arrests, response times, and clearance rates (Davis, 2012). Even CompStat, perhaps the best-known contemporary example of the power of performance measurement in police agencies, focus predominantly on the volume of crime reported in their jurisdiction as the bottom line for policing (Moore & Braga, 2003; Sparrow, 2015). Although these measures, obtained mostly from the reported crime statistics, have retained their prominence, a limitation of using this kind of data is that the focus is narrow because crime control is just one of several components of the police mission, as well as reported crime overlooks unreported crimes.

Police perform a variety of tasks. Other than the traditional tasks of pursuing, arresting and charging criminals, preventing crimes, and dealing with traffic-related offences and accidents, police are further expected to resolve various conflicts in their communities, reduce or prevent social disorder, and construct and maintain community relations (Archbold, 2012). In addition to the traditional crime measures, other specific indicators such as citizen and employee satisfaction should be taken into consideration (Maslov, 2016). There needs to be a balance among multiple measures reflecting the complex array of responsibilities of police agencies (Kiedrowsk et al., 2014; Moore & Braga, 2003a).

Policing experts have emphasized the importance of the multi-dimensional nature of police performance and provided some ideas about how to specify the performance dimensions and measures in line with the principles of the Balanced Scorecard (BSC) approach (Davis, 2012; Mastrofski, 1999; Moore & Braga, 2003a; Sparrow, 2015). The BSC, developed by Kaplan and Norton (1992), is one of the best-known and most widely used conceptual frameworks for measuring organizational performance. Developed from an organization’s vision and strategy, this framework includes measures (key performance indicators, KPIs)[1], targets, and initiatives distributed among four perspectives: financial results, innovation and learning, internal processes, and customer satisfaction. The BSC helps managers to view organizations from the four perspectives and to obtain comprehensive information on the organization's performance in a single system.

Since its application by Thanassoulis (1995), the Data Envelopment Analysis (DEA) has become one of the most widely used techniques for assessing police performance. DEA is a non-parametric technique for measuring a relative performance of a set of comparable units (e.g., police agencies). It allows each unit to identify a benchmark group and handle multiple inputs and outputs simultaneously. It has been argued that the BSC and DEA are complementary to each other. Nevertheless, the research to date has not been able to integrate the DEA and the BSC approaches for an enhanced assessment of police organizational performance (Amado et al., 2012).

A number of conceptual and methodological issues need to be addressed to develop the integrated police performance measurement system. On the conceptual side, performance measures should be (1) multidimensional to capture the complexity inherent in modern policing, (2) a mix of outputs and outcomes based on surveys to overcome the limitations of traditional measures, and (3) comparative to make it possible to compare an agency's performance over multiple time periods and/or to compare an agency to other agencies (Davis, 2012; Sparrow, 2015). These measures, then, should be incorporated into the conceptual framework for measuring both overall and (sub-) process performance of police organization. The Balanced Scorecard approach is chosen for this study as we have seen many successful stories not only in the private sector but also in public sector in recent years. Considering that the BSC was originally devised for private-sector firms, however, some modifications need to be made for application to public-sector organizations such as police agencies.

To support the proposed conceptual framework consistent with the BSC approach, the methodological considerations focus on the DEA that can overcome some limitations of the BSC. This study pays more attention to the Dynamic-Network (DN) DEA model developed by Tone and Tsutsui (2001, 2009, 2010, 2014) than other DEA models that have been used in the DEA-based police performance studies. The DN model can fill some gaps in the literature and facilitate the integration of the BSC with the DEA methodology by allowing us (1) to employ a non-radial slack-based measure approach that can present better practical implications; (2) to connect the causal chain links by using outputs of one model as inputs of the following model; (3) to consider various forms of interconnecting activities or carry-overs, characterized by the desirable, the undesirable, the discretionary, and the non-discretionary link value or carry-over case; and (4) to estimate both overall and process efficiency over multiple time periods in a unified mathematical and programming DEA model altogether.

This study aims to develop an integrated assessment framework for police organizations by incorporating the concepts of the BSC into the DEA methodology, and to test practical relevance of the framework by using a case study of 31 police stations within the Seoul Metropolitan Police Agency (SMPA). A well-designed performance measurement framework may provide more useful information for organizational management and improvement, and ultimately satisfy citizens and other stakeholders the police serve and protect.

LITERATURE REVIEW

Police Performance Measurements

Why Measure?

An effective performance measurement system is an integral part of modern police management (Sparrow, 2015). Police administrators utilize the performance measurement system as a tool to make the police not only externally accountable to stakeholders and the public, but also internally accountable to all their members. A comprehensive assessment of police performance can increase external and internal accountability to government, society, and the public at large (Larson, 1978; Marteache & Maxfield, 2011; Moore & Braga, 2003b).

As for external accountability of the police, performance measures can help administrators provide more accurate information and defend against criticisms or requests from the public, their representatives, and other entities to which the agencies are accountable (Moore & Braga, 2003b). Municipal executives and legislators demand observable, empirical and measurable evidence to justify budget requests made by police departments (Davis, 2012). Since the 2008 global financial crisis, austerity has become “the new normal” all over the world. Many municipal officials have changed their budgeting process and have increasingly adopted the practice of outcome-based budgeting based on the relationship between funding and expected results (Probst, 2009). As police departments are the largest funded government agency in most municipalities and this ratio grows, finding ways to operate the public funds efficiently (as well as effectively and fairly) is a top concern of police administrators in an era of austerity. Police agencies should provide relevant, reliable, and comparable performance information for demonstrating value for money (Gascón, 2010).

As for internal accountability of the police, performance measures can help police executives monitor their operations in communities and promote adherence to agency policies and strategic plans. Also, performance indicators can aid executives in managing several personnel issues, including stress, job satisfaction, cynicism, and morale among police officers (Lilley & Hinduja, 2007; Moore & Braga, 2003b). Moreover, "by defining what is measured, executives send a signal to their command about what activities are valued and what results are considered important" (Davis, 2012, p. 1). Performance measurement systems are believed to ensure that the police are held externally and internally accountable in contemporary law enforcement.

A Brief History of Police Performance Measurement

Historically, police executives have measured their agency performance through traditional indicators such as crime rates, clearance rates, response times, and measures of enforcement productivity (Moore & Braga, 2003a; Sparrow, 2015). As traditional policing focused their duties primarily on crime control, these measures became institutionalized over many years (Alpert & Moore, 1993; Maguire, 2003). Information systems were developed to record police performance in relation to the traditional measures. In 1930, the Uniform Crime Reports (UCR) program was developed to collect crime data from police agencies throughout the nation in the United States. The primary objective of this program is to "generate reliable information for use in law enforcement administration, operation, and management" (https://www.fbi.gov/services/cjis/ucr). Starting in 1939, the International City Managers' Association (ICMA) also began collecting data from participating police agencies around the world as part of its Municipal Yearbook Series (Maguire, 2003); data includes measures of crimes, arrests, and efficient use of resources. Since national or international standardized systems were developed in the early 20th century, the traditional measures of police agency performance have become entrenched within the field of policing. As stated by Geoffrey Alpert and Mark Moore (1993), these "generally accepted accounting practice became enshrined as the key measures to evaluate police performance… these are the statistics by which police departments throughout the United States are now held accountable" (p.110). Maguire (2003) also supported the fact that the UCR data has now become one of the country's leading social indicators for comparative performance measurement of police agencies in the United States.

Traditional measures remain crucial as part of an overall system for measuring police performance. A significant proportion of today's police organizations seem to remain narrowly focused on the same categories of indicators that have dominated the field for decades (Sparrow, 2015). COMPuter STATistics (COMPSTAT) is perhaps the best-known contemporary example of the power of performance measurement in law enforcement management (Roberts, 2006). First implemented in 1994 by Commissioner William Bratton of the New York City Police Department, COMPSTAT is now being adopted by agencies around the United States and in other nations (Maguire & Uchida, 2000). COMPSTAT generates crime statistics for each precinct and uses them to hold precinct commanders accountable for addressing crime in their area (Bratton, 1999). COMPSTAT made it possible for the department to set ambitious goals, to continually monitor progress toward those goals, and eventually to drive its officers to higher levels of accomplishment and foster a revitalized organizational culture (Bratton, 1999). This process requires regular COMPSTAT meetings where district or precinct managers and officers discuss and analyze crime problems and the strategies used to address those problems. All members of the organization are held accountable for crime problems in the area that they work and to find solutions to those problems. COMPSTAT is believed to change police organization by way of the culture of police organizations and the manner in which they function (Archbold, 2012).

From an administrative perspective, COMPSTAT represents a significant change; however, from a measurement perspective, it relies on the same crime data that have been criticized as inadequate measures of police performance for the past three decades. Traditional measures emphasize only the crime control aspect of policing and do not adequately address the many other police agency activities (Maguire & Uchida, 2000; Moore & Braga, 2003a). These measures reflect an increasingly outdated model of police tasks and fail to capture many important contributions that police make to their communities. More importantly, these measures may misguide police managers and lead them and their organizations towards purposes and activities that are less valuable than others. Although COMPSTAT-like systems allow precinct commanders to talk about their special efforts to deal with serious crime problems, it provides little room for them to talk about problem-solving efforts focused on non-crime problems (Moore & Braga, 2003b). As George Kelling pointed out, "COMPSTAT appropriately focuses on crime, but … the danger is that COMPSTAT doesn't always balance the focus with the other values that policing is supposed to pursue" (PERP, 2013, p.1). As NYPD Assistant Commissioner Ronald J. Wilhelmy wrote in a November 2013 internal NYPD strategy document:

> we cannot continue to evaluate personnel on the simple measure of whether crime is up or down relative to a prior period. Most importantly, CompStat has ignored measurement of other core functions. Chiefly, we fail to measure what may be our highest priority: public satisfaction. We also fail to measure quality of life, integrity, community relations,

administrative efficiency, and employee satisfaction, to name just a few other important areas.

Police Performance Measures Conceptualized

Police performance measurements need to incorporate the principals of community-oriented policing. Historically, crime control is the "bottom line" for policing. Since the advent of community and problem-oriented policing, the roles and functions of the police have expanded considerably from crime fighter to problem solver and neighborhood ombudsperson. As defined by Alpert and Moore (1993), community policing includes "building a strong relationship with the community, attacking fear of crime through enhancing neighborhood quality of life, encouraging police officers to focus on problems rather than on incidents, and decentralizing authority" (Davis et al., 2015, p. 471). They argued that performance measures need to capture how agencies are performing in these aspects of community policing through examining what agencies are doing to promote community policing and how the community is responding. Performance measurement systems, based entirely on the traditional indicators such as response time and clearance rates, prevent police agencies not only from moving towards a strategy of community policing but also from assessing police efforts to address community concerns. Moore and Braga (2003a) argued that only by adopting a comprehensive performance measurement system that incorporates the precepts of community policing will police executives have a chance to spur their agencies towards a community problem-solving strategy and improve their accountability and performance.

Police performance should be measured across multiple dimensions to capture public values produced by modern police organizations. While reducing crime is the single most important core function of the police, there are many other dimensions of performance that are valued and should be measured (Moore & Braga, 2003a; Moore & Braga, 2004). Herman Goldstein (1977), for example, defined eight important functions of the police, including "patrolling the streets; responding to calls for service; investigating crimes; arresting suspected offenders; regulating traffic; responding to citizen requests for assistance; handling crowds and demonstrations; and providing a variety of emergency medical and social services" (p.35). It has been argued that the functions of the police are broader than simply controlling crime (Goldstein, 1990; Bayley, 1994). Goldstein's framework has pushed the police not only to view their role more broadly than as part of the criminal justice system, but also to lay a foundation for adapting community and problem-oriented policing (Sparrow, 2015). It should be noted that the mission or strategy of a police department is not a fixed or permanent thing. Instead, it is something to be decided by the department leader in light of environmental circumstances-both the task environment of problems that the police confront as well as the authorizing environment of public expectations and demands of the police (Moore, 1995).

In recent years, several attempts have been made to understand multiple dimensions of police performance. Moore and Braga (2003a) proposed seven dimensions of value that can be used to evaluate their police departments. The dimensions come from answering the question of 'what citizens should value (and measure) in policing.' Unlike Goldstein's work, they produced a set of concrete performance measures for these conceptual dimensions by using existing measures, as well as proposing *new* performance measurement scheme. They argued that the improved measures of police performance make police departments more accountable and improve their performance. Table 1 shows the Moore and Braga's seven dimensions of police performance and possible indicators.

Table 1

Moore and Braga's Seven Dimensions of Police Performance and Possible Indicators

PERFORMANCE DIMENSION	INDICATORS
Reduce criminal victimization	Report crime rates Victimization rates
Call offender to account	Clearance rates Conviction rates
Reduce fear and enhance personal security	Reported change in levels of fear Reported changes in self-defense measures
Guarantee safety in public spaces	Traffic fatalities, injuries and damage Increased utilization of parks and public spaces Increased property values
Use financial resources fairly, efficiently and effectively	Cost per citizen Deployment efficiency/fairness Scheduling efficiency Budget compliance Overtime expenditures Civilianization
Use force and authority fairly, efficiently, and effectively	Citizen complaints Settlements in liability suits Police shootings
Satisfy customer demands and Achieve legitimacy with those policed	Satisfaction with police services Response times Citizen perceptions of fairness

Drawing heavily on the Moore and Braga’s work, Davis and his colleagues (2015) also developed a set of performance indicators to evaluate nine dimensions of policing, including: (1) delivering quality services; (2) fear, safety, and order; (3) ethics and values; (4) legitimacy/customer satisfaction; (5) organizational environment/commitment to high standards; (6) reducing crime and victimization; (7) efficient use of resources; (8) responding to offenders; and (9) judicious use of authority. When attempting to produce a standardized suite of performance measures along these conceptual dimensions of police performance, they focused primarily on cost effective survey methods to gather subjective information and data regarding "community opinion of the police and satisfaction of persons with recent police contacts" (Davis et al., 2015, p.479). Like Moore and Braga, they highlighted both the need for a performance measurement framework that capture the multi-dimensional nature of policing and the importance of producing a specific set of performance metrics derived from survey data.

The comprehensive view of what should be measured in police departments is informed by the idea of "balanced scorecard" approach, which was originally devised to assess a private-sector organizational performance. The BSC is one of the most prominent performance assessment framework, proposed by Kaplan and Norton (1992), and has a number of features. One key feature of the BSC is to look at the company's organizational performance from various dimensions and focuses on a set of non-financial measures organized in these dimensions, rather than single measures of financial performance represented by the bottom line for the private sector (Kaplan & Norton, 1996). A comprehensive set of performance dimensions and measures can provide a more balanced view of the organization based on four perspectives such as financial, customer, internal process, and learning and growth perspectives (Kaplan & Norton, 1992). Given that private sector companies have faced and resolved similar problems of focusing exclusively on the single bottom line (financial) measure, it is worth borrowing their experience in constructing performance dimensions and measures of policing (Moore & Braga, 2003a). This dissertation will transfer the useful managerial concepts of the BSC from the private sector to the public sector (i.e., police organizations). It will be discussed in more detail in the next section.

Designing Police Performance Measures

Once the valuable dimensions of police performance have been identified, it is necessary to develop concrete performance measures for these dimensions. Constructing specific and useful measures of police performance helps hold the police accountable and promote tangible improvements in police operations. Ideally (and theoretically), it would be possible to construct a single, simple, summary measure for each conceptual dimension of value; but, practically, it might be difficult (and impossible) to develop just one, perfect measure for each of the dimensions (Moore & Braga, 2003a). There might be several suitable measures that could be used to precisely evaluate how well the police were performing regarding that particular dimension of performance. Some important considerations have been suggested in developing a suite of performance measures that should capture the complexity of modern policing (Davis, 2012; Moore & Braga, 2003a; Sparrow, 2015).

Outcome, Output, Process and Input Measures

A basic consideration in developing performance measures is the distinctions between input, process, output, and outcome measures. Police managers have been encouraged not only to accept the idea of a "balanced scorecard" to capture the complex expectations that modern society has of the police, but also to develop a useful performance measurement system that includes multiple indicators of the ultimate results of policing (outcomes), police efforts to produce these results (outputs or processes), and the investments made in the police (inputs) (Moore & Braga, 2003b; Roberts, 2006; Walters, 1998).

Outcomes are the valuable end results that the police try to achieve in a society. It is these results that may provide both the fundamental justification for policing and the *fundamental* basis for evaluating performance. Recent evidence (from the Government Performance Review and Results Act of 2010) has emphasized the importance of outcome measures in that they are useful in testing innovative theories or programs of policing. Most desired social outcomes are not under the direct control of the police as many of the factors that result in the outcomes lie beyond the control of an organization (Davis, 2012; Moore & Braga, 2003a). An example would be citizens' feelings of safety. The police can enhance perceptions of safety through promising fear-reduction strategies and practices such as community and personalized policing, environmental design, and strategic communication to name a few (Cordner & Melekian, 2010). However, some external factors can exercise a greater influence over the feeling of safety than actions taken by the police. One notorious murder case in a neighborhood may undermine any special efforts made by police officers for enhanced private security.

Outputs are the products or services that the police produce right at the boundary of an organization. They are under the direct control of the police. It can therefore be assumed that the police may have more control over outputs than they do over outcomes. The outputs of policing are valuable not only as ends in themselves also as the means to achieving socially desired results (Davis, 2012; Moore & Braga, 2003a). For example, the number of arrests for violent crimes can be an output. Police agencies can put more or less effort into patrolling hotspots where crimes are most concentrated or adopting new computer mapping and crime analysis technology; as a result, they can increase or decrease the number of arrests.

Processes are the activities that the police generate at individual transactions with citizens. These specific transactions between police and citizens are important because their quality can be directly observed and evaluated as a particular dimension of performance, as well as because the quality can bring about other desired ends. Process measures can be used to understand the intermediate procedures in providing a product or service. Citizens and society as a whole have expected that the police deal justly and competently with an issue that concerns them (Bayley, 2005; Moore & Braga, 2003a).

Inputs are the human and capital resources used to produce the valuable results of policing: enhanced security, reduced crime, sturdier justice, and economic progress. The police use public money allocated to them by citizens who authorize and support their operations. Police departments are expensive businesses to create and maintain. In a time of tight budgets, it is expected that the police use financial resources fairly, efficiently and effectively in meeting their goals (Drake & Simpler, 2003; Moore & Braga, 2003a; Skogan, 1976).

Survey Measures

Another important consideration in designing performance measures is to think about potential data sources. Some can be derived from existing agency's administrative data, and others need to be collected using surveys or other methods. Traditional measures of police performance, including crime rates, the number of arrests and fines issued, clearance rates, and calls for service response time, are insufficient to capture diverse policing outcomes. In addition, some policing experts have indicated that these traditional measures, obtained mostly from crime statistics, fail to provide an accurate and adequate amount of information related to police performance. Sparrow (2015) pointed out some limitations of reported crime statistics, including (1) narrowing the focus of police activity, (2) manipulating crime statistics, and (3) overlooking unreported crimes. To overcome those limitations, it has been suggested that broader and subjective performance measures through survey methods (e.g., community, contact, and employee surveys) and direct observations should be considered in performance measurement systems (Gorby, 2013; Maslov, 2016; Maguire, 2004).

In recent years, there have been several attempts to develop survey measures that can supplement the traditional measures of police performance. Survey measures make more valid interpretations than data derived from the agency records (Davis, 2012). Maslov (2016) attempted to develop subjective performance measures by integrating survey methods (e.g., public opinion polling) into the framework for the Moore and Braga's seven dimensions of police performance. Mastrofski (1999) also presented six domains of performance (i.e., attentiveness, reliability, responsiveness, competence, manners, and fairness) and proposed a way in which police gather data on performance when they interact with citizens. Police officers complete a checklist related to the six domains of performance, show it to citizens and stakeholders, and obtain confirmation from them. These researchers both emphasized the need for survey measures and the importance of providing an innovative methodology to gather data on performance. It is necessary to see if both forms of objective and subjective measures show the same results.

Composite Measures (Weighting Scheme)

After constructing (and collecting) a suite of measures along multiple dimensions of performance that seem important, police researchers and practitioners may face the difficult task of how to create composite performance measures, which aggregate multiple individual performance measures into a single, summary score (Maguire, 2004). While individual performance measures are valuable in themselves (Moore & Braga, 2003a), composite performance measures are also valuable: not only for evaluating an overall performance of the police, but also for evaluating police departments along different dimensions, each of which can be observed for the department as a whole, over time, and in comparison with other departments (Maguire, 2004). In this respect, the police have to be interested in the way in which the composite measures are constructed, as well as in the quality of them on the different dimensions as both a valuable end and a valuable means.

Constructing composite performance measures is a difficult and complicated task that requires a strong conceptual and methodological foundation (National Quality Forum, 2013). To many, an obvious obstacle of measuring police performance is not just that there are multiple dimensions or measures, not just that they seem to compete (sometimes conflict) with one another, but also that they are hard to measure and combine together into a single composite score. Moore and Braga (2003a) argued, “it is impossible to know how to add the positive and negative effects together to get a net bottom line because the values are incommensurable” (p.26). However, some performance values seem to be more important than others. For example, reducing criminal victimization and calling offenders to account can be recognized (by a police chief) as a more important function of policing than reducing fear in a certain police department. In some cases, it can be useful to give more weighting to certain values when computing composite performance scores (Maguire, 2004).

Some weighting schemes have been used for producing composite performance scores of policing. One method is to use the analytic hierarchy process (AHP) that is an effective tool for dealing with complex decision making. It helps decision makers set priorities and make the best decision. Ranking scores are obtained through a paired comparison technique. Park and his colleagues (2018) used the AHP analysis along with experts’ surveys and interviews to prioritize the fundamental qualities that police officers should have for community-oriented policing. Mark Moore and his colleagues (1992) also used a similar approach for ranking the most important innovations in policing. Another method is to employ the data envelopment analysis (DEA), which is a mathematical optimization technique to measure the relative performance of decision-making units (DMU) with multiple inputs and outputs. By using an objective weighting scheme (to assign a set of weights where each unit has the highest possible efficiency rating in comparison to the other units), DEA can produce a composite measure over a set of individual measures. DEA has been used as an effective tool not only for constructing composite measures of police performance but also for performance evaluation, benchmarking, and decision making.

As DEA is the main component of this dissertation, it will be discussed in more detail in the next section.

Fair Comparisons

The fact that police organizations exist in different environments makes comparative performance measurement in policing very difficult. According to open systems theory, organizations are strongly influenced by their environment (Bertalanffy, 1972; Katz & Kahn, 1978). The environment refers to “everything external to an organization that is important for its functioning and survival” (Maguire & Uchida, 2000, p. 544). The external environment includes a wide variety of needs and influences that can affect the organization, but which the organization cannot directly control. Influences can be political, economic, ecological, societal and technological in nature. It has been argued that a highly effective organization regularly try to exchange feedback with its environment, analyze that feedback, adjusts internal system as needed to achieve the system’s goals and then transmit necessary information back out to the environment (Bertalanffy, 1972; Katz & Kahn, 1978).

Open systems theory has contributed to virtually all modern theories of organization, including contingency theory, institutional theory, and resource dependency theory (Bastedo, 2004). For example, contingency theorists argue that organizations are structured and operate in certain ways that best fit the environment in which they are embedded. Contingency theory is the dominant theoretical framework used to study practices, behaviors, and structures of police organizations (Archbold, 2012). This theory has been used by researchers to understand the environment in which the police operate and how that environment impacts police organizations (Dahle & Archbold, 2015; Goltz, 2006; Xu, 2008; Zhao et al., 2003). It has been argued that police organizations adopt their organizational structures and operational activities that are most effective and efficient in achieving specific goals. A good fit between a police organization and its environment results in higher performance (Zhao et al., 2003); on the other hand, police organizations that fails to make the appropriate adjustments to the environmental contingencies they face will not prosper, and in some cases, will not survive (Donaldson & Lex, 1995; Mastrofski, 1999).

Several studies have emphasized some of the effects that the environment might have on the police performance. For example, Goltz (2006) and Xu (2008) focused their attention on the role of the environment in developing a conceptual framework for police organizational performance. They considered the influence of both organizational and environmental factors on the police performance. Using structural equation modeling (SEM) and data envelopment analysis (DEA), Goltz (2006) examined a multi-dimensional conceptual framework that explains the relationships among three constructs: environmental constraints, the design structures of police organizations, and organizational performance indicators. This conceptual model was deeply rooted in contingency theory, which is based on the idea that an organization is shaped by its environment. This study evaluated the rationale behind contingency theory by determining the direct or indirect effect of environmental constraints on organizational design structure and performance in the context of policing. Similarly, Xu (2008) also assumed that organizational factors coupled with environmental factors affect organizational performance of local law enforcement agencies. As can be seen in the figure above, this conceptual model indicates that the organizational characteristics determined the organizational performance but was moderated by environmental and demographic variables.

There are two primary methods to control environmental factors when measuring and comparing performance: One is to form “peer groups” of similar agencies, and the other is to calculate “risk-adjusted” performance measures (Davis, 2012; Maguire, 2004). One way of adjusting for context is that comparisons should be made within the peer group to account for the different environmental inputs. In some countries, the police are organized at three hierarchical levels: local (municipal or county), regional (state or provincial), and national. Agencies at each level have some general characteristics in common that facilitate fair comparisons, including roles and responsibilities, size, type, jurisdiction, workload, and the characteristics of the communities they serve (for a discussion, see Aristovnik et al., 2013). It has been suggested that small agencies at local level should be compared with other small agencies at the same level (Davis, 2012). The other way of adjusting for the risk is to use statistical techniques, such as hierarchical linear modeling (HLM) in regression analysis and two-stage approach in data envelopment analysis (DEA); the latter DEA approach will be discussed in more detail in the next section of ‘Controlling for Non-Discretionary Inputs.’ When measuring police performance, data are often organized at individual or situational, organizational, neighborhood or regional levels. “The decisions and actions of officers are situated within these larger contexts, and they affect the quality of policing” (National Research Council, 2004). HLM is an appropriate statistical technique to analyze variance in outcome variables when the predictor variables are at the differing hierarchical levels (Woltman, Feldstain, MacKay, & Rocci, 2012). This technique accounts for the shared variance in hierarchically structured data and provides measurable evidence to make fair comparisons of police agencies with differing external factors.

Balanced Scorecard (BSC)

Basic Concepts of BSC

The Balanced Scorecard, developed by Kaplan and Norton at Harvard Business School in 1992, is one of the best-known and most widely used strategic planning and performance management system (Amado et al., 2012). The BSC provides a conceptual framework for translating an organization's vision and strategy into a set of objectives, measures, targets, and initiatives distributed among four major managerial perspectives: financial, customer, internal business process, and learning and growth (Kaplan & Norton, 1992). As well as enabling managers to clarify their vision and strategy, link objectives and measures, and develop targets and strategic initiatives, the BSC provides answers to the following questions: How should we appear to our shareholders? (Financial Perspective); How should we appear to our customers? (Customer Perspective); Which business processes must we excel at? (Internal Business Process); How will we sustain our ability to change and improve? (Learning and Growth Perspective). Progress in meeting strategic objectives is measured by comparing the actual results on each of the performance criteria against the target (Kaplan & Norton, 1992). This forms the basis for evaluating areas that require management efforts to remedy shortfalls and area where the organization is achieving its objectives (accomplish critical management processes / enhance strategic feedback and learning).

The BSC is a comprehensive and simple performance measurement tool. The term 'balance' attempts to capture various aspects of performance measurements including financial and non-financial performance measurements, long-term and short-term objectives, external and internal perspectives as well as quantitative (objective) and qualitative (subjective) measures (Chiang & Lin, 2009; Eilat, Golany & Shtub, 2008). It also considers the interests of the key stakeholders including owners, customers, and employees (Kaplan & Norton, 1996). By focusing on the four key perspectives of organizational performance and making explicit the links between them in a single system, the BSC provide top managers with a comprehensive view of their organization and prevents sub-optimization of organizational performance (Amado et al., 2012; Eilat et al., 2008).

Balanced Scorecard in Public Sectors (in Police Organizations)

The Balanced Scorecard (BSC) was originally devised to meet performance measurement challenges of private-sector firms. Although some of the struggles have been encountered in the quest for quality performance measurement, the greatest challenges to measure performance in the private sector is an almost exclusive reliance on financial measures of performance. Several concerns about the limitations of the excessive use of financial measures have been addressed in the most for-profit enterprises. Traditional measures of corporate performance focus on the bottom-line financial result derived from the short-term analysis of operational and financial data, providing irrelevant or misleading information. Tracking one single dimension of performance for a short period of time reflect neither an integrated or holistic view of organizational performance nor the demands of modern business environment in which the relationships with employees, customers, shareholders, and other key stakeholders can create the bulk of values. Balanced performance information is necessary to overcome the limitations of using single dimensional performance measures (Kaplan & Norton, 1992; Kaplan & Norton, 1996).

Along with the private-sector business, public organizations also produce products and services. Furthermore, the public organization's performance must take multiple values into account and is achieved in co-production with many stakeholders (Cole & Parston, 2006). Some public agencies have recognized many benefits of measuring their performance from a variety of perspectives and have begun to adapt the Balanced Scorecard approach to their circumstances (Northcott & Taulapapa, 2012). The City of Charlotte, North Carolina is the best-known example for an early adopter of the Balanced Scorecard system and its success in a local government organization (Niven, 2008). Many other recent success stories of the Balanced Scorecard in the public sector can be found in Defense Finance and Accounting Service (DFAS), Federal Aviation Administration Logistics Center, Department of Energy Federal Procurement System, and Department of Energy Federal Personal Property Management Program (https://balancedscorecard.org/bsc-basics/examples-success-stories/). It has been demonstrated that the Balanced Scorecard can fill some voids in the measurement efforts of public organizations (Moore & Braga, 2003a; Niven, 2008).

Some modifications have been suggested to facilitate the Balanced Scorecard in the public sector because public and private sector organizations have different organizational values and strategies (Moore & Braga, 2003a, Niven, 2008). For profit-seeking businesses, the most important bottom-line performance is to improve shareholder value, so the financial perspective is placed at the top of the Balanced Scorecard model (Kaplan & Norton, 1996). On the other hand, in spite of the fact that public organizations are also accountable for fiscal responsibility and stewardship, they place the Customer perspective at the top of the model. Financial resources are necessary for not only private- but also public-sector enterprises to successfully operate and meet customer requirements and needs, but financial measures in the public sector Balanced Scorecard model can be best seen as an indicator of customer success (Niven, 2008). It is not that performance measures are an end in themselves, but that the true value of performance measures can stem from comparing the results with the assumptions we make about the relationships among the measures within broader perspectives of the Scorecard model (Moore & Braga, 2003a). These measures link together in a chain of cause-and-effect relationships from the performance drivers in the financial perspective all the way through to improved customer outcomes in customer perspective (Niven, 2008). A well-designed Balanced Scorecard may help translate an organization's strategy through the objectives and measures toward the mission by making the relationships among the measures explicit (Amado et al., 2011). The linkage of measures throughout the Balanced Scorecard is constructed with a series of if-then statements:

A good example suggested by Niven (2008) shows, "If we increase our revenue, then we will have the resources to acquire the very best talent available. If we attract the best talent, then we will have the means necessary to develop and promote diverse offerings. If we develop and promote more diverse offerings, then we will be able to present them to the public" (p.38).

Data Envelopment Analysis (DEA)

Basic Concepts of DEA

Data envelopment analysis (DEA) is a non-parametric linear programming approach for measuring the relative performance of a set of comparable units, called Decision Making Units (DMUs), with multiple inputs and multiple outputs (Cooper et al., 2011). Incorporating multiple inputs and outputs often makes performance measurements and comparisons difficult (Davis, 2012; Maguire, 2004). One advantage of the DEA is that it has ability to process multiple elements and evaluate multi-criteria systems. It solves the problem by transforming the multiple inputs and outputs to a single virtual input and output and by assigning a set of weights where each unit has the highest possible efficiency rating in comparison to the other units. Consequently, the relative performance, especially efficiency, can be measured as: Relative Efficiency = a weighted average of the outputs / a weighted average of the inputs. DEA provides a comprehensive index to recap the interaction between measures of performance (Cook & Seiford, 2009; Cook & Zhu, 2005). Another advantage of DEA is that it requires very few assumptions. For example, it neither requires parametrized families of probability distributions nor a specification of a cost or production function. Although the use of inputs and outputs has a strong indication of a production function or process, they simply represent performance metrics for performance evaluation (and benchmarking against the best-practices). This feature allows for richer models and increases the acceptability of its results (Cook & Seiford, 2009; Cook & Zhu, 2005).

DEA is a best practice technique that determines baselines and benchmarks for performance improvements. Using all the data available (e.g., resources used, services provided, and other quantitative and qualitative information), DEA constructs an efficiency (or productivity) frontier, and a set of efficient and inefficient DMUs emerge by comparing their actual operating results. The efficiency (or productivity) frontier serves as an empirical standard of excellence. It is used to identify a best practice group of DMUs and determine which DMUs are inefficient compared to the best practice group and the amount of inefficiencies. One of the interesting features of DEA is that it allows each inefficient DMU to identify a benchmarking group that requires the same objectives and priorities but performs better. Also, it provides strong indications of what type or amount of changes in inputs and outputs are needed to make inefficient units efficient. Specifically, it can analyze by what percentage the inputs should decrease in order to achieve a given output level and by what percentage the output should increase given original levels of inputs in order to reach efficiency (Rickards, 2003). It clearly and objectively indicates performance improvement possibilities by measuring the efficiency of a particular DMU against a projected point on a productivity (or efficiency) frontier.

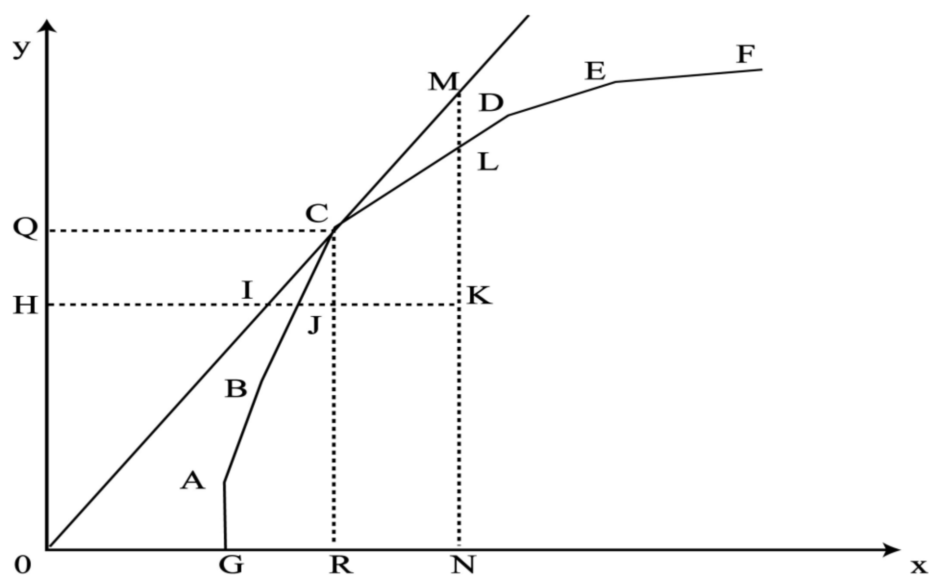
The original DEA was developed by Charnes, Cooper, and Rhodes (1978) (i.e., the CCR model), and extended by Banker, Charnes, and Cooper (1984) (i.e., the BCC model). To make a better comparison of the inefficient unit with the peer units, two different types of returns to scale

(RTS) was considered in both models, respectively. The CCR model assumes constant returns to scale, and the BCC model accounts for variable returns to scale.

Figure 1 below illustrates two different forms of efficient frontiers produced by the CCR and BCC models for a set of comparable units with a single input (x) and a single output (y). The units that have an efficiency rating of 1.0 are deemed as efficient and the convex envelope connecting them is called the efficient frontier. In Fig 4, the linear line (0ICM) represents the efficient frontier of CCR model with constant returns to scale while the convex line (GACDEF) depicts the efficient frontier produced by the BCC model with variable returns to scale. The units inside the efficient frontier are identified as inefficient and their relative efficiency rating is based on the distance from the efficient frontier. For example, the unit K is regarded as an inefficient, and the distance from the efficiency frontiers refers to the degree of the inefficiency of the unit K.

Figure 1

Technical and Scale Efficiencies in the CCR and BCC Model



In DEA two model orientations, the input-oriented and output-oriented model, are proposed for the inefficient unit to become efficient through resource conservation without detriment to its output or output augmentation without additional resources. In an input-oriented model, an inefficient DMU could reduce all its inputs simultaneously or equal-proportionally without sacrificing or reducing its outputs; while in an output-oriented model, an inefficient DMU could expand all its outputs simultaneously or equal-proportionally without increasing its input use.

Figure 1 also shows how to measure the overall, technical, and scale efficiencies through the single-input (x) and the single-output case. The efficiency scores obtained by solving the CCR and BCC models are referred to as overall efficiency (OE) and technical efficiency (TE), respectively. The scale efficiency (SE) is defined as the ratio of OE to TE. Under the input orientation model, for example, the technical efficiency of the unit K is *HI/HK* in the CCR model and *HJ/HK* in the BCC model. The scale efficiency of the unit K is obtained by the ratio *HI/HJ*. Under the output orientation model, in contrast, the technical efficiency of the unit K is provided by NK/NM in the CCR model and NK/NL in the BCC model. The scale efficiency of the unit K is provided by the ratio *NL/NM*.

DEA-Based Police Performance Studies

Several DEA-based police performance studies can be found in the literature. Several attempts have been made to fill some methodological gaps in the existing literature that can be further subcategorized into three groups: (1) controlling for non-discretionary inputs, (2) controlling for changes in the technical and technological efficiency over time, (3) combining additional methodologies with the DEA to improve the efficiency estimates.

Controlling for Non-Discretionary Inputs

Thanassoulis (1995) was the first to propose Data Envelopment Analysis (DEA) as a new approach for an assessment of police performance. He used an output-oriented CCR DEA model to analyze 41 police forces in England and Wales. While this study has contributed to this growing interest of research on measuring police performance through DEA, an important limitation to this pilot study needs to be acknowledged. He indicated that “many inefficient forces covered rather dissimilar policing environments to those of their efficient peers and so they were not strictly comparable to one another on performance” (Thanassoulis, 1995, p. 655-656).

In order to address this concern, several efforts have been reported to make comparisons more meaningful. Given the underlying models, introduced by Carnes et al. (1978) and Banker et al. (1984), did not consider non-discretionary inputs that are beyond their control (the so-called exogenous, environmental, or contextual factors), a number of researchers, such as Banker & Morey (1986), Ray (1991), Ruggiero (1996), and Muniz (2002), have developed models to control for these non-discretionary inputs for DEA analyses. Among them, the most applied model in the DEA-based police performance literature is the Ray’s (1991) two-stage approach.

Two-stage DEA. In the first stage, an original DEA technique is commonly used to measure the relative efficiency of decision-making units (DMUs). Then, in the second stage, the first-stage DEA scores are regressed on the inputs of interest beyond their managerial control. A regression model is estimated for the efficiency scores to examine the effect of environmental factors on the efficiency of DMUs.

Ruggiero (1998) argues that a second-stage DEA regression analysis makes it possible to disentangle inefficiency from environmental effects by adjusting the error term, leading to a measure of net technical efficiency. However, it should be noted that this approach requires a priori functional form specification for the second-stage regression. Misspecification leads to distorted measurement. The choice of regression model for the second-stage DEA analysis has been controversial and recently much disputed subject in the literature (Banker, Natarajan & Zhang, 2019; Hoff, 2007). Several regression approaches to second-stage DEA analysis have been applied, including ordinary least-squares (OLS), Tobit, and Truncated, Fractional regression analysis. These approaches have been echoed in policing literature.

Since its introduction by Ray (1991), a number of researchers have used a standard DEA model in the first stage and an OLS regression model in the second stage. Sun (2002), for example, used the two-stage process not only to measure the relative efficiencies of 14 Taipei City police precincts but also to examine the effect of environmental factors on the efficiency scores. Specifically, he estimated an output-oriented CRS and VRS DEA model in the first stage, followed by ordinary least-squared (OLS) regression for the second-stage DEA efficiency analysis.

Given DEA efficiency scores obtained from the first-stage DEA analysis are typically defined on the interval between 0 and 1, the most commonly used method to model DEA score is a two-limit Tobit model, with limits at zero and unity (Carrington et al., 1997; Drake & Simpler, 2005). By adopting input-oriented CCR and BCC models in the first stage and the Tobit regression in the second stage, Carrington et al. (1997) estimated the relative efficiency of the New South Wales (NSW) Police Service in 1994-1995.

Two earlier studies by Carrington et al. (1997) and Sun (2002) indicated that differences in operating environments (e.g., resident population and location factors) do not have a significant influence upon the efficiency of police patrols or precincts. Although Drake and

Simple (2005) found some evidence that environmental factors (i.e., socio-economic variables) may affect variation in the level of public's fear of crime as a police performance indicator, they failed to consider that not only is it important to identify the factors, but to include them in the comparative or benchmarking process.

In contrast to earlier findings, a number of researchers using a second-stage regression approach have reported environmental factors affect the technical efficiency of police agencies (Barros, 2007; Garcia-Sanchez, 2007; Hu et al., 2011). Barros (2007) employed second-stage Tobit regression models, in the second stage allowing efficiency scores to be regressed with socio-economic variables characterized by Lisbon police force precincts. He found that poverty levels were negatively associated with the precincts' efficiency. Also, Hu et al (2011) used Tobit regression to study the effects of socio-economic factors on the technical efficiency of regional police agencies in Taiwan. They identified three socio-economic determinants of police efficiency scores: the proportion of people aged from 15 to 64 years old, the rate of social increase to population, and the high education rate. The former two factors had significantly negative effects and the latter one had a significantly positive effect on the technical efficiency.

Three-stage DEA. Ruggiero (1998) reviewed existing methods for measuring efficiency while controlling for non-discretionary inputs that affect production. He highlighted potential limitations of these methods and developed a new approach to overcome existing weaknesses. He broke from the traditional regression approach of accounting for multiple environmental factors and presented an improved three-stage model to measure performance in the presence of multiple non-discretionary factors. In the first stage, a BCC model was applied with the given inputs and outputs. The resulting index captured not only TE but also the effects of environmental factors. In the second stage, following Ray (1991), ordinary least square (OLS) regression was used, with the index obtained in the first stage regressed on the environmental factors. In the third stage, a BCC model with an additional constraint was applied. This constraint restricts the comparison set by removing from the potential reference set any DMU having a more favorable environment than the DMU under analysis.

Since its introduction by Ruggiero (1998), his three-stage DEA analysis has become one of the two most commonly used approach for studying the influence of environmental factors on the efficiency of police organizations. Several researchers have attempted to identify the most appropriate regression model to control for the non-discretionary inputs, employing the three-stage models (Aristovnik et al., 2013; Garcia-Sanchez et al., 2013; Gorman & Ruggiero, 2008; Wu et al., 2010). Gorman and Ruggiero (2008) and Aristovnik and colleagues (2014) employed the three-stage approach and compare the OLS and Tobit regression in the second-stage analysis. The Tobit and OLS results are very similar, with Gorman and Ruggiero (2008, p.1035) arguing that these results are "virtually the same and the resulting correlation between environmental cost indices form these regressions was essentially one."

The Tobit regression has been widely used in many DEA studies when performance measures were regressed in the second-stage analysis. A small number of researchers, have argued that the Tobit regression is an inappropriate approach in that efficiency scores were not obtained from a censoring process, but were fractional data. It has been suggested by McDonald (2009), Simar and Wilson (2007), and Ramalho, Ramalho, and Henriques (2010) that the OLS, Truncated, and Fractional[2] regression models are a more consistent estimator than the Tobit model, respectively. In order to ensure fair comparisons among the 52 Spanish police forces, for example, Garcia-Sanchez and coworkers (2013) estimated the truncated regression model. Based on the results from the second-stage regression analysis, they considered several non-discretionary inputs, such as area, immigration rate, and youth index, which negatively affect police efficiency in their three-stage model.

Controlling for the Dynamics over Times

Measurement of intertemporal efficiency change has long been a subject of concern in DEA. The window analysis by Klopp (1985) was the first approach for this purpose. Based on Malmquist (1953), Fare, Grosskopf, Norris, and Zhang (1994) developed the Malmquist index in the DEA framework.

Very few studies have attempted to account for the dynamics of police efficiency scores over time. The exceptions are the works of Sun (2002), Barros (2006), Garcia-Sanchez et al. (2013), and Hadad, Keren, and Hanani (2015). Sun (2002) was the first study to emphasize a longitudinal efficiency analysis of the police over time. He used a DEA window analysis to analyze trends and potential stability problems over the 3-year period from 1994-1996. The window analysis allows to deal with the degrees of freedom problems in a case when "there is an insufficient number of DMUs in comparison to the number of relevant inputs and outputs in the model" (Cooper et al., 2011, p. 13). Given that his study involved 7 inputs and outputs to assess 14 police precincts with a three-year window, the technique seemed appropriate by treating each unit in a different year as a different unit and increasing the discriminatory power of the results. One advantage of the window analysis is that it avoids the problem of the degree of freedom (i.e., sample size), but it has a drawback assuming that there are no technological changes within each of the window.

Instead of the window analysis, some other authors used Malmquist index (MI) to evaluate the productivity change of a DMU (i.e., a police agency or precinct) between two time periods. The Malmquist index can break the productivity change down into two components:

technical efficiency (i.e., the product of "Catch-up" terms)[3] and technological progress (i.e., the product of "Frontier-shift" terms). The technical efficiency concerns the degree that a DMU attains for improving its efficiency while the technological progress reflects the change in the efficient frontiers surrounding the DMU between the two time periods (Cao and Yang, 2009). Estimating an input-oriented Malmquist productivity index, for instance, Barros (2006) examined the productivity growth in the Lisbon police force for the period 2000-2001. Garcia-Sanchez et. Al (2013) also used the MI combined with a multi-stage approach in order to control for both non-discretionary inputs and a dynamic evolution over the studied period. More recently, Hadad et al. (2015) developed a new measure by combining the DEA methodology with the Malmquist Index. Based on the new DEA/MI measure, they assessed all 13 police stations in the south of Israel and identified that "a police station with relatively low efficiency can be ranked high if it has a high improvement index" (Hadad et al., 2015, p.5). It is not only the efficiency but also the rate of improvement that should be considered when measuring the performance of police stations.

Ranking: Super Efficiency DEA vs Reference Set Frequency

Several additional methodologies have been combined with the DEA to sharpen the police efficiency estimates. DEA has become an increasingly popular method not only for performance evaluation, but also for benchmarking against best-practice. In a Sloan Management & Review article by Sherman and Zhu (2013), DEA is used for used for benchmarking in Operations Management, where a set of measures is selected to benchmark the performance of manufacturing and service operations… In the circumstance of benchmarking, the efficient DMUs, as defined by DEA, may not necessarily form a production frontier, but rather lead to a best-practice frontier.

DEA can determine the efficient frontier that consists of the best performing DMUs, and a projection to the frontier can provide valuable and practical information for the DMU managers. The identification of a role model (i.e., the target DMU) makes it possible that the managers of inefficient DMUs have a fair comparison, identify problems, and decide what to do (e.g., strategy).

From a typical DEA procedure, all efficient DMUs (i.e., the peer group / the reference set) will be given an efficiency rating of 100 (or 1.00) and so they cannot be differentiated in terms of the efficiency score (Verma & Gavirneini, 2006). To overcome this problem, additional procedures have been used to compare and rank the relative efficiency of all DMUs, including super efficiency DEA, reference set frequency, and a combined DEA/MI measure. Wu et al. (2010), for example, adopted the super efficiency score suggested by Anderson and Peterson (1993). Sun (2002) and Garcia-Sanchez (2007) used the frequency with which an efficient unit appears as a referent for inefficient ones. To see if the two methods gave the same measurement, Verma and Gavirneini (2006) used both procedures that seems to be complementary. Moreover, Barros (2006) and Hadad et al. (2015) take improvement trends (e.g., efficiency changes over time) into account by combing the Malmquist Productivity Index when they rank the relative efficiency of police agencies.

Integrating the BSC and DEA

In recent years, the Balanced Scorecard (BSC) has become one of the most prominent strategic planning and performance measurement system for many businesses and organizations in private and public sectors around the world (Amado et al., 2012). BSC has been selected by the editors of Harvard Business Review as one of the most influential business ideas of the past 75 years (Bibleet al., 2006). BSC provides management with a comprehensive picture of an organization’s performance from the four key perspectives such as financial, customer, internal business process, and learning and growth. Moreover, the BSC ensures an integrated strategic planning by clarifying an organization’s vision and strategy, aligning objectives, initiatives and key performance indicators (KPIs), and monitoring process and progress toward goals. It is an ongoing process that needs analysis, initiative identification, and successful implementation to keep the organization competitive (Kaplan & Norton, 1992; Kaplan & Norton, 1996).

Despite its advantages and widespread use, numerous authors have identified disadvantages of the BSC approach (Gomes & Romao, 2017). One of the disadvantages is that a large number of performance measures may confront managers with complex optimization problems (Fletcher & Smith, 2004). It lies in the fact that it does not specify a mathematical model of an objective weighting scheme (Rickards, 2007). The BSC analysis may fail to provide a comprehensive index to summarize the interaction between the key performance indicators embedded in the four perspectives of the BSC (Banker et al., 2004; Neves & Lourenco, 2008). It has also been argued that the BSC lacks a standardized baseline or benchmarks to compare performance (Banker et al., 2000; Lee, 2012). Without a benchmarking exercise, it is difficult to establish baselines, define best practices, identify improvement opportunities, and create a competitive environment within the organization.

Recently, several studies have been conducted to overcome these disadvantages of BSC, and researchers have found that data envelopment analysis (DEA) can complement the complexities of BSC (Amado et al., 2012; Eilat et al., 2008; Lee, 2012; Shafiee, Lotfi & Saleh, 2014). Data envelopment analysis (DEA), a mathematical programming technique, is used to evaluate the relative performance of DMUs with multiple inputs and outputs. This technique analyzes the efficiency of DMUs by computing a comparative ratio of weighted outputs to weighted inputs for each DMU. Through use of an objective weighting scheme for the construction of composite measures of performance, DEA can directly incorporate multiple inputs and outputs and provide a comprehensive index to encapsulate the key performance measures from the different performance dimensions. DEA uses all the data available to construct a best practice empirical frontier, to which each inefficient DMU is compared. Using a non-parametric linear programming approach for the estimation of efficiency frontier, DEA measures the efficiency of each DMU against a projected point on an efficient frontier and sets appropriate benchmarks for each DMU. Consequently, it clearly and objectively indicates which DMUs should be able to improve performance and the amount of resource savings and/or output augmentations that these inefficient DMUs must achieve to meet the level of efficiency of the best practice DMUs.

BSC and DEA are complementary to each other. On the one hand, BSC is a conceptual framework that can provide appropriate strategic measures of performance for DEA (Amado et al., 2012; Chiang & Lin, 2009; Lee, 2012). As Kaplan and Norton (1992) stated, 'what you measure is what you get' (p.71), it is important to choose a very small number of key performance indicators (KPIs) and to track them on a regular basis. By translating each strategic objective into one or two measures, an organization is able not only to avoid information overload but also to focus on the things that matter most. Thus, an organization can more easily measure and monitor progress towards strategic targets. On the other hand, DEA is a quantitative analysis tool for measuring the efficiency of DMUs (Chiang & Lin, 2009; Lee, 2012). Using multiple inputs and outputs based on the BSC indicators, DEA can calculate efficiency frontier and identify appropriate benchmarks for each DMU. DEA can transform performance measures into managerial information by estimating maximum output levels for given input levels and/or alternatively minimum input levels for given output levels in order to reach efficiency.

Incorporating the advantages of both the conceptual framework (BSC) and the quantitative analysis technique (DEA) into a comprehensive framework helps to provide structured information regarding the performance of each DMU and practical ways to improve it. Also, the BSC-DEA approach can measure the relative performance of DMUs through a balanced set of performance measures and trace the sources of performance results and performance changes. Moreover, the integrated BSC-DEA framework can determine the baseline and benchmarks to compare performance and reveal opportunities for reciprocal learning between DMUs or over time (Amado et al., 2012). Combining the BSC and the DEA approaches help to determine where there is room for improving organizational performance. According to Eilat and his colleagues (2008), the integrated BSC-DEA model address three common goals that organizations are trying to accomplish: (1) achieving strategic objectives – effectiveness goal; (2) optimizing the usage of resources in generating desired outputs – efficiency goal; and (3) obtaining balance – balance goal.

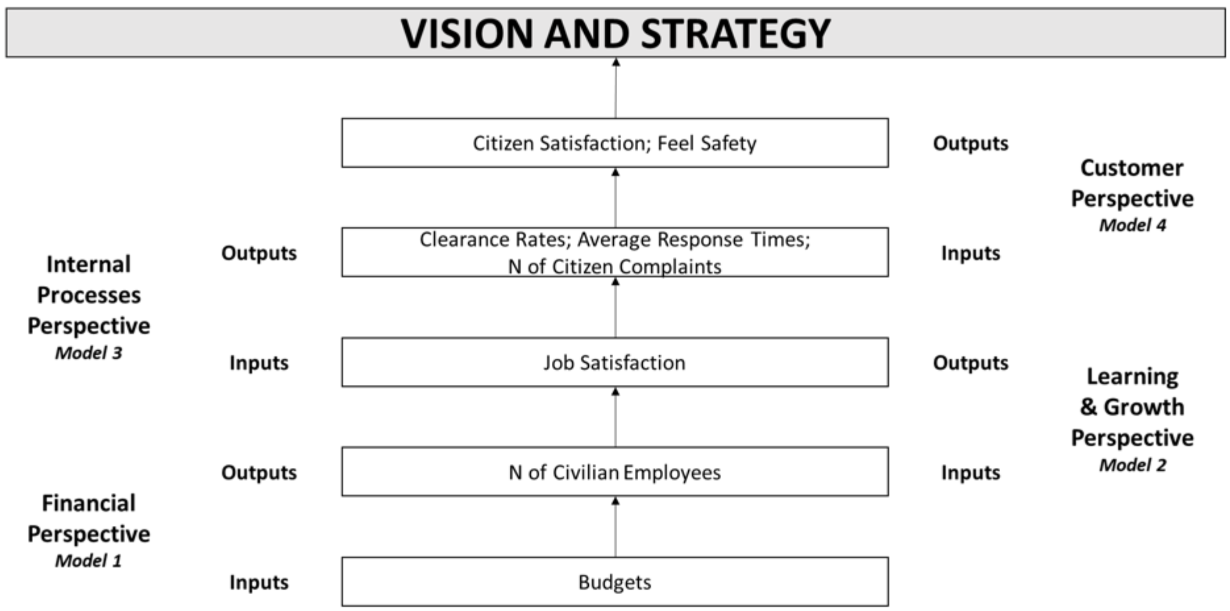
DEVELOPING A BALANCED SCORECARD FOR MEASURING THE PERFORMANCE OF A POLICE AGENCY

The original BSC approach, developed for an assessment of private-sector firms, does not make clear the order of the four perspectives of performance, but has more emphasis on the financial perspectives than the others. It does make sense in that the bottom line financial result is the reason for the existence of for-profit business. There have been many arguments regarding the relationships among those perspectives. Recently, Amado and colleagues (2012) developed a structured conceptual framework of the BSC by placing the financial perspective on the top, followed by customers, internal processes, and learning and growth. They assumed that "it is important for the department to improve its performance in the learning and growth perspective as this enables the department to improve its internal processes, which in turn enables improvements in customer satisfaction and subsequently create desirable results in the financial perspectives" (Amado et al., 2012, p.12).

Although an efficient use of public funds is important in evaluating the performance of the police, it cannot be the bottom line or the desirable outcome for police organizations. Many policing experts have argued that it should be the top priority to fulfil the needs and demands of customers (e.g., citizens and taxpayers as both customers and investor of police service and funds). In this respect, it is necessary to develop a conceptual framework that is distinct from the private sector but allows us to describe the interconnected relationships among the four perspectives of performance. Various interconnected models can be developed in the BSC framework for police organizations, but this study employs the well-structured conceptual model suggested by Niven (2008) in his book, *Balanced Scorecard: Step-by-Step for Government and Non-Profit Agencies*. By using if-then hypotheses and trade-offs between the models, four interconnected models are developed in Figure 2 by placing the customer perspective on the top, subsequently followed by internal processes, learning and growth, and financial perspective. Some evidence for these interconnected models and tradeoffs can be found in the literature.

Figure 2

Integrating DEA and BSC Approaches for Measuring Police Performance



Key Performance Indicators (KPIs) as Tradeoffs between Four Perspectives

Financial perspective has been often considered as the foremost driver of police performance assessment framework. The key objective of the financial perspective is how efficiently as well as fairly and effectively the police use public money to produce their valuable results[4]. The most commonly used financial indicator found in the literature is total expenditures in policing, mostly for personnel and equipment. The underlying assumption lies in the rationale that "more can be done with more." Some previous studies have examined the assumption that additional financial resources lead to higher level of performance in police agency (Choi, 2011; Xu, 2008; Zhao et al., 2003; Zhao et al., 2011;). Zhao et al. (2010), for example, examined the effects of additional resources (e.g., COPS grants) on police arrest, and indicated that "the hiring grant, the largest part of the COPS funding project, had consistently significant impact on police arrests after controlling for the socioeconomic variables and crimes" (p.165). One possible implication of this result is that additional financial resources or their efficient use lead to increases in the number of personnel and equipment, which in turn has a positive effect on police productivity (e.g., arrests, citations, and crime rates).

Police work by its very nature is labor intensive, and personnel costs account for approximately 85-90% of a police department's total operating budget in any U.S. agency (Swanson et al., 1998). Some efforts have been made to improve the efficient use of public resources: one of them is civilianization, the process of replacing sworn officers with nonsworn personnel for certain positions, such as dispatchers, research and planning specialists, crime-data analysts, and computer technicians (Walker and Katz, 2012). The civilianization effort makes it possible that police departments save money since in many cases civilian employees are less expensive than sworn officers (Bayley, 1994) and that sworn police officers focus on critical police work that requires a trained and experienced officer (Walker and Katz, 2012), which altogether contributes to professionalism in police force and enhance their operational efficiency (Maguire, 2003).

Increase in public funds or resources itself might not improve the performance of the internal process and customer perspective, but rather contribute to police officers' motivation and capabilities as core values in the learning and growth perspective in the progress toward the desired outputs and outcomes. In this sense, the number of civilian employees can be a good tradeoff indicator because it can be considered both as an output for the financial perspective and an input for the learning and growth perspective. The first interconnected model of this study assumes that additional resources (i.e., total expenditures) increase the number of personnel and equipment (i.e., civilian employees) which subsequently improve the motivation and capabilities (i.e., job satisfaction level) of police employees.

Learning and growth constitute the essential foundation for success of any knowledge-worker organization, including police organizations, and holds the key to future sustainable success. Nevertheless, the learning and growth perspective is often overlooked during the development of the BSC. A good example can be found in the policing literature: the most cited performance measurement tool in policing is the seven dimensions of police performance developed by Moore and Braga (2003), but this well-known framework does not include the employee learning and growth (related) perspective for the bottom line of policing. The current study attempts to fill some voids in the literature by incorporating measures in the learning and growth perspective into the police performance measurement framework.

According to Niven (2008, p.222), the learning and growth perspective should contain "measures relating to human capital (training, retention, succession), information capital (access to information), and a climate for positive action (communication, satisfaction, alignment)." Employee satisfaction has been the most widely used indicator to measure the success in the learning and growth perspective (Amado et al., 2012; Niven, 2008). Police organizations conduct opinion surveys to find out how police officers feel about their job. Administrators and managers are concerned about the job satisfaction of their employees. Job satisfaction is one of the widely studied and measured constructs in the organizational behavior and management literature because it is believed to have relationships with job-related variables (Rose et al., 2009; Judge et al,, 2002; Spector, 1997). Job satisfaction is positively associated with job performance and organizational commitment as well as physical health, psychological well-being and life satisfaction. In contrast, job satisfaction is negatively related to burnout, withdrawal behavior, and counterproductive behavior including absenteeism, turnover and perceived stress. In this respect, this study uses the officer job satisfaction as a tradeoff indicator between the learning and growth perspective and the internal processes perspective. Specifically, officers' job satisfaction level is considered as an output for the learning and growth perspective and simultaneously as an input for the internal process perspective.

The internal process perspective focuses on all the activities and key processes that drive values expected by the public. When developing measures for this perspective, it is important to identify those activities/processes and develop the best possible measures that enable to track organizational progress. The priority of the perspective is to continuously improve operational performance and achieve desirable organizational outcomes.

Previous studies on police performance have highlighted some traditional measures that are highly correlated with desirable policing outcomes. Some crime-focused indicators have been used as key measures of police performance, and some examples but are not limited to: crime rates, arrests, response times, and clearance rates. These measures are under the direct control of the police, so they can be directly and easily observed and evaluated. If the police quickly respond to citizens' calls for service, succeed in apprehending suspects with legitimate actions, and call them to account, we can say that the police are successful in producing justice. These activities and processes are valuable for evaluating police performance, but they cannot be an end in themselves. Some important duties of the police are hardly captured through crime statistics or CompStat. Many policing experts argue that measures of enforcement outputs should be used as the means to achieve other desired results. In other words, a set of desired social outcomes should be an ultimate basis for evaluating police performance. This study uses one measure the police may have more control over as a tradeoff indicator between the internal process perspective and the customer perspective. In detail, average response time is utilized as outputs for the internal process perspective and concurrently as inputs for customer perspective.

Lastly, the customer perspective is placed at the top of the integrated assessment framework of police performance. When choosing measures of the customer perspective, we must define our target customer and their demands. Citizens and taxpayers are customers of public funds and authority. Citizens and their representatives demand accountability from the police. In addition to results from other perspectives, the police have measured and reported progress in meeting the citizens' demand for accountability.

Citizen or community surveys are effective tools to provide more comprehensive information on police performance because they focus on socially desirable outcomes rather than policing outputs that are often derived from administrative data maintained by the police department. Citizen satisfaction and feelings of safety (or fear) surveys are two most widely used measures for the customer perspective from the policing literature. Choi (2011), for example, suggests four types of police performance based on two dimensions: the nature of performance indicator (i.e., output or outcome) and the objective of police strategy (i.e., crime reduction or citizen satisfaction). In his study, citizen trust in police was used as an outcome-citizen satisfaction related measure. Similarly, Lee (2013) considered citizen satisfaction as police performance. She empirically analyzed the impact of two policing strategies (e.g., traditional policing strategy and community policing strategy) on citizen satisfaction and indicated that traditional policing strategy had a negative effect on citizen satisfaction[5]. Besides the citizen satisfaction and feelings of safety survey, contact surveys are also useful in holding the police accountable for service to the public. Mastrofski (1999) has emphasized the importance of transactions with the public, and suggested six domains of performance indicators, such as attentiveness, reliability, responsiveness, competence, manners, and fairness. Maguire (2003, p.23) argues that "these kinds of surveys can be very revealing, particularly when they are focused on different kinds of contacts." This study uses two types of citizen surveys (i.e., citizen satisfaction with the police and feelings of safety survey) by integrating some benefits from contact surveys into the citizen surveys to better capture citizens' perception on socially desired outcomes.

DEA with Network and Dynamic Structures

Since the seminal work of Charnes, Cooper, Rhodes (1978) and Banker, Charnes, and Cooper (1984), traditional DEA models have been widely used as an effective method for measuring the relative performance of DMUs. The traditional DEA models are often considered as a "black-box" model because it does not take the underlying production processes into account in transforming multiple inputs into multiple outputs. That is, the traditional models only provide overall measures of performance by neglecting internal products or linking activities. As suggested by Amado et al. (2012, p.4), however, "overall measures fail to capture the efforts of different processes and sub-processes within the organization and might inhibit valuable managerial information." This limitation can also be found in that of the Balanced Scorecard. Without considering a series of internal processes among multidimensional perspectives of performance toward identified goals, the Balanced Scorecard cannot obtain useful information for performance improvement.

The multidimensional nature of performance can be best captured using several DEA models, attempting to move away from a black box and capture the dynamics of the transformation processes and sub-processes within the organization. Network DEA models have been developed for considering the internal processes and structures, allowing for measuring process or divisional efficiencies as well as overall efficiency in a unified framework. Their models were first proposed by Fare and Grosskopf (1991) and then extended by several authors, with many theoretical developments and practical application being reported. A comprehensive review can be found in Cook, Liang, and Zhu (2010) and Kao (2014). In his recent review, Kao (2014) classified not only the family of Network DEA models into the following nine types: (1) Independent model; (2) System distance measure model; (3) Process distance measure model; (4) Factor distance measure model; (5) Slack-based measure model; (6) Ratio-form system efficiency model; (7) Ratio-form process efficiency model; (8) Game theoretical model; and (9) Value-based model; but also the network structures into the following seven types: (1) Basic two-stage structure; (2) General two-stage structure; (3) Series structure; (4) Parallel structure; (5) Mixed structure; (6) Hierarchical structure; and (7) Dynamic structure.

In terms of the classification of the model and structure, various network DEA models have been developed and they are complementary to each other. Choosing an appropriate model is an important step in meeting the defined goals and objectives through the network DEA. This empirical study employs the Dynamic-Network (DN) DEA models within which Tone and Tsutsui (2009, 2010, 2014) have incorporated the concepts of (1) slack-based measures, (2) series structure, and (3) dynamic structure.

Slack-based Measure Model (SBM)

Tone and Tsutsui (2009) proposed a slack-based measure (SBM) model to measure divisional or process efficiencies along with the overall efficiency of the network system. The slack-based measure approach has a number of attractive features: the first advantage is to use a non-radial method for measuring efficiencies when inputs and outputs may change nonproportionally. The second advantage is to employ the weighted slack-based measure approach that takes the importance of each division or process into account. The last advantage is to decompose the overall efficiency into divisional or process ones.

Series Structure

The series structure refers to a number of processes connected in sequence, where each process consumes the exogenous inputs and intermediate products produced by the preceding process, and produces exogenous outputs and intermediate products for the succeeding one to use. Although a series system can have as many processes as desired, except for theoretical studies, the largest system that has appeared in the literature has only five processes.

Tsutsui and Goto (2009), for example, applied a weighted SBM model to evaluate the performance of multi-functional, vertically integrated electric power companies in the United States. Assuming that five processes (e.g., generation, transmission, distribution, sales, and general administration) are involved in supplying electricity to customers, they estimated the process efficiencies seeking to optimize the overall management efficiency.

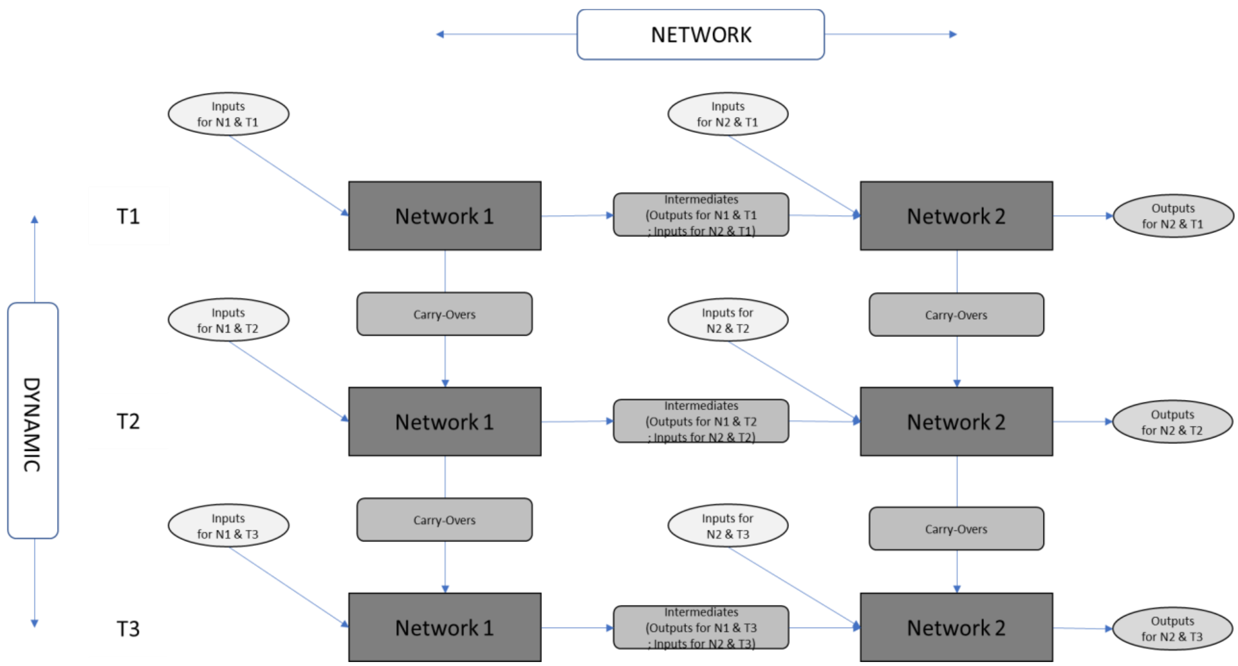
Dynamic Structure

Dynamic structures concern the repetition of a single-period system connected by carryovers, where the single-period structure can be any of those discussed in the preceding subsections. Structurally, the dynamic structure resembles the series one in that each period can have a network structure. Physically, the dynamic structure is a special type of series structure, because the inputs, outputs, and intermediate products (carryovers) are the same for every period.

Tone and Tsutsui (2010) developed a dynamic model by transforming their SBM models with series-network structures into those with dynamic-network structures. The dynamic SBM model incorporates carry-over activities between two consecutive terms into the DEA model, instead of linking activities between two interrelated divisions or processes in the original network SBM model. This model makes it possible to measure the overall efficiency of DMUs as well as period specific efficiency based on the long time optimization during the whole period.

Figure 3

Dynamic Model with Network Structure: A Slack-Based Measure Approach



A Combined Model: A Dynamic-Network DEA

Tone and Tsutsui (2014) have attempted to integrate the network and dynamic models in the slack-based measure (SBM) framework, and proposed a combined model, called a dynamic-network (DN) data envelopment analysis model. This combined model enables us to measure (1) the overall efficiency over the entire observed period, (2) dynamic change of period efficiency, and (3) dynamic change of divisional efficiency. In addition, each DMU has carry-over variables that take into account a positive or negative factor in the previous period. The model can be implemented in input-, output-, or non-oriented forms under the CRS or VRS assumptions on the production possibility set.

METHODOLOGY

Research Site: Seoul, South Korea

Seoul, formally the Seoul Special City, is the capital and largest metropolis of South Korea. With an estimated 2015 population of 9,904,312 distributed over 233.69 square miles (605.25 km^2), Seoul is the most populous and densely populated city in South Korea. Seoul has been described as the political, financial, cultural hub of South Korea. As of 2015, 50.3 percent of the city population was female, and 80 percent of the population was 15 years old or older. Moreover, 58.1 percent of the economically active population (aged 20 to 49) held a college degree or post-baccalaureate degree. More than half of the population were married (51.5%), and there were 274,957 foreigners residing in the city (2.7%). Finally, 4.7% of the working age population (aged 16 to 64) was unemployed, and the average monthly household income was approximately US$4,000 (Seoul Statistics, 2020).

Seoul is comprised of 25 autonomous districts (i.e., *GU*) and 424 administrative neighborhoods (i.e., *DONG*). Each district in Seoul has been governed independently by a democratically elected mayor and legislative body since the 1995 implementation of the local autonomous system in South Korea. Table 2 describes the population and area of each district in Seoul. All districts in Seoul are largely urbanized (Kim & Han, 2012). Figure 4 shows the map of 25 administrative districts of Seoul. Most of the administrative districts coincide with police jurisdictions, but there are two police stations in the shaded six districts (in Figure 5).

Table 2

The Population and Area of Districts in Seoul (2015)

Districts	Population	Area	Population Density	Dong
Jung	134,409	9.96	13,494	15
Jongno	161,922	23.91	6,771	17
Seodaemun	325,871	17.61	18,506	14
Yongsan	245,102	21.87	11,209	16
Seongbuk	461,617	24.58	18,781	20
Dongdaemun	370,312	14.22	26,050	14
Mapo	390,887	23.84	16,394	16
Yeongdeungpo	406,779	24.53	16,583	18
Seongdong	307,161	16.86	18,218	17
Dongjak	413,247	16.35	25,269	15
Gwangjin	372,104	17.06	21,807	15
Gangbuk	330,704	23.6	14,015	13
Geumcheon	254,654	13.02	19,560	10
Jungnang	415,677	18.5	22,474	16
Gangnam	572,140	39.5	14,484	22
Gwanak	525,607	29.57	17,776	21
Gangseo	602,104	41.44	14,531	20
Gangdong	448,471	24.59	18,238	18
Guro	449,600	20.12	22,347	15
Seocho	451,477	46.98	9,610	18
Yangcheon	281,845	17.41	27,681	18
Songpa	664,946	33.88	19,629	27
Nowon	571,212	35.44	16,119	19
Eunpyeong	495,937	29.7	16,696	16
Dobong	350,272	20.67	16,948	14

Figure 4

Administrative Districts of Seoul (N=25)

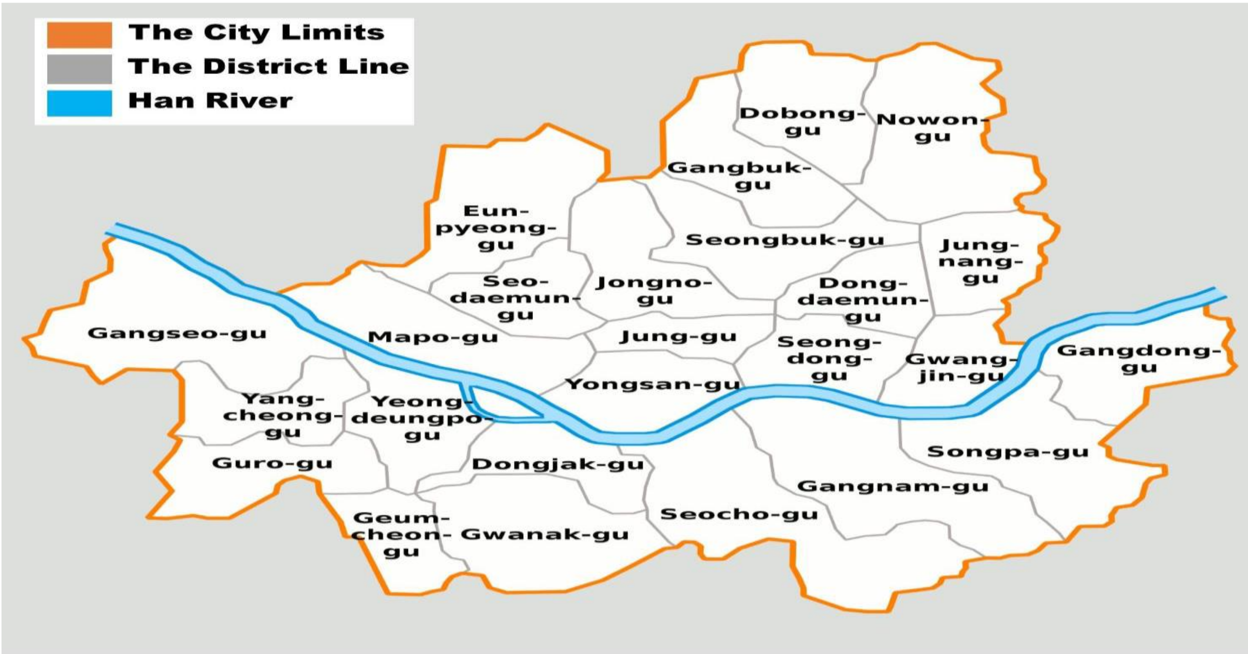
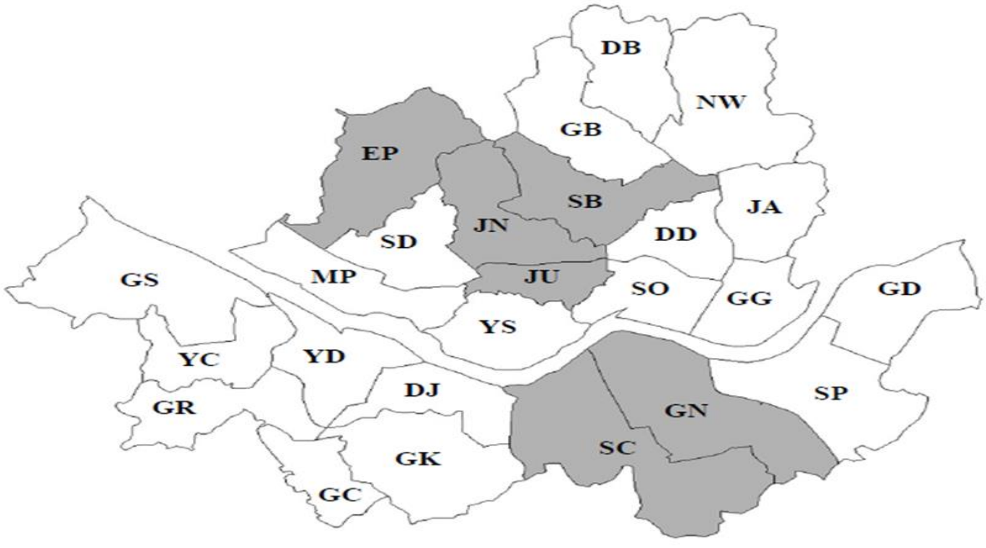


Figure 5

Police Jurisdictions in Seoul (N=31)



The territorial jurisdiction of most police stations coincides with the boundaries of the 25 administrative districts of Seoul, with the exception of the following six districts: Gangnam (GN), Seocho (SC), Eunpyeong (EP), Jongno (JN), Jung (JU), and Seongbuk (SB). Figure 5 shows the map of 31 police stations across the 25 administrative districts. The shaded districts represent that two police stations are operating in each district to provide more effective and efficient services. Table 3 presents the number of police officers, jurisdiction size, population served, and reported number of major crimes of each police station in the SMPA (Police Statistical Yearbook, 2016).

Table 3

Descriptive Statistics of 31 Police Stations in the SMPA (2015)

Police Stations	Police Officers	Jurisdiction Size	Population Served	Crimes
Jungbu	497	7.2	108,561	2,698
Jongno	620	18.6	82,565	2,158
Namdaemun	472	2.8	19,400	2,098
Seodaemun	648	17.6	309,935	3,876
Hehwa	468	5.3	74,191	2,471
Yongsan	663	21.9	235,838	3,485
Seongbuk	518	16.6	259,117	2,244
Dongdaemun	765	14.2	363,444	3,956
Mapo	766	23.9	386,677	5,229
Yeongdeungpo	890	24.6	382,734	6,441
Seongdong	624	16.9	295,959	3,326
Dongjak	635	16.4	407,340	3,576
Gwangjin	657	17.1	363,182	5,749
Seobu	388	8.4	204,257	2,207
Gangbuk	621	23.6	334,995	3,709
Geumcheon	534	13	256,430	3,558
Jungnang	728	18.5	418,629	4,873
Gangnam	768	14.2	208,648	4,925
Gwanak	776	28.9	495,603	6,274
Gangseo	809	41.4	585,951	4,917
Gangdong	704	24.6	475,929	5,037
Jongam	396	8	207,237	2,007
Guro	732	20.1	425,679	4,992
Seocho	641	36.9	297,888	3,731
Yangcheon	723	17.4	486,559	4,286
Songpa	991	33.9	664,709	7,392
Nowon	763	35.4	582,060	4,892
Bangbae	349	10.1	152,502	1,365
Eunpyeong	467	21.3	294,645	2,818
Dobong	525	20.7	353,509	2,914
Suseo	583	25.3	370,342	3,057

Decision Making Units (DMUs): Police Stations within the SMPA

In DEA, the term Decision-Making Unit (DMU) refers to any entity that is to be evaluated in terms of its abilities to convert inputs into outputs. DEA was originally developed for evaluating the relative performance of a set of peer entities called DMUs, which perform the same function in terms of inputs they use and outputs they produce. DEA has been successfully applied to many different types of law enforcement agencies engaged in a wide variety of activities in many contexts worldwide. Nevertheless, Korean police have rarely used this technique for evaluating their performance.

The Korean police system is ideal for DEA due to its organizational structure and uniformity in operating policies. South Korea adopts a highly centralized police system. It has only one national police force, the Korean National Police (KNP), which is responsible for stablishing and managing all of the police services for the entire nation. The KNP is divided into 16 regional headquarters (i.e., metropolitan and provincial police agency), each of which is further subdivided into police stations and police substations (i.e., mini-police stations). Figure 6 shows the hierarchical structure of the KNP, referring to the organization's chain of command. The Korean National Police exerts strong leadership over local police forces and promotes common standards. Police agencies in South Korea are similar in function, purpose, size, and training.

When using the proposed performance measurement system in other police agencies around the world, performance dimensions and possible indicators will likely change as a result of cultural differences in both the police agencies and societies, and also as a result of the variation in the structure of policing systems across various countries. For example, police stations are more homogenous in South Korea, while police agencies in the United States are not standardized in any way and are highly fragmented.

Figure 6

The Hierarchical Structure of KNP

The SMPA is the most important component of the KNP. The SMPA is the primary law enforcement agency for the city of Seoul. Since it was established in 1946, the SMPA has become the largest police force in the country. It employs around 26,000 uniformed police officers and has 31 police stations covering the entire city. The SMPA plays a disproportionately important role. It is responsible for about one-fifth of the Korean population but face about 50 percent of all violent crimes across the nation. The SMPA has the heaviest responsibility for dealing with difficult order maintenance problems and serious crime, which are disproportionately concentrated in cities. Also, it is asked to provide a wide range of emergency services.

The main focus of this study is on police stations. Under the control of the SMPA, there are 31 police stations, enabling fair comparisons and benchmarking. The KNP has evaluated their agency-level performance at the police station level. Also, many DEA-based police performance studies have used police stations or precincts as their DMUs (See Table 4). For example, Akdogan (2012) used DEA to measure the efficiency of police stations in the city of Ankara in Turkey. More recently, Aristovnik and colleagues (2014) measured the relative efficiency of police stations in Slovenia. It has been argued that police stations or precincts can act as a viable and important sub-organizational level of analysis in police organizations (Hassell, 2007; Klinger, 1997; Taniguchi & Salvatore, 2018). Hassell (2007) found that police practices vary at the station/precinct level of analysis. Klinger's ecological theory of policing (1997) provides a theoretical explanation for understanding how police behavior varies across police precincts and how crime patterns develop and are sustained in local communities.

Local police stations provide the lion's share of police service in South Korea. At the local level, these stations strive to enhance the quality of life by enforcing the law, maintaining order, and providing miscellaneous services to the citizens on a day-to-day basis. Since they patrol their precincts and respond to requests for service, they are most visible to the citizens and have the most direct contact with them (Aristovnik et al., 2014). In this respect, local police stations appear to be the appropriate DMUs for measuring the agency-level police performance.

Table 4

Previous DEA-based Police Performance Studies

Authors	Country	Unit of Analysis / Decision-Making Units	Data
Hadad et al. (2015)	Israel	Police Station (Local Level)	All 13 police stations in the south of Israel over 4 years
Aristovnik et al. (2014)	Slovenia	Police Station (Local Level)	76 (local) organization units among 128 units in 2010
Aristovnik et al. (2013)	Slovenia	Police Directorates (Regional Level)	11 Police Directorates (PDs) in 2005 and in 2010
Garcia-Sanchez et al. (2013)	Spain	Police Directorates (Regional Level) - Provincial Level	52 Spanish provinces in the 2001-2006 period
Garcia-Sanchez (2009)	Spain	Local Level - Town Level	113 towns on the mainland with over 50000 in habitants in the year 2000
Garcia-Sanchez (2007)	Spain	Police (Regional Level) - Provincial Level	52 cities that represent all of the provincial capitals in Spain in 1999
Akodogan (2012)	Turkey	Police Station (Local Level)	20 Police Stations in the City of Ankara in 2006 (Not all police stations were included)
Hu et al. (2011)	Taiwan	Police Directorates (Regional Level)	23 regional police agencies in Taiwan from 2003 to 2007
Wu et al. (2010)	Taiwan	Police Directorates (Regional Level)	22 administrative districts across the entire island (out of 25 administrative districts and 7 cities)
Rodgers (2008)_Di	USA	Police Station Department (Local Level)	61 police departments in the State of Texas, where the population exceeded 50000 in year 2008
Gorman and Ruggiero (2008)	USA	Police Directorates (Regional Level)-States	49 continental states for the year 2000
Gupta et al. (2008)	India	Police Directorates (Regional Level)-States	All 28 states and 7 union territories of India for year 2006 under 12 crime heads
Barros (2007)	Portugal	Police Station (Local Level)-Lisbon police force precincts	33 Lisborn police precincts for the period 2000-2002
Barros (2006)	Portugal	Police Station (Local Level)-Lisbon police force precincts	panel data for the years 2000, 2001 and 2002 on 33 precincts
Verma and Gavirneni (2006)	India	Police Directorates (Regional Level)-States	All the 25 states in India for the year 1997
Goltz (2006)_Di	USA	Police Station (Local Level)	342 municipal and county police organizations in the state of Florida for the year 2005

Table 4. *Previous DEA-based Police Performance Studies (Continued)*

Authors	Country	Unit of Analysis / Decision-Making Units	Data
Drake & Simpler (2005)	England and Wales		41 police forces in England and Wales
Drake & Simpler (2003)			
Drake & Simpler (2002)			
Drake & Simpler (2001)	England and Wales		England and Wales police force
Drake & Simpler (2000)	England and Wales		England and Wales police force
Sun (2002)	Taiwan	Police Station (Local Level)	the 14 police precincts in Taipei city for the years 1994, 1995, and 1996
Diez-Ticio & Mancebo´n (2002)	Spain	Police Station (Local Level)	47 Spanish metropolitan police forces in 1995
Nyhan & Martin (1999)	USA	Regional Level	36 county and municipal police forces; 20 major US cities
Carrington et al. (1997)	Australia	Police Precincts/Patrols (Local Level)	NSW 163 police patrols for the period 1994-1995
Thanassoulis (1995)	England and Wales	Regional Level	41 police forces in England and Wales using data for the years 1992–1993

Data

The data used in this empirical investigation concerns the 31 police agencies under the SMPA from 2013 to 2015 in a balanced panel. Data from this time frame are used because more recent data was not available at the time that this study began. All the data were secondary and collected from annual reports and surveys of the KNPA. The primary source of the data comes from the Police Statistical Yearbook, the Police Job Satisfaction Survey, the Police Customer Satisfaction Survey and the Public Safety Survey.

First, the Police Statistical Yearbook is an annual publication that contains crime and police data at the level of local police agencies (i.e., police stations in South Korea); it includes arrests, clearances, response time, budgets, and law enforcement employee data. It is used to provide some traditional measures for assessing the financial and internal business perspectives, which previous studies have focused on to measure the performance of police forces.

Second, the Police Job Satisfaction Survey is a nationally representative police opinion survey that aims to provide comprehensive information on police officers' job satisfaction and work attitudes related to various aspects of their job. The survey is distributed to all sworn officers nationwide by intranet (a network of computer within the KNPA) for two weeks (e.g., Apr 15 to Apr 28 in 2013; Apr 28 to May 11 in 2014; Apr 27 to May 10 in 2015). The response rate is approximately 40% (e.g., 33.7% in 2013; 42.6 % in 2014; 40.8% in 2015). It is used to measure the learning and growth perspective.

Finally, community surveys such as the Police Customer Satisfaction Survey and the Public Safety Survey are also used to assess the customer perspective of police performance. Community surveys can provide police with reliable feedback from citizens about perceptions of police performance. Because in-house survey may make results suspect, community surveys have been conducted by research organizations in which trained personnel and resources such as telephones and computers are allocated and located. To avoid certain unforeseen circumstances where a major crime event occurs during the survey period, efforts have been made to carry out the survey as quickly as possible. Survey administrators (e.g., usually a managerial level of police personnel) monitor adherence to a series of quality control procedures such as interview time and manner on a random basis. Sample size for the Police Customer Satisfaction Survey ranges from 14,000 in 2013 to 36,420 in 2015; and for the Public Safety Survey is about 9,000 during the year 2013-2015.

These data all together are used by Congress to affect police policy decisions and allocate funding to police force, so the data should be accurate. Given that the DEA does not account for measurement errors in the data, the correct variable specification is required (Kawaguchi et al., 2014).

Variables

Most of the previous studies, which have evaluated the organizational performance of the police forces by adopting a DEA method, have focused on the products of internal business process; that is, crime control in the field of policing. These studies usually adopt not only the traditional DEA model (BB model) but also a limited set of police and crime data. Table 5 below gives a brief summary of DEA models and measures (e.g., inputs, outputs, and other variables) used to measure police organizational performance in the literature. The capital letter E in the parenthesis (i.e., (E)) refers to environmental factors that are usually considered as non-discretionary inputs in DEA.

As seen in the Table 5, the previous studies do not consider performance measures and products for the financial, learning and growth, and customer perspectives. However, the DN

DEA model enable us to consider the measures and products in line with the four interconnected perspectives of the BSC. We can observe the activities or products of all the perspectives separately from the internal business process perspective.

Table 5

Inputs and Outputs Used in the Previous DEA-based Police Performance Studies

Authors	Inputs	Outputs
Hadad et al. (2015)	Annual number of property crime, violent crime, burglaries, and traffic accidents with injuries Annual total cots (personnel and operational costs) Population; Number of vehicles	Annual "clear-up" rate for property crime Number of drunken driving cases exposed Number of traffic reports
Aristovnik et al. (2014 & 2013)	Annual number of occupied employment posts; Material resources (e.g., equipment); Number of workstations and police vehicle radio stations (E) Criminal offenses; Violation of public order regulations; Road accidents	Solved criminal offences Road accidents involving serious injury and minor injury Average response time of police patrol Use of instruments of restraint and warning shots
Garcia-Sanchez et al. (2013)	Number of police officers Number of vehicles (E) Unemployment rate; Immigration; Population; Young people; Area	Percentage of solved crimes (property, person, sexual, and safety)
Garcia-Sanchez (2009 & 2007)	Total staff Number of vehicles	Number of Kilometers traveled by police vehicles; Number of those arrested taken before the count; Number of objects recovered; Number of interventions made Number of accusations formulated; Number of vehicles removed from the public highway; Number of breathalyze tests carried out; Number of accident reports drawn up
Akodogan (2012)	Number of personnel; Number of police cars (E) Population; Area; Number of critical entities (e.g., school) ; Number of incoming documents (both judicial and managerial); Number of incidents	Number of processed judicial and managerial documents; Number of outgoing documents Number of solved incidents
Hu et al. (2011)	(E) Proportion of people aged from 15 to 64 years old; Higher education rate; Rate of social increase to population	Rate of violent crime, larceny crime, other crimes Clearance rate of violent crime, larceny crime, and other crimes

Table 5. *Inputs and Outputs Used in the Previous DEA-based Police Performance Studies (Continued)*

Authors	Inputs	Outputs
Wu et al. (2010)	Labor cost; General running and operating cost; Equipment purchasing cost	Number of "clear-up" crimes of burglaries, violent crimes, and other crimes Number of road traffic accidents Number of general and special services
Gorman and Ruggiero (2008)	Number of sworn officers and other employees Number of vehicles (E) Percentage of single mothers; Property rates; Percentage of individuals in the labor force; Population	Rates of murders, other violent crimes, and total property crimes
Gupta et al. (2008)	Civil and armed police force strength Total police expenditure	Number of persons arrested Crime rates
Barros (2007 & 2006)	Number of police officers; Number of police cars Cost of labor and others (E) Number of crimes (theft, burglary, car robbery, and drug)	Percentage of "clear-up" crimes (theft, burglary, stolen cars, and drug) Number of raids, stop operations, and minor offences with fines
Verma and Gavirneni (2006)	Total expenditure Number of police officers and investigating officers (E) Total number of investigated cases	Number of persons arrested, charge sheeted, and convicted Number of trials completed
Drake & Simpler (2005)	Number of burglaries, vehicle crimes, and robberies Total budget	Civilian days lost Aggregate offenses cleared
Drake & Simpler (2003 & 2002 & 2001 & 2000)	Employment costs Premises-related expenses Capital and other costs	Clear up rates Number of traffic offenses Number of breathalyzer tests administered
Sun (2002)	Number of police officers Number of burglaries, offence crimes, and other crimes (E) Location of a police precinct; Jurisdiction area of a police precinct; Population; Population of young people	Number of "clear ups" of burglaries, offence crimes, and other crimes
Diez-Ticio & Mancebo´n (2002)	Total manpower Vehicles (E) Population	Clear up rate of solved violent and property crime offenses

Table 5. *Inputs and Outputs Used in the Previous DEA-based Police Performance Studies (Continued)*

Authors	Inputs	Outputs
Nyhan & Martin (1999)	Total department costs Total personnel (sworn officers and civilians) (E) Population; Median income	Total crimes Response time Crime clear up rate
Carrington et al. (1997)	Number of police officers and civilian employees Number of police cars	Number of offences, arrests, summons, major car accidents Kilometers traveled by police cars
Thanassoulis (1995)	Number of police officers Number of violent crimes, burglaries, and other crimes	Number of "clear ups" of violent crimes, burglaries, and other crimes

The selection of appropriate inputs, outputs, and other variables to model the police agencies' performance is an important but complicated task. Based on the literature review and the data availability, the inputs, outputs, links and carry-overs of the DN model are selected and described in Figure 7 and Table 6 below.

Figure 7

An Integrated Police Performance Framework

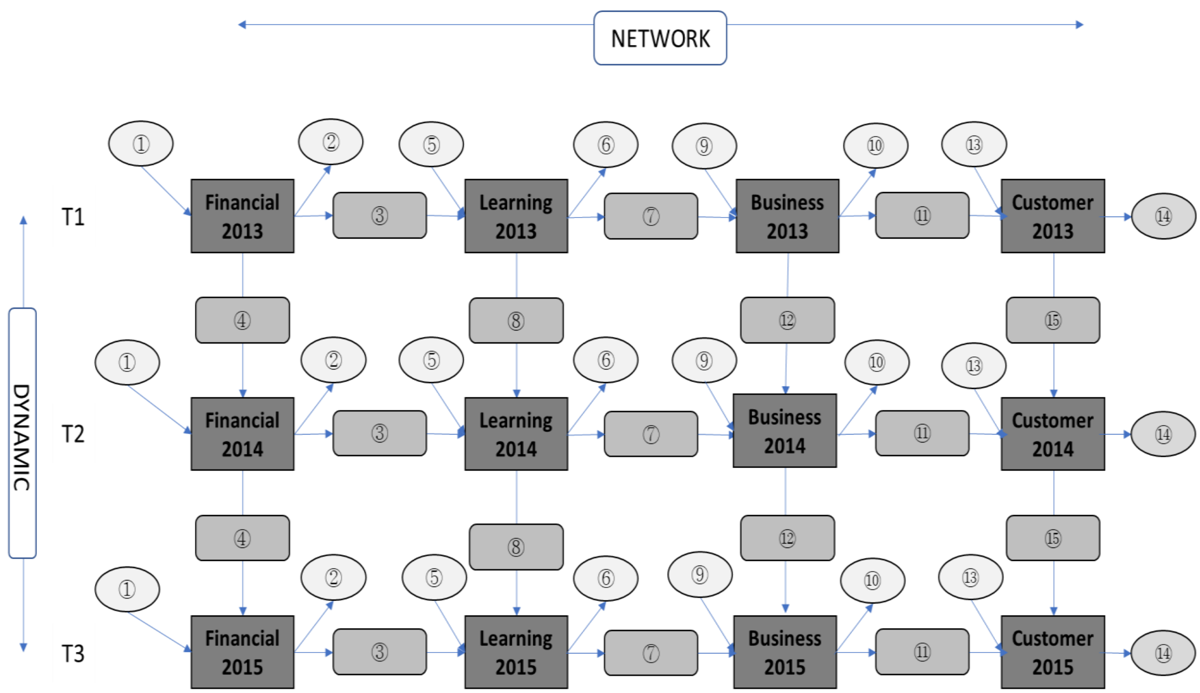


Table 6

Inputs, Outputs, Links and Carry-overs of the DN Model

Variables names		
Financial	Input	① Budget
	Output	② Number of police officers
	Link	③ Number of civilian employees
	Carry-over	④ Carry-over budget
Learning & Growth	Input	⑤ Number of equipment
	Output	⑥ Officers' perceptions of learning environment
	Link	⑦ Officers' job satisfaction
	Carry-over	⑧ Officers' perceptions of learning opportunities
Internal Business	Input	⑨ Number of crimes occurred
	Output	⑩ Number of crimes cleared
	Link	⑪ Average response time to calls for service
	Carry-over	⑫ Number of crimes unsolved
Customer	Input	⑬ Number of citizen crime prevention groups
	Output	⑭ Citizens' satisfaction of police services
	Carry-over	⑮ Citizens' feelings of safety

The definition, operationalization, and data sources for the selected variables are escribed below.

Dimension I: Financial Perspective

Measure 1 (Input): Budget

Definition: *Budget* is the cost of policing at the police station level. It includes personnel expenditures (such as salaries, benefits, and overtime for officers and civilian employees), operating costs, and capital spending.

Operationalization: This indicator was operationalized as the US dollars.

Source: Police Management and Administrative Statistics (PMAS)

Measure 2 (Output): Number of police officers

Definition: the number of sworn police officers, representing the police officer strength in each of the police stations under the SMPA.

Operationalization: Use the original data in the PMAS, without operationalization

Source: Police Management and Administrative Statistics (PMAS)

Measure 3 (Link): Number of civilian employees

Definition: the number of non-sworn police officers, representing the degree of the civilianization in each of the police stations under the SMPA.

Operationalization: Use the original data in the PMAS, without operationalization

Source: Police Management and Administrative Statistics (PMAS)

Measure 4 (Carry-over): Carry-over budget

Definition: *Carry-over budget* is the funds unused during a financial year which are transferred to the budget for the following year.

Operationalization: This indicator was operationalized as the US dollars.

Source: Police Management and Administrative Statistics (PMAS)

Dimension II: Learning and Growth Perspective

Measure 1 (Input): Number of equipment

Definition: the number of patrol cars and motorcycles, facilitating organizational learning through the use of advanced technology in patrol cars and motorcycles

Operationalization: Use the original data in the PMAS, without operationalization

Source: Police Management and Administrative Statistics (PMAS)

Measure 2 (Output): Officers' perceptions of learning environment

Definition: the degree of organizational environment for learning

Operationalization: This indicator was measured by a single item, 'How satisfied are you with the learning environment your affiliated police station provides?' Response ranged from 1 (very satisfied) to 7 (never satisfied). The 7-point Likert scale was converted to a 100-point scale.

Source: Police Job Satisfaction Survey (conducted biannually, Use the survey result conducted in the first half of the years 2013-2015).

Measure 3 (Link): Officers' job satisfaction

Definition: the extent to which police officers feel satisfied with their job

Operationalization: It is presented in Appendix * how this indicator was measured and operationalized.

Source: Police Job Satisfaction Survey (conducted biannually, Use the survey result conducted in the first half of the years 2013-2015).

Measure 4 (Carry-over): Officers' perceptions of learning opportunities

Definition: the degree to which police officers have opportunities for any course of study, education or training made available by the affiliated police station (as a learning provider)

Operationalization: This indicator was measured by a single question, 'Do you think your affiliated police station provide learning opportunities for your duties and self-improvement?' Response ranged from 1 (very satisfied) to 7 (never satisfied). The 7-point Likert scale was converted to a 100-point scale.

Source: Police Job Satisfaction Survey (conducted biannually, Use the survey result conducted in the first half of the years 2013-2015).

Dimension III: Internal Business Perspective

Measure 1 (Input): Number of crimes occurred

Definition: the number of major crimes (such as murder, robbery, rape, theft, and violence) occurred and reported to each police station

Operationalization: Use the original data in the PMAS, without operationalization

Source: Police Management and Administrative Statistics (PMAS)

Measure 2 (Output): Number of crimes cleared

Definition: the number of major crimes (such as murder, robbery, rape, theft, and violence) cleared among the reported major crimes

Operationalization: Use the original data in the PMAS, without operationalization

Source: Police Management and Administrative Statistics (PMAS)

Measure 3 (Link): Average response time calls for service **(Undesirable)**

Definition: Use the original data in the PMAS, without operationalization

Operationalization: Use the original data in the PMAS, without operationalization

Source: Police Management and Administrative Statistics (PMAS)

Measure 4 (Carry-over): Number of crimes unsolved **(Undesirable)**

Definition: the number of major crimes (such as murder, robbery, rape, theft, and violence) unsolved among the reported major crimes

Operationalization: Use the original data in the PMAS, without operationalization

Source: Police Management and Administrative Statistics (PMAS)

Dimension IV: Customer Perspective

Measure 1 (Input): Number of citizen crime prevention groups

Definition: the number of citizen-volunteer patrol groups for crime prevention in each of the police stations under the SMPA

Operationalization: Use the original data in the PMAS, <u>without operationalization</u>.

Source: Police Management and Administrative Statistics (PMAS)

Measure 2 (Output): Citizens' satisfaction of police services

Definition: the citizens' satisfaction levels of various police services: (1) civil complaints; (2) calls for police services; (3) investigation for traffic accidents; and (4) investigation for criminal cases

Operationalization: It is presented in Appendix * how this indicator was measured and operationalized.

Source: Police Customer Satisfaction Survey

Measure 3 (Carry-over): Citizens' feelings of safety

Definition: the citizens' perceptions of safety in their neighborhood

Operationalization: This indicator was measured by constructing a scale that included four questions about citizens' feelings of safety: (1) How do you feel safe with crimes (including murder, robbery, and so on) in your neighborhood? (2) How do you feel safe with traffic accidents in your neighborhood? (3) How do you think people in your neighborhood abide by law and order? and (4) How do you feel safe in your neighborhood considering all the question above? Responses to each question ranged from 1 (Completely unsafe) to 10 (completely safe). The responses to these four items were summed and then averaged to create the scale.

Source: Public Safety Survey (conducted biannually, Use the survey result conducted in the second half of the years 2013-2015).

Dynamic-Network Data Envelopment Analysis Model

This study presents a dynamic-network data envelopment analysis model to measure the performance of police agencies. The DN DEA model can calculate both the efficiencies of each police agency and the dynamic changes of the efficiencies at the same time. This is the first practical application of the DN DEA model in the field of police management. The current analysis calculates efficiency scores from 2013 to 2015 not only for the 31 police agencies in the Seoul Metropolitan Police Agency, S. Korea, but also for four internal perspectives that contribute to the overall success of the agencies.

The DN model makes it possible to integrate the proposed conceptual framework, derived from the BSC approach, into the DEA method. In contrast to traditional DEA models, the DN model facilitates a multi-step production structure. Previous studies, which have employed the DEA method to assess the police agency's performance, have not considered the intermediate products in policing. In the DN model, the intermediate products can be used as link variables to combine the structures. Compared with the traditional 'Black-Box" model, one of the advantages of using the DN model is to reflect a network of interconnecting activities. Another advantage the DN model has over the BB model is to prevent inadequate correspondence between inputs and outputs. Police budget, for example, does not directly affect the production of customer satisfaction; instead, this input may correspond with the number of civilian employees in a police agency as an output. Such a wrong relationship between the police budget and the customer satisfaction may cause an unexpected bias in the estimation of police efficiency performance. The DN model, therefore, conceptually eliminates bias in the efficiency estimation by reflecting the multi-stage production structure in real situation and avoiding the inadequate matches between inputs and outputs. This study adopts three link variables in the DN model, including the number of civilian employees, the level of police officers' job satisfaction, and the average response time to calls for service.

In addition to the existence of a link variable, a carry-over variable is one of the key differences between the DN model and the BB model. A DMU functions over several terms, so there is a possibility of carry over effect that some intertemporal factors can influence its performance. For instance, an unused police budget in previous term can be carried over the next term; and then, the carry over balance can affect the police activities or services in the next term. The carry-over variable enables to consider the effect of intertemporal factors in the following terms, so the DN model can reduce estimation bias caused by the carry over effect. This study adopts four carry-over variables in the DN model, including carry-over budget, officer's perceptions of learning opportunities, number of crimes unsolved, and citizen's feelings of safety for the financial, learning and growth, internal business process, and customer perspective, respectively.

The DNSBM DEA, with input orientation and variable returns to scale, was performed to simultaneously measure both the overall performance over the entire observed period(s) and dynamic changes of perspective-period performance. Network structure considers the relationship between the four perspectives of the Balanced Scorecard (BSC) model and the dynamic structure overs the 2013-2015 period. Input-orientation is defined as minimizing the level of inputs while maintaining at least the same level of outputs.

Model Solution

It is assumed that the number of DMUs is n ($j = 1, ..., n$), with each DMU being divided into a number of k, (k = 1, . . . , K), and time periods t, (t = 1, . . . , T). Each DMU has an input and output in period t through a carry-over (link) to the next period $t + 1$.

Set m_k and r_k as the input and output for each division K, in which (*k*,*h*)*i* indicates division *k* to *h*, and L_{hk} denotes the set of *k* and *h*.

Inputs and Outputs:

$X_{ijk}^{t} \in R_{+}(i = 1, \dots, m_k; j = 1, \dots, n; K = 1 \dots, K; t = 1, \dots, T)$: indicates input *i* in period *t* for division *k* in DMU_j.

$y_{rjk}^{t} \in R_{+}\left(r = 1, \dots, r_{k'}\ j = 1, \dots, n; K = 1 \dots, K; t = 1, \dots, T\right)$: indicates output *r* in period *t* for division *k* in DMU_j.

If part of the output is not good, it is considered an input to division *k*.

Links:

$Z_{j(kh)t}^{t} \in R_{+}(j = 1; \dots; n; l = 1; ..; L_{hk}; t = 1; \dots; T)$: denotes the link between division *k* and division *h* in DMU_j in period *t*, where L_{hk} is the number of links between *k* and *h*.

$Z_{j(kh)t}^{t} \varepsilon R_{+}(j = 1; \dots; n; l = 1; \dots; L_{kh}; t = 1; \dots; T)$

Carry-overs:

$Z_{jkl}^{(t,t+1)} \in R_{+}\left(j = 1, \dots, n; l = 1, .., L_{k'}\ k = 1, \dots k, t = 1, \dots, T - 1\right)$: denotes the carry-overs from division *k* to *h* in DMU_j from period *t* to *t* + 1 where L_k is the number of carry-overs from division *k*.

Objective Function

Overall Efficiency:

$$\theta_0^* = \min \frac{\sum_{t=1}^{T} W^t \left[\sum_{k=1}^{K} W^k \left[1 - \frac{1}{m_k + linkin_k + nbad_k} \left(\sum_{i=1}^{m_k} \frac{S_{iok}^{t-}}{x_{iok}^{t}} + \sum_{(kh)_l=1}^{linkin_k} \frac{s_{o(kh)_l in}^{t}}{z_{o(kh)_l in}^{t}} + \sum_{k_l=1}^{nbad_k} \frac{s_{ok_l bad}^{(t,(t+1))}}{z_{ok_l bad}^{(t,(t+1))}}\right]\right]}{\sum_{t=1}^{T} W^t \left[\sum_{k=1}^{K} W^k \left[1 + \frac{1}{r_k + linkout_k + ngood_k} \left(\sum_{r=1}^{r_k} \frac{s_{rok}^{t+}}{y_{rok}^{t}} + \sum_{(kh)_l=1}^{linkout_k} \frac{s_{o(kh)_l out}^{t}}{z_{o(kh)_l out}^{t}} + \sum_{k_l=1}^{ngood_k} \frac{s_{oklgood}^{(t,(t+1))}}{z_{oklgood}^{(t,(t+1))}}\right]\right]}$$

Constraints

$$x_{ok}^{t} = X_{k}^{t}\lambda_{k}^{t} + s_{ko}^{t-}(\forall k, \forall t)$$
$$y_{ok}^{t} = Y_{k}^{t}\lambda_{k}^{t} - s_{ko}^{t+}(\forall k, \forall t)$$
$$e\lambda_{k}^{t} = 1(\forall k, \forall t)$$
$$\lambda_{k}^{t} \geq 0, s_{ko}^{t-1} \geq 0, s_{ko}^{t+} \geq 0, (\forall k, \forall t)$$
$$Z_{(kh)free}^{t}\lambda_{h}^{t} = Z_{(kh)free}^{t}\lambda_{k}^{t}(\forall (k,h) \text{ free }, \forall t)$$
$$Z_{(kh)free}^{t} = \left(Z_{1(kh)free}^{t}, \dots, Z_{n(kh)free}^{t}\right) \in R^{L_{(h)free} \times n}$$
$$Z_{o(kh)fix}^{t} = Z_{(kh)fix}^{t}\lambda_{h}^{t}(\forall (k,h)fix, \forall t)$$
$$Z_{o(kh)fix}^{t} = Z_{(kh)fix}^{t}\lambda_{k}^{t}(\forall (k,h)fix, \forall t)$$
$$Z_{o(kh)in}^{t} = Z_{(kh)in}^{t}\lambda_{k}^{t} + S_{o(kh)in}^{t}((kh)in = 1, \dots, linkin_{k})$$
$$Z_{o(kh)out}^{t} = Z_{(kh)out}^{t}\lambda_{k}^{t} - S_{o(kh)out}^{t}((kh)out = 1, \dots, \text{ linkout }_{k})$$
$$\sum_{j=1}^{n} z_{jk_{1}\alpha}^{(t,(t+1))}\lambda_{jk}^{t} = \sum_{j=1}^{n} z_{jk_{1}\alpha}^{(t,(t+1))}\lambda_{jk}^{t+1}(\forall k; \forall k_{l}; t = 1, \dots, T-1)$$
$$Z_{ok_{l}good}^{(t,(t+1))} = \sum_{j=1}^{n} z_{jk_{l}good}^{(t,(t+1))}\lambda_{jk}^{t} - s_{ok_{l}good}^{(t,(t+1))}k_{l} = 1, \dots, \text{ ngood }_{k}; \forall k; \forall t\Big)$$
$$Z_{ok_{l}bad}^{(t,(t+1))} = \sum_{j=1}^{n} z_{jk_{l}bad}^{(t,(t+1))}\lambda_{jk}^{t} - s_{ok_{l}bad}^{(t,(t+1))}k_{l} = 1, \dots, \text{ nbad }_{k}; \forall k; \forall t\Big)$$
$$Z_{ok_{l}\text{ free}}^{(t,(t+1))} = \sum_{j=1}^{n} z_{jk_{l}\text{ free}}^{(t,(t+1))}\lambda_{jk}^{t} - s_{ok_{l}\text{ free}}^{(t,(t+1))}k_{l} = 1, \dots, \text{ nfree }_{k}; \forall k; \forall t\Big)$$
$$Z_{ok_{l}fix}^{(t,(t+1))} = \sum_{j=1}^{n} z_{jk_{l}fix}^{(t,(t+1))}\lambda_{jk}^{t} - s_{ok_{l}fix}^{(t,(t+1))}k_{l} = 1, \dots, nfix_{k}; \forall k; \forall t\Big)$$
$$s_{ok_{l}g\text{ ood}}^{(t,(t+1))} \geq 0, s_{ok_{l}bad}^{(t,(t+1))} \geq 0, s_{ok_{l}}^{(t,(t+1))}: \text{ free } (\forall k_{l}; \forall t)$$

Period and Division Efficiencies

Period efficiency:

$$\tau_{0}^{t^{*}} = \min \frac{\sum_{k=1}^{K} W^{k}\left[1 - \frac{1}{m_{k}+linkin_{k}+nbad_{k}}\left(\sum_{i=1}^{m_{k}} \frac{S_{iok}^{t-}}{x_{iok}^{t}} + \sum_{(kh)_{l}=1}^{linkin_{k}} \frac{s_{o(kh)_{l}in}^{t}}{z_{o(kh)_{l}in}^{t}} + \sum_{k_{l}=1}^{nbad_{k}} \frac{s_{ok_{l}bad}^{(t,(t+1))}}{z_{ok_{l}bad}^{(t,(t+1))}}\right)\right]}{\sum_{k=1}^{K} W^{k}\left[1 + \frac{1}{r_{k}+linkout_{k}+ngood_{k}}\left(\sum_{r=1}^{r_{k}} \frac{s_{rok}^{t+}}{y_{rok}^{t}} + \sum_{(kh)_{l}=1}^{linkout_{k}} \frac{s_{o(kh)_{l}out}^{t}}{z_{o(kh)_{l}out}^{t}} + \sum_{k_{l}=1}^{ngood_{k}} \frac{s_{ok_{l}good}^{(t,(t+1))}}{z_{ok_{l}good}^{(t,(t+1))}}\right)\right]}$$

Division Efficiency:

$$\delta_{0t}^{*} = \min \frac{\sum_{t=1}^{T} W^{t}\left[1-\frac{1}{m_k+linkin_k+nbad_k}\left(\sum_{i=1}^{m_k}\frac{S_{iok}^{t-}}{x_{iok}^{t}}+\sum_{(kh)_l=1}^{linkin_k}\frac{s_{o(kh)_l in}^{t}}{z_{o(kh)_l in}^{t}}+\sum_{k_l=1}^{nbad_k}\frac{s_{ok_l bad}^{(t,(t+1))}}{z_{ok_l bad}^{(t,(t+1))}}\right]\right.}{\sum_{t=1}^{T} W^{t}\left[1+\frac{1}{r_k+linkout_k+ngood_k}\left(\sum_{r=1}^{r_k}\frac{s_{rok}^{t+}}{y_{rok}^{t}}+\sum_{(kh)_l=1}^{linkout_k}\frac{s_{o(kh)_l out}^{t}}{z_{o(kh)_l out}^{t}}+\sum_{k_l=1}^{ngood_k}\frac{s_{ok_l good}^{(t,(t+1))}}{z_{ok_l good}^{(t,(t+1))}}\right]\right.}$$

Division Period Efficiency:

$$p_{0k}^{t*} = \frac{1-\frac{1}{m_k+linkin_k+nbad_k}\left(\sum_{i=1}^{m_k}\frac{S_{iok}^{t-}}{x_{iok}^{t}}+\sum_{(kh)_l=1}^{linkin_k}\frac{s_{o(kh)_l in}^{t}}{z_{o(kh)_l in}^{t}}+\sum_{k_l=1}^{nbad_k}\frac{s_{ok_l bad}^{(t,(t+1))}}{z_{ok_l bad}^{(t,(t+1))}}\right.}{1+\frac{1}{r_k+linkout_k+ngood_k}\left(\sum_{r=1}^{r_k}\frac{s_{rok}^{t+}}{y_{rok}^{t}}+\sum_{(kh)_l=1}^{linkout_k}\frac{s_{o(kh)_l out}^{t}}{z_{o(kh)_l out}^{t}}+\sum_{k_l=1}^{ngood_k}\frac{s_{ok_l good}^{(t,(t+1))}}{z_{ok_l good}^{(t,(t+1))}}\right.} (\forall k; \forall t)$$

$$Z_{ol_k}^{(0,1)} = \sum_{j=1}^{n} Z_{jlk}^{(0,1)} \lambda_{jk}^{l} (\forall l_k)$$

From the above results, the overall efficiency, period efficiency, division efficiency and division period efficiency can be determined.

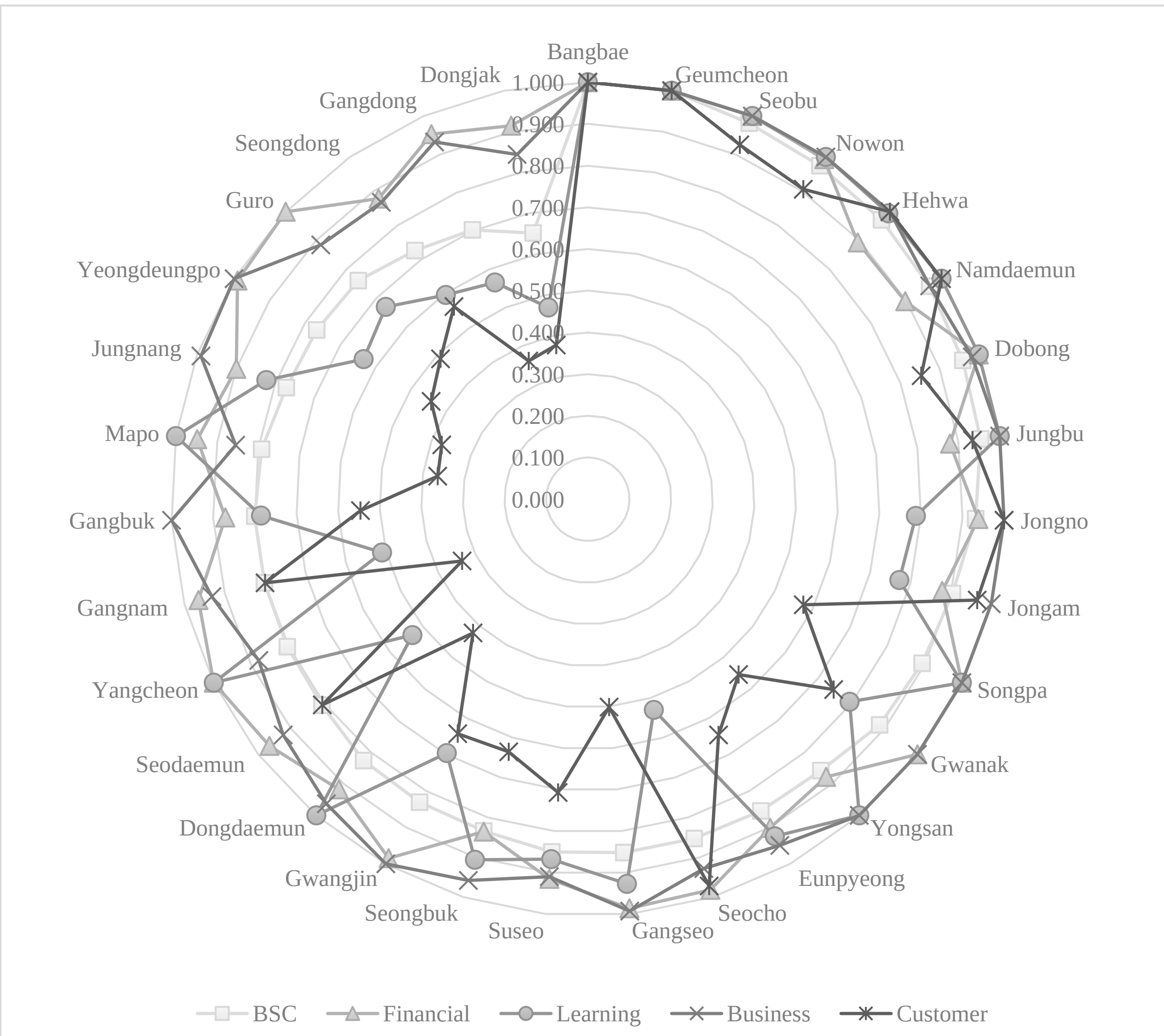
RESULTS

Table 7 presents the performance scores obtained for the 31 police stations under the Seoul Metropolitan Police Agency (SMPA). The overall performance score of the BSC model over the entire observed period was .857. It ranged from .651 (i.e., Dongjak had the lowest rank) to 1.0000 (i.e., Bangbae and Guemcheon had the highest rank); the standard deviation was small (SD=.091). Twenty-four out of the 31 police stations scored above the 80 percent performance rating. In contrast, there were seven underperforming police stations below the level of the 80 percent performance rating, including Mapo (.792), Jungnang (.771), Yeongdeungpo (.767), Guro (.761), Seongdong (.726), Gangdong (.703), and Dongjak (.651). These results reveal that the SMPA as a whole operated efficiently in terms of the overall BSC model.

Table 7 also presents the perspective-period performance scores for each of the 31 police stations compared. The overall performance scores of the four perspectives over the entire observed period are displayed in a radar chart (Figure 8), in order to facilitate the visualization of the tradeoffs between the scores obtained in each of the BSC models. The overall performance score of the financial perspective was .94. It ranged from .84 (i.e., Seongbuk) to 1.00 (i.e., Bangbae, Guemcheon, Seobu, Dobong, Songpa, Gwanak, Yangcheon, and Guro**)**; the standard deviation was small (=.05), indicating that they are clustered closely around the average. The overall performance score of the learning and growth perspective was. .83. It ranged from .47 (i.e., Dongjak) to 1.00 (i.e., Bangbae, Geumcheon, Seobu, Nowon, Hehwa, Namdaemun, Dobong, Jungbu, Songpa, Yongsan, Dongdaemun, Yangcheon, and Mapo**)**; the standard deviation was relatively large (=.18), indicating that they are far from the average. The overall performance score of the internal business process perspective was .96. It ranged from .84 (i.e., Dongjak) to 1.00 (i.e., Bangbae, Geumcheon, Seobu, Nowon, Hehwa, Jungbu, Jongno, Jongam, Songpa, Gwanak, Yongsan, Gwangjin, Gangbuk, and Yeongdeungpo); the standard deviation was relatively small (=.05). The overall performance score of the customer perspective was .69. It ranged from .34 (i.e., Yangcheon) to 1.00 (i.e., Bangbae, Geumcheon, Hehwa, Namdaemun, and Jongno); the standard deviation was relatively large (=.24).

Figure 8

Overall Performance Scores Obtained in Each of the BSC Perspectives for the Entire Period



One advantage of the DNSBM DEA model makes it possible to observe the performance scores separately for individual perspectives of the BSC model in each year of the observed periods. Figure 9 shows the dynamic changes in the overall performance scores for the four individual perspectives annually from 2013-2015. A similar dynamic change pattern was observed, showing that the scores slightly increased from 2013 to 2014 and then decreased from 2014 to 2015. For example, the annual average performance scores of the BSC model slightly fluctuated between 2013 and 2015 (.85 in 2013, .87 in 2014, .85 in 2015); those of the financial perspective were .95 in 2013, .96 in 2014, and .93 in 2015; those of the learning and growth perspective were .80 in 2013, .85 in 2014, and .84 in 2015; those of the internal business perspective were .95 in 2013, .97 in 2014, and .96 in 2015; those of the customer perspective were .69 in 2013, .73 in 2014, and .66 in 2015.

Figure 9

Dynamic Changes in the Overall Performance Scores for the Four Perspectives

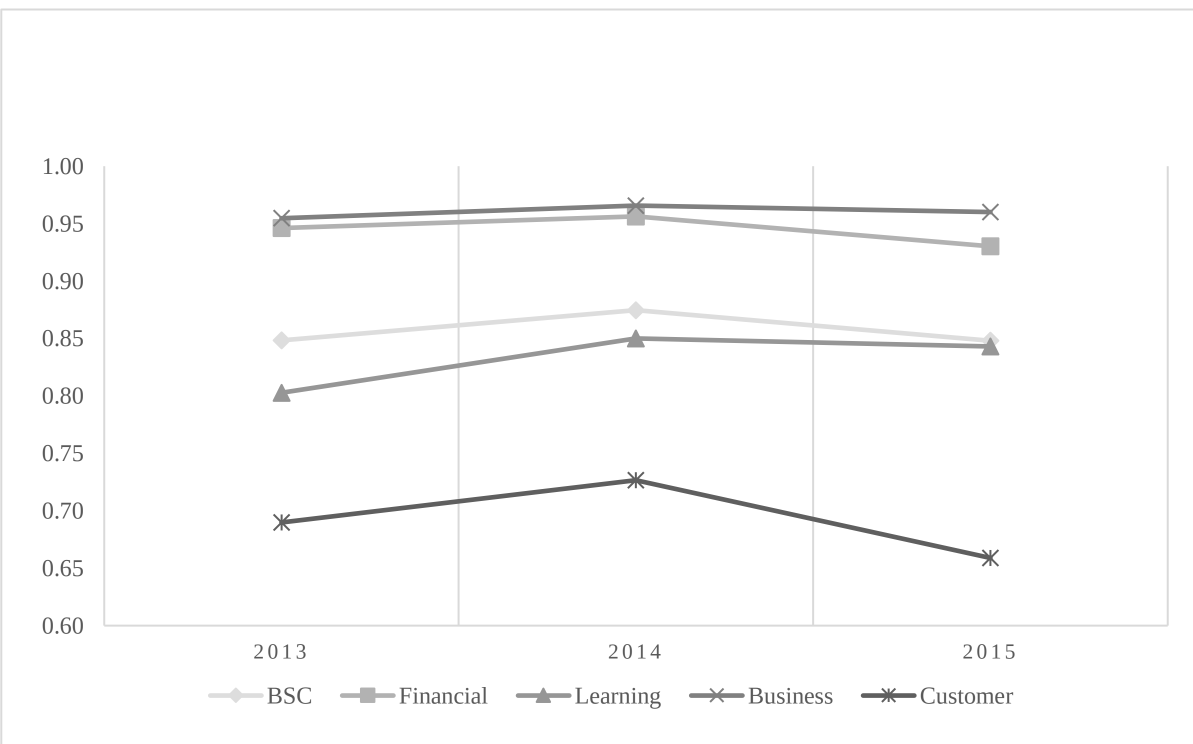


Table 7

Overall and Perspective-period Performance Scores over the Entire Observed Period

DMU	BSC					Financial				Learning & Growth				Internal Business				Customer			
	2013	2014	2015	Average	Rank	2013	2014	2015	Average	2013	2014	2015	Average	2013	2014	2015	Average	2013	2014	2015	Average
Bangbae	1.00	1.00	1.00	1.000	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Geumcheon	1.00	1.00	1.00	1.000	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Seobu	1.00	0.98	0.96	0.981	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.93	0.85	0.92
Nowon	0.92	1.00	1.00	0.975	4	0.98	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.72	1.00	1.00	0.91
Hehwa	0.97	0.98	0.96	0.972	5	0.90	0.92	0.86	0.89	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Namdaemun	0.96	0.99	0.95	0.966	6	0.84	0.96	0.89	0.90	1.00	1.00	1.00	1.00	1.00	1.00	0.90	0.97	1.00	1.00	1.00	1.00
Dobong	0.99	0.99	0.90	0.959	7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.95	1.00	1.00	0.98	1.00	0.98	0.58	0.85
Jungbu	0.98	0.98	0.90	0.953	8	0.93	0.91	0.80	0.88	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.80	0.93
Jongno	0.87	0.92	1.00	0.932	9	0.82	1.00	1.00	0.94	0.68	0.68	1.00	0.79	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Jongam	0.90	0.94	0.87	0.904	10	0.86	0.93	0.85	0.88	0.76	0.84	0.72	0.77	1.00	1.00	1.00	1.00	1.00	1.00	0.90	0.97
Songpa	0.85	1.00	0.83	0.894	11	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.40	1.00	0.32	0.58
Gwanak	0.90	0.91	0.85	0.885	12	1.00	1.00	1.00	1.00	1.00	1.00	0.38	0.79	1.00	1.00	1.00	1.00	0.61	0.63	1.00	0.75
Yongsan	0.85	0.92	0.80	0.858	13	0.89	0.93	0.82	0.88	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.52	0.75	0.39	0.55
Eunpyeong	0.90	0.88	0.78	0.856	14	0.94	0.92	0.84	0.90	1.00	1.00	0.77	0.92	0.98	0.90	0.96	0.95	0.69	0.69	0.56	0.65
Seocho	0.86	0.86	0.84	0.853	15	0.95	1.00	1.00	0.98	0.57	0.48	0.54	0.53	0.94	0.94	0.91	0.93	1.00	1.00	0.92	0.97
Gangseo	0.78	0.85	0.93	0.852	16	1.00	1.00	0.96	0.99	0.78	1.00	1.00	0.93	0.98	1.00	1.00	0.99	0.35	0.40	0.75	0.50
Suseo	0.74	0.83	0.98	0.850	17	0.90	0.91	0.93	0.92	0.67	0.94	1.00	0.87	0.86	0.90	0.97	0.91	0.55	0.57	1.00	0.71
Seongbuk	0.82	0.82	0.86	0.834	18	0.83	0.79	0.89	0.84	0.99	0.73	1.00	0.91	0.88	1.00	1.00	0.96	0.59	0.76	0.55	0.63
Gwangjin	0.77	0.83	0.90	0.831	19	1.00	0.96	1.00	0.99	0.59	0.50	1.00	0.70	1.00	1.00	1.00	1.00	0.49	0.86	0.58	0.64
Dongdaemun	0.80	0.84	0.84	0.826	20	0.92	0.94	0.89	0.92	1.00	1.00	1.00	1.00	0.89	1.00	1.00	0.96	0.39	0.41	0.48	0.42
Seodaemun	0.84	0.84	0.74	0.807	21	0.97	0.97	0.96	0.97	0.42	0.49	0.68	0.53	0.97	0.93	0.88	0.92	1.00	0.96	0.46	0.81
Yangcheon	0.80	0.82	0.79	0.804	22	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.87	0.88	0.89	0.88	0.33	0.40	0.28	0.34
Gangnam	0.73	0.87	0.80	0.802	23	0.99	1.00	0.91	0.97	0.44	0.69	0.40	0.51	0.91	0.98	0.91	0.93	0.58	0.82	1.00	0.80
Gangbuk	0.75	0.85	0.81	0.801	24	0.89	0.89	0.83	0.87	0.47	0.99	0.89	0.79	1.00	1.00	1.00	1.00	0.62	0.51	0.52	0.55
Mapo	0.79	0.79	0.79	0.792	25	0.99	0.93	0.92	0.95	1.00	1.00	1.00	1.00	0.87	0.85	0.84	0.86	0.31	0.39	0.39	0.36
Jungnang	0.80	0.77	0.74	0.771	26	0.92	0.88	0.90	0.90	0.78	0.87	0.81	0.82	1.00	1.00	0.97	0.99	0.49	0.33	0.29	0.37
Yeongdeungpo	0.73	0.83	0.74	0.767	27	0.99	1.00	0.98	0.99	0.60	0.74	0.55	0.63	1.00	1.00	1.00	1.00	0.34	0.56	0.43	0.44

Table 7. *Overall and Perspective-period Performance Scores over the Entire Observed Period (Continued)*

DMU	BSC					Financial				Learning & Growth				Internal Business				Customer			
	2013	2014	2015	Average	Rank	2013	2014	2015	Average	2013	2014	2015	Average	2013	2014	2015	Average	2013	2014	2015	Average
Guro	0.79	0.71	0.78	0.761	28	1.00	1.00	1.00	1.00	0.60	0.59	0.82	0.67	0.86	0.89	0.90	0.88	0.70	0.38	0.39	0.49
Seongdong	0.84	0.69	0.65	0.726	29	0.91	0.88	0.84	0.88	0.56	0.66	0.57	0.60	0.88	0.82	0.90	0.87	1.00	0.41	0.28	0.56
Gangdong	0.69	0.75	0.67	0.703	30	0.96	1.00	0.90	0.95	0.53	0.61	0.55	0.57	0.93	0.96	0.91	0.93	0.35	0.40	0.33	0.36
Dongjak	0.64	0.68	0.64	0.651	31	0.94	0.92	0.87	0.91	0.44	0.52	0.45	0.47	0.83	0.88	0.82	0.84	0.35	0.38	0.40	0.38
Overall									**Overall**				**Overall**				**Overall**				**Overall**
Average	0.85	0.87	0.85	0.857		0.95	0.96	0.93	0.94	0.80	0.85	0.84	0.83	0.95	0.97	0.96	0.96	0.69	0.73	0.66	0.69
St Dev	0.10	0.10	0.10	0.091		0.06	0.05	0.07	0.05	0.22	0.19	0.21	0.18	0.06	0.05	0.06	0.05	0.27	0.26	0.28	0.24
Min	0.64	0.68	0.64	0.651		0.82	0.79	0.80	0.84	0.42	0.48	0.38	0.47	0.83	0.82	0.82	0.84	0.31	0.33	0.28	0.34
Max	1.00	1.00	1.00	1.000		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

A series of Spearman's rank correlation analyses were conducted to statistically examine the relationships between the ranks (or performance scores) obtained from various DEA models. The Spearman's rank correlation coefficient, a nonparametric measure of rank correlation, assesses statistical dependence between the rankings of two models (Corder & Foreman, 2014). The correlation coefficient refers to the strength and direction of a relationship between the two models and ranges in value from -1 (negative direction) to +1 (positive direction). This test is useful when Pearson's product-moment correlation cannot be applied due to violations of normality and a linear relationship, or when ordinal variables such as rank variables are being used.

Table 8 presents the results of the rank correlations among the four performance perspectives of the BSC model. The ranks determined by the overall performance scores over the entire observed period were used for the rank correlation analysis. It is expected that there are significant associations not only between two interconnected perspectives (e.g., financial-learning, learning-business, business-customer) but also between the overall BSC model and its four perspectives because the integrated BSC-DEA model was developed in a hypothesis of cause-and-effect relationships from the financial perspective all the way through to the customer perspective. The results show that there were statistically significant rank correlations between most of the models. However, there was no rank correlation between the BSC model and the financial perspective, which was not statistically significant, $r_s = .2172$, $p = .2405$; between the financial perspective and the learning and growth perspective, which was not statistically significant, $r_s = .1519$, $p = .4147$.

Table 8

Spearman's Rank Correlation Coefficient over the Entire Observed Period

Overall	BSC	Financial	Learning	Business	Customer
BSC	1				
Financial	0.2172	1			
Learning	0.6644***	0.1519	1		
Business	0.6478***		0.3784**	1	
Customer	0.8255***			0.4723***	1

$p < 0.1$*. $p < 0.05$**. $p<0.01$***.

Table 9

Spearman's Rank Correlation Coefficient for 2013

2013	BSC	Financial	Learning	Business	Customer
BSC	1				
Financial	0.0490	1			
Learning	0.6951***	0.172	1		
Business	0.5294***		0.3544*	1	
Customer	0.7530***			0.3631**	1

$p < 0.1$*. $p < 0.05$**. $p<0.01$***.

Table 10

Spearman's Rank Correlation Coefficient for 2014

2014	BSC	Financial	Learning	Business	Customer
BSC	1				
Financial	0.3433*	1			
Learning	0.6072***	0.0866	1		
Business	0.6395***		0.3949**	1	
Customer	0.8420***			0.4977***	1

$p < 0.1$*. $p < 0.05$**. $p<0.01$***.

Table 11

Spearman's Rank Correlation Coefficient for 2015

2015	BSC	Financial	Learning	Business	Customer
BSC	1				
Financial	0.3351*	1			
Learning	0.6218***	0.1913	1		
Business	0.6194***		0.3954**	1	
Customer	0.8148***			0.3845**	1

$p < 0.1$*. $p < 0.05$**. $p<0.01$***.

The Table 9 to Table 11 present the results of the rank correlations among the four perspectives of the BSC model for each year of the observed periods. The ranks determined by the annual average performance scorers for 2013, 2014 and 2015 were used for a series of the Spearman's rank correlation analyses, respectively. The results from Table 9 to Table 11 show that there were statistically significant rank correlations between most of the models for 2013, 2014 and 2015. However, for 2013 there was no statistically significant rank correlation between the BSC model and the financial perspective and between the financial and learning perspectives. For 2014 and 2015 there was no statistically significant rank correlation between the financial and learning perspectives. An unexpected result is that the rankings of the 31 police stations for the financial perspective are not consistent with those for the learning and growth perspective because they are assumed to be interconnected following the cause-and-effect relationships hypothesized in the BSC literature. However, it should be noted that a strong correlation does not necessarily mean that there is a causal relationship (Mukaka, 2012).

More attention should be paid to the results of the customer perspective. Compared to other three perspectives of the BSC model, the customer perspective presents relatively low performance scores. Only five police stations such as Bangbae, Guemcheon, Hehwa, Namdaemun, and Jongno were considered efficient for the entire observed period. Nine out of the 31 police stations consistently report below the average performance score (.69) for the entire observed period. Also, the dynamic change pattern in the annual average performance scores between 2013 and 2015 was equally noticeable in this perspective but relatively large (see Figure 8 above). In contrast, various forms of dynamic changes in the annual performance scores for each of the sample police stations were found within the perspective. For example, Dobong police station showed a dramatic decrease in the annual performance scores (from 1.00 in 2013 to .58 in 2015); in contrast, Gangnam police station showed dramatic increase in the scores (from .58 in 2013 to 1.00 in 2015). Nevertheless, the Spearman's rank correlation analysis indicated that there were statistically significant rank correlations across the observed period 2013-2015. This means not only that performance rankings among the sample police stations were consistent across time but also that they operated consistently less efficiently, thus drawing managerial attention to the customer perspective as the main source of inefficiency observed in Table 7.

In addition to the performance score estimations, DEA also provides benchmarking information, which can be used to improve the performance of the DMUs (e.g., police stations). This benchmarking information gives DEA a distinct advantage over other performance measurement techniques including ratio analysis and the least-squares regression (Sherman & Zhu, 2006). These two kinds of information, the performance level and the benchmarking information, are inseparable. The performance level is measured by "the distance between the observed DMU and the reference DMU, which serves as a benchmarking target" (Baek & Lee, 2009, p.256). The SBM model maximizes the average improvements of relevant factors (e.g., inputs, outputs, links, and carry-overs) for the observed DMU to reach the efficient frontier. That is, the target is a strong efficient point on the frontier which is the farthest to the observed DMU (Tone, 2001).

Table A1-A3 in Appendix A present a set of benchmarks for each police station. The information in this table can be interpreted as follows. For a police station which is inefficient (i.e., less than 1.00) regarding a particular perspective-period model, the benchmarks for learning are indicated. For a police station which is efficient (i.e., equal to 1.00) regarding a particular perspective-period model, the number that is indicated represents the number of police stations for which this unit is a benchmark. It should be noted that almost all efficient DMUs (i.e., police stations) consider themselves to be their own benchmarks; however, some efficient police stations benchmark other efficient police stations as well.

For example, Jungnang police station (PS18) is inefficient in terms of the 2013 financial perspective, and its benchmarks for learning are PS9 (Gangseo), PS11 (Guro), PS13 (Gwangjin), and PS27 (Songpa). However, in terms of the 2013 internal business perspective, this police station is classified as efficient and can perform as a benchmark to other police stations, including PS3 (Dongdaemun), PS7 (Gangdong), PS9 (Gangseo), PS11 (Guro), PS19 (Mapo), and PS21 (Nowon). Interestingly, reciprocal relationships are observed between Jungnang (PS18) and Gangseo (PS9) police stations and between Jungnang (PS18) and Guro (PS11) police stations in a single year but different perspectives. Furthermore, in terms of the 2015 internal business perspective, Dongdaemun police stations (PS3) is classified as efficient and can perform as a benchmark to Jungnang police station (PS18) in the same perspective but different years. These results reveal opportunities for reciprocal learning between police stations.

DEA provides practical and objective information on how to make inefficient police stations become efficient. The benchmark DMUs on the efficiency frontier can help determine the potential improvements for inefficient units off the frontier. By calculating the Euclidean maximum distance between them, DEA can compute the potential improvements or the technical inefficiencies, which are called as slacks in DEA. The amount of slacks and benchmark targets for input-oriented VRS model are presented in Table B1-B3 and Table C1-C3 in Appendix B and C, respectively. Slacks exist only for those police stations identified as inefficient. Because the SBM-DEA directly deals with the slacks (i.e., input excesses and the output shortfalls), the benchmark target levels can be obtained by adding the slack amounts from the original dataset

(Tone, 2001). For example, Jongno police station had the lowest performance score (82% performance rating) for the 2013 financial perspective. In order for the Jongno police station to become inefficient for this particular perspective-period model given the values for original input, input slack, and input target are 1432639.30, -264086.22, and 1168553.08, respectively, the annual budget (i.e., input for the financial perspective) may be decreased by 18.43 % (e.g., (input excess / original input) * 100) while maintaining their output levels. The main idea behind identifying the slacks and the benchmark targets is to reduce the estimated input excess or output shortfalls and to provide information on how to reduce inefficiency.

Table 12 presents the mean potential improvements estimated by the input-oriented VRS (DNSBM DEA) model for inputs and outputs. Excess implies potential to lower inputs and shortage implies potential to raise outputs when inefficient police stations are benchmarked against efficient police stations compared in the sample. Input-oriented modeling of potential improvements focuses on minimizing the level of inputs for given outputs. Use of SBM enables the analysis to capture the potential non-radial changes in inputs and outputs (Avkiran, 2014).

Table 12

Mean Potential Improvements Identified by DNSBM DEA for Key Inputs and Outputs (%)

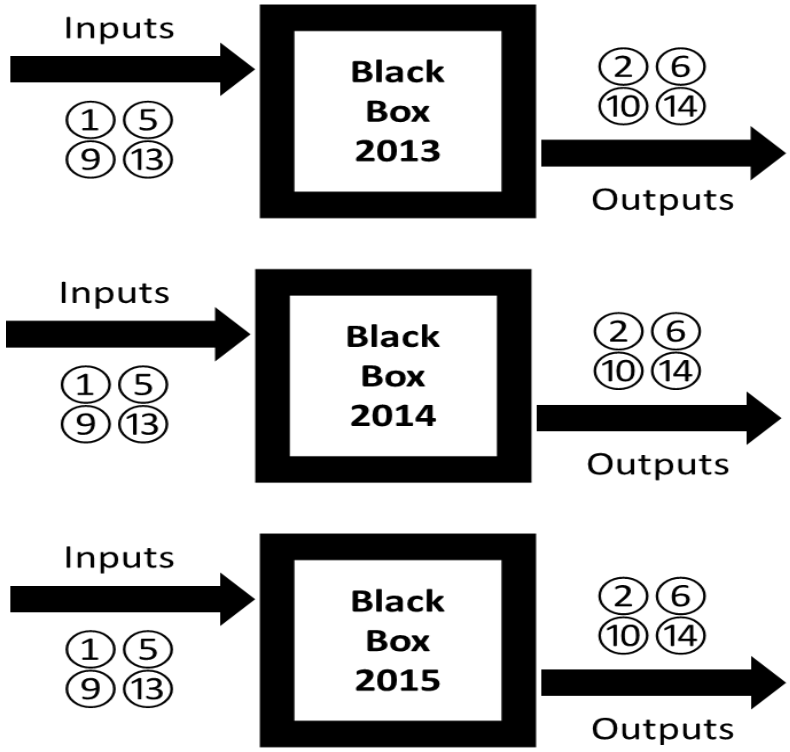
	Financial Perspective		**Learning and Growth Perspective**		**Internal Business Perspective**		**Customer Perspective**	
	Input (Excess)	Output (Shortage)	Input (Excess)	Output (Shortage)	Input (Excess)	Output (Shortage)	Input (Excess)	Output (Shortage)
2013	5.38	0	21.47	1.63	4.72	0	35.57	0.98
2014	4.31	0	21.47	1.57	3.53	0	32.46	1.15
2015	6.90	0	18.89	0.22	4.32	0	38.76	2.14
Mean	5.53	0.00	20.61	1.14	4.19	0.00	35.60	1.42

For the financial perspective, the annual budget can be decreased by 5.53% on average while maintaining their output levels (i.e., the number of police officers) unchanged. For the learning and growth perspective, the number of equipment can be decreased by 20.61 on average while officers' perception on learning environment can be increased by 1.14% on average; compared to the financial and internal business perspectives, this perspective cannot reach the efficient frontier through input reduction only. For the internal business perspective, the number of crimes occurred can be decreased by 4.19% on average while maintaining their output levels (i.e., the number of crimes cleared) unchanged. For the customer perspective, the number of community crime prevention groups can be decreased by 35.60% while citizens' satisfaction of police service can be increased by 1.42%; similar to the learning and growth perspective, this perspective also fails to reach the efficient frontier through the input reduction only. Relatively higher level of input slack values in the customer perspective as well as the learning and growth perspective indicates considerable scope for improvements.

The advantage of the DNSBM DEA model can be confirmed by comparing the results with those of a traditional BB DEA model. Figure 13 shows the structure of the BB model, for which the four BSC perspectives are aggregated into a single black box. The inputs and outputs of the BB model are exactly the same as those of the DN model. The inputs are denoted as Input ① ⑤ ⑨ ⑬ whereas the outputs are denoted as ② ⑥ ⑩ ⑭ as in Table 6 above. However, the variables for links and carry-overs do not apply in the case of the BB model because it neglects the internal structure of police stations and carry-over activities between two consecutive terms. The BBSBM DEA models, with input orientation and VRS assumptions, were run to measure the relative performance scores of the 31 police stations in Seoul. A series of analyses were conducted for each year that make up the three-year panel study.

Figure 10

Structure of the Black Box (BB model)



The results are summarized in Table 13. Estimated by the BB models over the three years, the average annual performance scores of the police stations had above 98 percent efficiency ratings. Compared with the average annual performance scores produced by the DN Model, those produced by the BB model were higher and remained stable over time (from .985 in 2013 to .980 in 2014 to .967 in 2015). There are many police stations which are evaluated as efficient in the BB model but in inefficient in the DN model (see Table 7 and Table 14). It is hard to compare both scores directly because problem schemes are different in the DN and BB models. Nevertheless, it is expected that the disparity comes from the characteristics of the applied models: the dynamic-network structure in the DN model and the aggregated one in the BB model.

Table 13

Average Performance Scores Estimated by BBSBM DEA

Black box DEA model				Dynamic-network DEA model				
	2013	2014	2015	Balanced Scorecard		2013	2014	2015
Average	0.985	0.980	0.967	Overall	Average	0.85	0.87	0.85
SD	0.044	0.050	0.055		SD	0.10	0.10	0.10
Maximum	1	1	1		Maximum	1.00	1.00	1.00
Minimum	0.802	0.807	0.841		Minimum	0.64	0.68	0.64
				Financial Perspective	Average	0.95	0.96	0.93
					SD	0.06	0.05	0.07
					Maximum	1.00	1.00	1.00
					Minimum	0.82	0.79	0.80
				Learning & Growth Perspective	Average	0.80	0.85	0.84
					SD	0.22	0.19	0.21
					Maximum	1.00	1.00	1.00
					Minimum	0.6035	0.409	0.3339
				Internal Business Perspective	Average	0.95	0.97	0.96
					SD	0.06	0.05	0.06
					Maximum	1.00	1.00	1.00
					Minimum	0.83	0.82	0.82
				Customer Perspective	Average	0.69	0.73	0.66
					SD	0.27	0.26	0.28
					Maximum	1.00	1.00	1.00
					Minimum	0.31	0.33	0.28

One of the drawbacks of the BB model is the neglect of internal and carry-over activities of police performance. It fails to identify how the processes in each part of a police station over time contribute to its success or failure. Thus, it does not explicitly identify successful or unsuccessful areas in which police stations should focus their attention for performance learning and improvement. Another drawback of the BB model is inadequate correspondence between inputs and outputs. For example, the initial inputs of the process (e.g., police budget from the financial perspective) do not directly influence the final outputs of the process (e.g., citizen satisfaction toward the police from the customer perspective). Such a mismatch between inputs and outputs may cause some unexpected biases in the estimation of police performance, resulting in wrong analysis and conclusions.

Table 14

Performance Scores for Each of the 31 Police Stations

	2013	2014	2015
Bangbae	1.000	1.000	1.000
Dobong	1.000	1.000	1.000
Dongdaemun	1.000	1.000	1.000
Dongjak	0.802	0.831	0.842
Eunpyeong	1.000	1.000	0.959
Gangbuk	1.000	1.000	1.000
Gangdong	0.929	1.000	0.891
Gangnam	1.000	1.000	1.000
Gangseo	1.000	1.000	1.000
Geumcheon	1.000	1.000	1.000
Guro	1.000	1.000	1.000
Gwanak	1.000	1.000	0.880
Gwangjin	1.000	1.000	1.000
Hehwa	1.000	1.000	0.878
Jongam	1.000	1.000	1.000
Jongno	1.000	1.000	1.000
Jungbu	1.000	1.000	1.000
Jungnang	1.000	1.000	1.000
Mapo	1.000	1.000	1.000
Namdaemun	1.000	1.000	1.000
Nowon	1.000	1.000	1.000
Seobu	1.000	1.000	1.000
Seocho	1.000	1.000	1.000
Seodaemun	1.000	0.917	0.897
Seongbuk	0.922	1.000	1.000
Seongdong	1.000	0.807	0.841
Songpa	1.000	1.000	1.000
Suseo	0.873	0.897	0.904
Yangcheon	1.000	0.932	0.890
Yeongdeungpo	1.000	1.000	1.000
Yongsan	1.000	1.000	1.000
Average	0.985	0.980	0.967
SD	0.044	0.050	0.055
Max	1.000	1.000	1.000
Min	0.802	0.807	0.841

CONCLUSIONS

Summary

This study set out to present an enhanced performance measurement system for police agencies by integrating the BSC and DEA approaches. The BSC provides the theoretical foundation for building a comprehensive performance measurement framework, while the DEA provides the analytical tool to test the theoretical framework. Specifically, this study adopted the Dynamic-Network (DN) DEA model, which provides the systematic account of the comprehensive police agency performance. A case-study approach was used to assess the feasibility of the integrated performance measurement system and to critically examine the ways in which performance information can be used for performance management in police agencies. Police stations under the Seoul Metropolitan Police Agency (SMPA) were chosen for conducting this case study.

The theoretical framework is based on the Balanced Scorecard (BSC) approach and the relationships between the financial, learning and growth, internal business process, and customer perspectives. Suggested by Niven (2008), the four performance perspectives are interconnected in a chain of cause-and-effect relationships from the financial perspective all the way through to the customer perspective. Four interconnected theoretical models were developed in the BSC framework and KPIs (i.e., Key Performance Indicators) were used as tradeoffs between the perspectives and models. For example, the number of civilian employees was used as tradeoffs between the financial perspective and the learning and growth perspective; it served as an output for the financial model and as an input for the learning and growth model. Along with the KPIs, appropriate inputs, outputs, and carry-overs were selected based on the literature review and researcher’s experience and knowledge.

In addition to the network performance system of a single period, this study also incorporates a dynamic structure in which every single-period network is linked over multiple periods. A combined dynamic-network model was established to examine the overall performance over the entire observed period, the dynamic changes of network (i.e., perspectives) performance, and the dynamic changes of period performance. The DNSBM DEA allows not only to estimate the combined theoretical model, but also to present the practical ways in which inefficient DUMs (e.g., police stations in this study) become more efficient. The DNSBM DEA reports the specific benchmarking objects and the target input and output levels for inefficient DMUs.

Discussion of Findings

The DNSBM DEA, with input-orientation and variable returns-to-scale, was run to assess the overall performance over the entire observed period as well as dynamic changes of the perspective-period performance. Data used in this empirical investigation includes the 31 police stations under the SMPA and covers the years of 2013-2015.

Table 7 shows that the overall performance score was .857 and it ranged from .651 (Dongjak had the lowest rank) to 1.00 (Bangbae and Guemcheon had the highest rank). As the BSC networks proceed through three periods, we can obtain the perspective-period performance scores as well. The overall performance score of the financial perspective was .94 and it ranged from .84 (Seongbuk had the lowest rank) to 1.00 (Eight police stations were classified as efficient). Those of the learning and growth perspective was .83 and it ranged from .47 (Dongjak had the lowest rank) to 1.00 (Thirteen police stations were classified as efficient). Those of the internal business process perspective was .96 and it ranged from .84 (Dongjak had the lowest rank) to 1.00 (Fourteen police stations were classified as efficient); Those of the customer perspective was .34 (Yangcheon had the lowest rank) to 1.00 (Five police stations were classified as efficient). The Spearman's test statistics in Table 8 revealed that there were statistically significant rank correlations between the BSC networks, except the relationship between the financial and learning and growth perspectives. Although the BSC networks as a conceptual framework assumed the cause-and-effect relationships between the four internal perspectives, statistically significant relationships identified using rank correlation coefficients should be interpreted for what they are: associations, not causal relationships (Mukaka, 2012).

Furthermore, DN-DEA enables us to examine the dynamic changes of the BSC networks and its internal perspectives. As can be seen in Figure 8 and Figure 9 (along with Table 7), all performance models show the same pattern where the annual average performance scores slightly increased from 2013 to 2014 and then decreased from 2014 to 2015. For example, the annual average performance scores obtained in the customer perspective were .69 in 2013, .73 in 2014, and .66 in 2015. A series of Spearman's rank correlation analyses for each observed period were conducted, and the results shown in Table 9 through Table 11 indicate that there were statistically significant rank correlations between the BSC networks and its perspectives across the observed period 2013-2015.

In addition, the DNSBM-DEA can determine benchmarks and amount of slacks (i.e., input or output inefficiency) to make inefficient police stations to become efficient. Table A1-A3 in Appendix A presents a best-practice peer group, and in turn Table B1-B3 and Table C1-C3 in Appendix B and C present the target levels for input, output, link and carry-over variables. Table 9 present the mean potential improvements identified by DNSBM-DEA for key inputs and outputs. These objective and practical information are very useful for organizational learning and performance improvement. In particular, interesting possibility for reciprocal learning between peer police stations compared were observed.

Lastly, the advantages of DNSBM-DEA over the BBSBM-DEA were confirmed by comparing the results with those of BBSBM-DEA. When network and dynamic dimensions are incorporated in a DN-DEA model, a more comprehensive information can be obtained and thus enables accurate estimate of organizational performance as well as identify potential improvements in more detail.

Contributions of the Study

This is the first study to integrate modelling approaches intended to establish an enhanced performance measurement system for police agencies. The proposed model combines two of the most popular methods used for organizational performance evaluation: the BSC and the DEA. As discussed in Chapter 2, the integration of these two methods creates many synergy effects because they are complementary to each other. It proves useful in providing structured information regarding various types of performance and ways to improve it. This approach for this study may be applied to other law enforcement agencies elsewhere in the world.

This study establishes a conceptual framework which aims to assess police organizations from multiple perspectives by adopting the BSC. Previous DEA-based police performance studies have narrowly focused on traditional police work (e.g., crime control) and used objective crime and policing data (e.g., UCR and NIBRS). To overcome these limitations, the present study employs both the objective and subjective (e.g., police officers' or citizens' survey) data distributed over the four perspectives of police performance: (1) financial; (2) learning and growth; (3) internal process; and (4) customer perspective. A balanced set of performance measurement system provides a more comprehensive assessment and understanding of modern police organizations.

This present study adds to a growing body of research that uses the DEA methodology for measuring the performance of police agencies. Since an original DEA model (i.e., an output-oriented CCR DEA model) was firstly proposed by Thanassoulis (1995) to analyze 41 police forces in England and Wales, some similar traditional DEA models have been also applied to previous studies assessing police organizational performance; however, one of the drawbacks of these traditional models is the omission of the network and/or dynamic structure of the DMUs (i.e., police agency). To overcome the limitations of traditional DEA, the current study, for the first time, adopts the DN DEA, which combines the network SBM (NSBM) and the dynamic SBM (DSBM) models. The DNSBM DEA model provides a powerful analytical tool for conducting a comprehensive analysis of police performance.

This study provides the first comprehensive assessment of police agencies in Seoul, South Korea by using the BSC and/or DEA approach. This work contributes to existing knowledge of police agency performance by conducting a case study of Seoul where there are 31 police stations that provide a wide range of policing services at the local level. This case study can facilitate the comparative study of police performance between Korea and other countries (or between Seoul and other cities in Korea) and also have a number of theoretical, methodological and practical implications for future practice and research. One possible implication of this is that other police departments can benchmark an effective performance measurement system or successful strategies to enhance police agency performance.

When applying the integrated BSC-DEA performance measurement system to other medium or small police departments, it should be noted that they are a highly homogeneous group of DMUs, providing similar police services and so producing various comparable performance measures based on the conceptual framework. The balanced panel data set is required for the DNSBM DEA; that is, missing data for various DEA models in a single performance measurement system is not allowed. Also, an appropriate sample size should be used to ensure the discriminatory power of the DEA analysis. Avkiran and McCrystal (2014) tested the sensitivity of efficiency estimates to sample size for the DNDEA and revealed that results become more discriminating as sample sizes increases. A commonly accepted DEA rule-of-thumb specifies a minimum sample size of [3 * (#inputs + #outputs)] (Avkiran & McCrystal, 2014; Bowlin, 1998). Accounting for three periods in this study, the minimum suggested sample size is at least 18 police stations, i.e. based on each perspective of the BSC model as depicted in Figure 9, [3*(1+1)] * 3 = 18.

Limitations of the Study

This study was limited by the absence of communication and cooperation with internal experts who have experiences and knowledge about an organization's performance measurement system when developing a BSC for the SMPA. Several workshops should be undertaken with the internal experts (e.g., the head and other managers of the organization) in order not only to clearly identify the mission, vision, and strategy of the organization but also to define the strategic objectives, critical success factors, and key performance indicators in unambiguous, measurable, and accurate term. A key benefit of developing a BSC with the internal experts through a disciplined strategic planning process is that it gives organizations more accurate structured information. Another key benefit is that participation of the internal experts in the process of developing a BSC strategic planning can make values more than just an academic exercise conducted at a single point in time by applying research to real-world problems.

Implications for Future Research and Practice

Conceptual Implication

This study adopts the BSC to establish an improved performance measurement system for modern police organizations; it serves as the conceptual framework for this study. Because the original BSC, developed by Kaplan and Norton (1992), was designed to address performance measurement challenges of private-sector firms, some modifications have been made to apply it to the public sector including police organizations. For example, the current study involves the same four perspectives but constructs a new interconnected BSC model as shown in Figure 3 in Chapter 2. Otherwise, *In Recognizing Value in Policing: The Challenge of Measuring Police Performance* (2002), Moore and Braga proposed the seven dimensions of the BSC for law enforcement. This study has offered a framework for the exploration of publicly valuable dimensions (or perspectives) of policing. Future research might explore the identification of different dimensions of police performance that should reflect a wide range of their responsibilities and meet citizens' demands for accountability.

Methodological Implication

In order to clarify the advantages of the DN DEA model over the BB model, this current study compares the results of both models in Chapter 5. Given that the DN DEA is a composite of the network SBM (NSBM) and the dynamic SBM (DSBM) models, it would be interesting to compare the results of these three models. The DN DEA model was developed to overcome the limitations of existing DEA models by proposing a dynamic DEA model involving network structure in each period within the framework of a slack-based measure approach (Tone & Tsutsui, 2014). By applying these slack-based DEA models to the empirical investigation of police performance measurement, we can observe the comparative advantages of the models.

These findings will enhance our understanding of the DN DEA model and help other researchers use a DEA method to measure the performance of DMUs.

Also, more research is required to investigate the Malmquist index under the DN SBM model. The Malmquist index was developed by Fare, Norris, and Zhang (1994) to compute performance change over time. It may be valuable to combine these two methods in order not only to obtain enhanced performance analysis and assessment but also to better capture the complexities of performance changes over time (Amado et al., 2012; Ozcan & Khushalani, 2017).

Practical Implication

Although developing a comprehensive and effective performance measurement framework is necessary for systematically evaluating the impact of policing efforts, it should be treated as part of a broad police management. Because performance evaluation is not just an academic exercise carried out at a single point in time, it should be an ongoing strategic approach that can reengineer the performance measurement system to drive improvement in policing. Therefore, an important practical implication is to incorporate performance measurement into the broader management system. This effort will ensure that police agencies are able to develop valid and reliable performance measures and use the information in everyday policing.

REFERENCES

Alpert, G. P., & Moore, M. H. (1993). Measuring police performance in the new paradigm of policing. *Performance measures for the criminal justice system*, *109*, 111.

Agarwal, S., Yadav, S. P., & Singh, S. P. (2014). Sensitivity analysis in data envelopment analysis. *International Journal of Operational Research*, *19*(2), 174-185.

Amado, C. A., Santos, S. P., & Marques, P. M. (2012). Integrating the Data Envelopment Analysis and the Balanced Scorecard approaches for enhanced performance assessment. *Omega*, *40*(3), 390-403.

Andersen, P., & Petersen, N. C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management science*, *39*(10), 1261-1264.

Archbold, C. A. (2012). *Policing: A text/reader*. Sage.

Avkiran, N. K. (2014). A tutorial on using dynamic network DEA to benchmark organizational performance. *Available at SSRN 2567557*.

Avkiran, N.K., & McCrystal, A. (2014). Dynamic network range-adjusted measure vs. dynamic network slacks-based measure. Journal of the Operations Research Society of Japan, 57(1), 1-14.

Aristovnik, A., Seljak, J., & Mencinger, J. (2013). Relative efficiency of police directorates in Slovenia: A non-parametric analysis. *Expert Systems with Applications*, *40*(2), 820-827.

Aristovnik, A., Seljak, J., & Mencinger, J. (2014). Performance measurement of police forces at the local level: A non-parametric mathematical programming approach. *Expert Systems with Applications*, *41*(4), 1647-1653.

Baek, C., & Lee, J. D. (2009). The relevance of DEA benchmarking information and the least-distance measure. *Mathematical and Computer Modelling*, *49*(1-2), 265-275.

Banker, R. D., & Morey, R. C. (1986). Efficiency analysis for exogenously fixed inputs and outputs. *Operations research*, *34*(4), 513-521.

Banker, R. D., Chang, H., & Pizzini, M. J. (2004). The balanced scorecard: Judgmental effects of performance measures linked to strategy. *Accounting Review, 79*(1), 1–23.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, *30*(9), 1078-1092.

Banker, R. D., Potter, G., & Srinivasan, D. (2000). An empirical investigation of an incentive plan that includes nonfinancial performance measures. *Accounting Review, 75*(1), 65–92.

Banker, R., Natarajan, R., & Zhang, D. (2019). Two-stage estimation of the impact of contextual variables in stochastic frontier production function models using data envelopment analysis: second stage OLS versus bootstrap approaches. *European Journal of Operational Research*, *278*(2), 368-384.

Barros, C. P. (2006). Productivity growth in the Lisbon police force. *Public Organization Review*, *6*(1), 21-35.

Barros, C. P. (2007). The city and the police force: analysing relative efficiency in city police precincts with data envelopment analysis. *International Journal of Police Science & Management*, *9*(2), 164-182.

Bastedo, M. N. (2004). Open systems theory.

Bayley, D. H. (1994). *Police for the Future*. Studies in Crime and Public Policy.

Bayley, D. H. (2005). *Changing the guard: Developing democratic police abroad*. Oxford University Press.

Bible, L., Kerr, S., & Zanini, M. (2006). The balanced scorecard: here and back: from its beginnings as a performance measurement tool. *Management Accounting Quarterly*, *7*(4), 18-23.

Bowlin, W.F. (1998). Measuring Performance: An Introduction to Data Envelopment Analysis (DEA). The Journal of Cost Analysis, 15(2), 3-27.

Bratton, W. J. (1999, July). Great expectations: how higher expectations for police departments can lead to a decrease in crime. In *Measuring What Matters: Proceedings from the Policing Research Institute Meetings, Washington, DC: National Institute of Justice* (pp. 11-26).

Cao, Y., & Yang, Z. (2009, July). Using DEA window analysis and the Malmquist index to evaluate the operations of Canadian Schedule I banks. In *2009 IEEE/INFORMS International Conference on Service Operations, Logistics and Informatics* (pp. 353-358). IEEE.

Carrington, R., Puthucheary, N., Rose, D., & Yaisawarng, S. (1997). Performance measurement in government service provision: the case of police services in New South Wales. *Journal of Productivity Analysis*, *8*(4), 415-430.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, *2*(6), 429-444.

Chiang, C. Y., & Lin, B. (2009). An integration of balanced scorecards and data envelopment analysis for firm's benchmarking management. *Total Quality Management*, *20*(11), 1153-1172.

Cole, M., & Parston, G. (2006). *Unlocking public value: A new model for achieving high performance in public service organizations*. John Wiley & Sons.

Cook, W. D., & Seiford, L. M. (2009). Data envelopment analysis (DEA)–Thirty years on. *European journal of operational research*, *192*(1), 1-17.

Cook, W.D., Tone, K., & Zhu J. (2014). Data Envelopment Analysis: Prior to Choosing a Model. OMEGA, The International Journal of Management Science, 44, 1-4.

Cook, W. D., & Zhu, J. (2006). *Modeling performance measurement: applications and implementation issues in DEA* (Vol. 566). Springer Science & Business Media.

Cook, W. D., Liang, L., & Zhu, J. (2010). Measuring performance of two-stage network structures by DEA: a review and future perspective. *Omega*, *38*(6), 423-430.

Cooper, W. W., Seiford, L. M., & Zhu, J. (2011). Data envelopment analysis: History, models, and interpretations. In *Handbook on data envelopment analysis* (pp. 1-39). Springer, Boston, MA.

Cordner, G., & Melekian, B. K. (2010). Reducing fear of crime strategies for police.

Corder, G. W., & Foreman, D. I. (2014). *Nonparametric statistics: A step-by-step approach*. John Wiley & Sons.

Dahle, T. O., & Archbold, C. A. (2015). "Just do what you can […] make it work!" Exploring the impact of rapid population growth on police organizations in western North Dakota. *Policing: an international journal of police strategies & management*.

Davis, R. C. (2012). *Selected international best practices in police performance measurement*. Rand Corporation.

Davis, R. C., Ortiz, C. W., Euler, S., & Kuykendall, L. (2015). Revisiting "Measuring What Matters: " Developing a Suite of Standardized Performance Measures for Policing. *Police Quarterly*, *18*(4), 469-495.

Donaldson, L., & Lex, D. (1995). *American anti-management theories of organization: A critique of paradigm proliferation* (Vol. 25). Cambridge University Press.

Drake, L. M., & Simper, R. (2003). An evaluation in the choice of inputs and outputs in the efficiency measurement of police forces. *The Journal of Socio-Economics*, *32*(6), 701-710.

Drake, L. M., & Simper, R. (2005). The measurement of police force efficiency: An assessment of UK Home Office policy. *Contemporary Economic Policy*, *23*(4), 465-482.

Eilat, H., Golany, B., & Shtub, A. (2008). R&D project evaluation: An integrated DEA and balanced scorecard approach. *Omega*, *36*(5), 895-912.

Fare, R., Grosskopf, S., Norris, M., & Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *American economic review*, *84*(1), 66-83.

Fletcher, H. D., & Smith, D. B. (2004). MANAGING FOR VALUE: DEVELOPING A PERFORMANCE MEASUREMENT SYSTEM INTEGRATING ECONOMIC VALUE ADDED AND THE BALANCED SCORECARD IN STRATEGIC PLANNING. *Journal of Business Strategies*, *21*(1).

García-Sánchez, I. M. (2007). Evaluating the effectiveness of the Spanish police force through data envelopment analysis. *European Journal of Law and Economics*, *23*(1), 43-57.

García-Sánchez, I. M., Rodríguez-Domínguez, L., & Parra-Domínguez, J. (2013). Yearly evolution of police efficiency in Spain and explanatory factors. *Central European Journal of Operations Research*, 1-32.

Gascón, G. (2010). *Making policing more affordable: Managing costs and measuring value in policing*. Harvard Kennedy School Program in Criminal Justice Policy and Management.

Geun Choi, C. (2011). The effects of social disorganization and police force size on police performance. *International Review of Public Administration*, *16*(3), 25-44.

Goldstein, H. (1990). Excellence in problem-oriented policing. *New York: McGraw-Hill. Im, K., Pesaran, M., & Shin, Y.(2003). Testing for unit roots in heterogeneous panels. J. Econometrics*, *115*, 53-74.

Goltz, J. (2006). Police organizational performance in the state of Florida: Confirmatory analysis of the relationship of the environment and design structure to performance.

Gomes, J., & Romão, M. (2017, April). Balanced Scorecard: Today's Challenges. In *World Conference on Information Systems and Technologies* (pp. 648-657). Springer, Cham.

Gorby, D. M. (2013). The failure of traditional measures of police performance and the rise of broader measures of performance. *Policing: A Journal of Policy and Practice*, *7*(4), 392-400.

Gorman, M. F., & Ruggiero, J. (2008). Evaluating US state police performance using data envelopment analysis. *International Journal of Production Economics*, *113*(2), 1031-1037.

Hadad, Y., Keren, B., & Hanani, M. Z. (2015). Combining data envelopment analysis and Malmquist Index for evaluating police station efficiency and effectiveness. *Police Practice and Research*, *16*(1), 5-21.

Hassell, K. D. (2007). Variation in Police Patrol Practices-The Precint as a Sub-Organizational Level of Analysis. *Policing: Int'l J. Police Strat. & Mgmt.*, *30*, 257.

Hoff, A. (2007). Second stage DEA: Comparison of approaches for modelling the DEA score. *European Journal of Operational Research*, *181*(1), 425-435.

Hu, J. L., Tung, Y. Y., Shieh, H. S., & Lo, K. (2011). A metafrontier efficiency analysis of regional police agencies in Taiwan. *Journal of Management research*, *11*(3), 168.

Judge, T. A., Parker, S. K., Colbert, A. E., Heller, D., & Ilies, R. (2002). Job satisfaction: A cross-cultural review.

Kao, C. (2014). Network data envelopment analysis: A review. *European Journal of Operational Research*, *239*(1), 1-16.

Kaplan, R. S., & Norton, D. P. (1992). *Measures that drive performance*. Harvard Business Review. Robert, K. S., & David, N. P. (1996). The balanced scorecard: translating strategy into action. *Harvard Business School Press, Boston, Massachusetts*, 9.

Kaplan, R. S., & Norton, D. P. (1996). Using the balanced scorecard as a strategic management system.

Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations* (Vol. 2, p. 528). New York: Wiley.

Kawaguchi, H., Tone, K., & Tsutsui, M. (2014). Estimation of the efficiency of Japanese hospitals using a dynamic and network data envelopment analysis model. *Health care management science*, *17*(2), 101-112.

Kiedrowski, J. S., Petrunik, M., Macdonald, T., & Melchers, R. (2014). *Canadian police board views on the use of police performance metrics*. Law Enforcement and Policing Branch, Public Safety Canada.

Kim, H. M., & Han, S. S. (2012). Seoul. *Cities*, *29*(2), 142-154.

Klinger, D. A. (1997). Negotiating order in patrol work: An ecological theory of police response to deviance. *Criminology*, *35*(2), 277-306.

Klopp, G. A. (1986). THE ANALYSIS OF THE EFFICIENCY OF PRODUCTIVE SYSTEMS WITH MULTIPLE INPUTS AND OUTPUTS.

Korean National Police Agency (2016). *In Police Statistics Yearbook.* Retrieved August 13, 2018 https://www.police.go.kr/files/infodata/200141/2016/201601.pdf

Larson, R. C. (1978). *Police accountability: Performance measures and unionism* (Vol. 2). Lexington Books.

Lee, J. Y. (2012). Combining balanced scorecard and data envelopment analysis in kitchen employees performance measurement: An exploratory study.

Lilley, D., & Hinduja, S. (2007). Police officer performance appraisal and overall satisfaction. *Journal of Criminal Justice*, *35*(2), 137-150.

Maguire, E. (2003). Measuring the performance of law enforcement agencies. *Calea Update*, *83*, 84.

Maguire, E. R. (2004). Measuring the performance of law enforcement agencies: Part 2. CALEA Update Magazine, 84.

Maguire, E. R., & Uchida, C. D. (2000). Measurement and explanation in the comparative study of American police organizations. *Criminal justice*, *4*, 491-557.

Malmquist, S. (1953). Index numbers and indifference surfaces. *Trabajos de estadística*, *4*(2), 209-242.

Marteache, N., & Maxfield, M. G. (2011). *Performance measurement and accountability systems*. Oxford University Press.

Maslov, A. (2016). *Measuring the Performance of the Police*. Miscellaneous Agency.

Mastrofski, D. (1999). Policing for people. In *Ideas in American Policing Series*.

Mastrofski, S. D. (2006). *Police organization and management issues for the next decade*. Washington DC: National Institute of Justice.

Mastrofski, S. D., & Willis, J. J. (2010). Police organization continuity and change: Into the twenty-first century. *Crime and Justice*, *39*(1), 55-144.

McDonald, J. (2009). Using least squares and tobit in second stage DEA efficiency analyses. *European Journal of Operational Research*, *197*(2), 792-798.

Moore, M. H. (1995). *Creating public value: Strategic management in government*. Harvard university press.

Moore, M. H., & Braga, A. (2003a). The "bottom line" of policing: What citizens should value (and measure!) in police performance. In *Police Executive Research Forum*.

Moore, M. H., & Braga, A. A. (2003). Measuring and improving police performance: The lessons of Compstat and its progeny. *Policing: An International Journal of Police Strategies & Management*.

Moore, M. H., & Braga, A. A. (2004). Police performance measurement: A normative framework. *Criminal justice ethics*, *23*(1), 3-19.

Moore, M. H., Spelman, W., & Young, R. (1992). *Innovations in Policing: A Test of Three Different Methodologies for Indentifying Important Innovations in a Substantive Field*. Taubman Center for State and Local Government, John F. Kennedy School of Government, Harvard University.

Mukaka, M. M. (2012). A guide to appropriate use of correlation coefficient in medical research. *Malawi medical journal*, *24*(3), 69-71.

Muñiz, M. A. (2002). Separating managerial inefficiency and external conditions in data envelopment analysis. *European Journal of Operational Research*, *143*(3), 625-643.

National Quality Forum. (2013). Composite performance measure evaluation guidance.

National Research Council. (2004). *Fairness and effectiveness in policing: The evidence*. National Academies Press.

Neves, J., & Lourenco, S. (2008). *Performance evaluation of worldwide hotel industry using*

Niven, P. R. (2008). *Balanced scorecard: Step-by-step for government and nonprofit agencies*. John Wiley & Sons.

Northcott, D., & Taulapapa, T. M. A. (2012). Using the balanced scorecard to manage performance in public sector organizations. *International Journal of Public Sector Management*.

Ozcan, Y. A., & Khushalani, J. (2017). Assessing efficiency of public health and medical care provision in OECD countries after a decade of reform. *Central European Journal of Operations Research*, *25*(2), 325-343.

Park, H. H., Lee, M. W., & Jo. Y. S., (2018). Fundamental qualities every police officer should have for the community-oriented policing: How to apply key traits, characteristics, and skill sets in the recruitment process. *The Journal of Police Policies, 32*(1), 43-74.

Police Executive Research Forum. (2013). COMPSTAT: Its origins, evolution, and future in law enforcement agencies.

Probst, A. (2009). Performance measurement, benchmarking & outcome-based budgeting for Wisconsin local government. *Local Government Center*.

Ramalho, E. A., Ramalho, J. J., & Henriques, P. D. (2010). Fractional regression models for second stage DEA efficiency analyses. *Journal of Productivity Analysis*, *34*(3), 239-255.

Ray, S. C. (1991). Resource-use efficiency in public schools: A study of Connecticut data. *Management Science*, *37*(12), 1620-1628.

Rickards, R. C. (2007). BSC and benchmark development for an e - commerce SME. *Benchmarking: An International Journal*.

Robert, K. S., & David, N. P. (1996). The balanced scorecard: translating strategy into action. *Harvard Business School Press, Boston, Massachusetts*, 9.

Roberts, D. J. (2006). *Law enforcement tech guide for creating performance measures that work: a guide for executives and managers*. Washington, DC: US Department of Justice, Office of Community Oriented Policing Services.

Rose, R. C., Kumar, N., & Pak, O. G. (2009). The effect of organizational learning on organizational commitment, job satisfaction and work performance. *Journal of Applied Business Research*, *25*(6), 55.

Ruggiero, J. (1996). On the measurement of technical efficiency in the public sector. *European Journal of Operational Research*, *90*(3), 553-565.

Ruggiero, J. (1998). Non-discretionary inputs in data envelopment analysis. *European Journal of Operational Research*, *111*(3), 461-469.

Shafiee, M., Lotfi, F. H., & Saleh, H. (2014). Supply chain performance evaluation with data envelopment analysis and balanced scorecard approach. *Applied Mathematical Modelling*, *38*(21-22), 5092-5112.

Sherman, H. D., & Zhu, J. (2006). *Service productivity management: Improving service performance using data envelopment analysis (DEA)*. Springer science & business media.

Sherman, H. D., & Zhu, J. (2013). Analyzing performance in service organizations. *MIT Sloan Management Review*, *54*(4), 37.

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*, *136*(1), 31-64.

Skogan, W. G. (1976). Efficiency and effectiveness in big-city police departments. *Public administration review*, 278-286.

Sparrow, M. K. (2015). Measuring performance in a modern police organization. *Psychosociological Issues in Human Resource Management*, *3*(2), 17-52.

Spector, P. E. (1997). *Job satisfaction: Application, assessment, causes, and consequences* (Vol. 3). Sage publications.

Sun, S. (2002). Measuring the relative efficiency of police precincts using data envelopment analysis. *Socio-Economic Planning Sciences*, *36*(1), 51-71.

Swanson, C. R., Territo, L., & Taylor, R. W. (1998). Stress and police personnel. *Police administration: Structures, processes, and behavior*, 297-326.

Thanassoulis, E. (1995). Assessing police forces in England and Wales using data envelopment analysis. *European Journal of Operational Research*, *87*(3), 641-657.

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European journal of operational research*, *130*(3), 498-509.

Tone, K., & Tsutsui, M. (2009). Network DEA: A slacks-based measure approach. *European journal of operational research*, *197*(1), 243-252.

Tone, K., & Tsutsui, M. (2010). Dynamic DEA: A slacks-based measure approach. *Omega*, *38*(3-4), 145-156.

Tone, K., & Tsutsui, M. (2014). Dynamic DEA with network structure: A slacks-based measure approach. *Omega*, *42*(1), 124-131.

Taniguchi, T. A., & Salvatore, C. (2018). Policing a negotiated world: A partial test of Klinger's ecological theory of policing. *Journal of Quantitative Criminology*, *34*(2), 345-366.

Tsutsui, M., & Goto, M. (2009). A multi-division efficiency evaluation of US electric power companies using a weighted slacks-based measure. *Socio-Economic Planning Sciences*, *43*(3), 201-208.

Verma, A., & Gavirneni, S. (2006). Measuring police efficiency in India: an application of data envelopment analysis. *Policing: An international journal of police strategies & management*, *29*(1), 125-145.

Von Bertalanffy, L. (1972). The history and status of general systems theory. *Academy of management journal*, *15*(4), 407-426.

Walker, S., & Katz, C. M. (2012). *Police in America*. McGraw-Hill.

Walters, J. (1998). *Measuring up: Governing's guide to performance measurement for geniuses (and other public managers)*. Governing Books.

Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in quantitative methods for psychology*, *8*(1), 52-69.

Wu, T. H., Chen, M. S., & Yeh, J. Y. (2010). Measuring the performance of police forces in Taiwan using data envelopment analysis. *Evaluation and Program Planning*, *33*(3), 246-254.

Xu, H. (2008). *Municipal police expenditure and organizational performance in the United States*. ProQuest.

Zhao, J. S., He, N., & Lovrich, N. P. (2003). Community policing: Did it change the basic functions of policing in the 1990s? A national follow-up study. *Justice Quarterly*, *20*(4), 697-724.

Zhao, J. S., Scheider, M. C., & Thurman, Q. (2003). A national evaluation of the effect of COPS grants on police productivity (arrests) 1995-1999. *Police Quarterly*, *6*(4), 387-409.

Zhao, J. S., Zhang, Y., & Thurman, Q. (2011). Can Additional Resources Lead to Higher Levels of Productivity (Arrests) in Police Agencies?. *Criminal Justice Review*, *36*(2), 165-182.

APPENDIX A. BENCHMARKS DETERMINED BY DNSBM DEA

(2013-2015)

Table A1

Benchmarks Determined by DN-DEA for 2013

2013	Financial	Learning	Business	Customer
(PS1) Bangbae	1	8	4	3
(PS2) Dobong	PS 2	4	1 PS 6; PS 10; PS 20; PS 22	5
(PS3) Dongdaemun	PS 9; PS 11; PS 27	PS 3	PS 6; PS 14; PS 18; PS 31	PS 20; PS 22
(PS4) Dongjak	PS 9; PS 11; PS 13; PS 22	PS 1; PS 5; PS 22	PS 6; PS 14; PS 16; PS 20	PS 2; PS 20
(PS5) Eunpyeong	4 PS 9; PS 13; PS 22	11	PS 6; PS 14; PS 15; PS 31	1
(PS6) Gangbuk	PS 9; PS 11; PS 13; PS 22	PS 1; PS 14	13	PS 2; PS 20; PS 22; PS 24
(PS7) Gangdong	PS 9; PS 11; PS 22	PS 5; PS 22	PS 1; PS 8; PS 10; PS 18; PS 27	PS 20
(PS8) Gangnam	PS 5; PS 9; PS 10; PS 21	1 PS 1; PS 22	3 PS 8; PS 10; PS 13; PS 30	3 PS 1; PS 20
(PS9) Gangseo	18	PS 2; PS 5; PS 29	PS 10; PS 13; PS 18	PS 20; PS 23
(PS10) Geumcheon	8	2	9	1
(PS11) Guro	11	PS 14; PS 15; PS 31	PS 10; PS 12; PS 13; PS 18; PS 22	PS 2
(PS12) Gwanak	PS 12	PS 12	1	PS 10; PS 15; PS 17; PS 20
(PS13) Gwangjin	10 PS 9; PS 11; PS 13; PS 22	PS 1; PS 5; PS 10; PS 22	4	PS 20; PS 22; PS 23
(PS14) Hehwa	PS 10; PS 15; PS 22	3	6	1
(PS15) Jongam	PS 1; PS 10; PS 22	1 PS 1; PS 5; PS 10; PS 22	3	2
(PS16) Jongno	PS 5; PS 9; PS 10; PS 22	PS 1; PS 22	2 PS 1; PS 6; PS 14; PS 16; PS 20	2
(PS17) Jungbu	PS 10; PS 13; PS 22	PS 17	1	1
(PS18) Jungnang	PS 9; PS 11; PS 13; PS 27	PS 5; PS 14; PS 31	6	PS 2; PS 14; PS 20; PS 22

Table A1. *Benchmarks Determined by DN-DEA for 2013 (Continued)*

2013	Financial	Learning	Business	Customer
(PS19) Mapo	PS 9; PS 11; PS 13; PS 27	PS 19	PS 6; PS 10; PS 13; PS 18; PS 30	PS 20
(PS20) Namdaemun	PS 5; PS 9; PS 10; PS 22	PS 20	3	18
(PS21) Nowon	3 PS 10; PS 11; PS 13; PS 27	PS 21	PS 6; PS 10; PS 17; PS 18; PS 30	PS 2; PS 16; PS 20; PS 24
(PS22) Seobu	15	10	4	4
(PS23) Seocho	PS 5; PS 9; PS 10; PS 21	PS 1; PS 5; PS 22	PS 1; PS 6; PS 8; PS 28; PS 31	5
(PS24) Seodaemun	PS 9; PS 11; PS 13; PS 22	PS 1; PS 22	PS 6; PS 8; PS 24; PS 28	2
(PS25) Seongbuk	PS 9; PS 22; PS 29	PS 2; PS 5; PS 29	PS 1; PS 6; PS 14; PS 15	PS 1; PS 20
(PS26) Seongdong	PS 9; PS 11; PS 22	PS 2; PS 5; PS 22	PS 6; PS 10; PS 14; PS 16; PS 22	PS 26
(PS27) Songpa	PS 27	1	1	PS 5; PS 8; PS 15; PS 20
(PS28) Suseo	PS 9; PS 11; PS 13; PS 22	PS 5; PS 22	2 PS 6; PS 15; PS 22; PS 31	PS 1; PS 16; PS 20; PS 23
(PS29) Yangcheon	PS 29	2	PS 2; PS 6; PS 10; PS 30; PS 31	PS 20
(PS30)Yeongdeungpo	PS 9; PS 21; PS 27	PS 2; PS 5; PS 8; PS 27	4	PS 8; PS 20; PS 23
(PS31) Yongsan	PS 9; PS 13; PS 22	2	1	PS 8; PS 20; PS 23

Table A2

Benchmarks Determined by DN-DEA for 2014

2014	Financial	Learning	Business	Customer
(PS1) Bangbae	5	11	1	10
(PS2) Dobong	1	7	PS 6; PS 10; PS 16; PS 20; PS 22; PS 26	PS 14; PS 15; PS 20; PS 21
(PS3) Dongdaemun	PS 10; PS 11; PS 12; PS 13; PS 27	PS 3	1 PS 5; PS 6; PS 16; PS 17; PS 21; PS 30	PS 1; PS 15; PS 20
(PS4) Dongjak	PS 10; PS 11; PS 13; PS 27	PS 1; PS 2; PS 5; PS 17; PS 31	PS 6; PS 14; PS 15; PS 22; PS 24	PS 1; PS 15; PS 20
(PS5) Eunpyeong	PS 1; PS 10; PS 15; PS 22; PS 23	8	1 PS 3; PS 14; PS 16; PS 17; PS 25	2
(PS6) Gangbuk	PS 10; PS 11; PS 23; PS 27	2 PS 1; PS 2; PS 6; PS 17; PS 31	10	PS 17; PS 20
(PS7) Gangdong	PS 7; PS 10; PS 11; PS 13; PS 27	PS 1; PS 2; PS 5; PS 17; PS 31	PS 6; PS 21; PS 24; PS 30	PS 15; PS 20; PS 23
(PS8) Gangnam	1	1 PS 2; PS 9; PS 29	2 PS 6; PS 8; PS 30	PS 1; PS 8; PS 15
(PS9) Gangseo	PS 9	1	1 PS 8; PS 12; PS 16; PS 21; PS 30	PS 1; PS 15; PS 20; PS 23
(PS10) Geumcheon	18	PS 10	1	2
(PS11) Guro	10	1 PS 1; PS 5; PS 22; PS 31	PS 6; PS 16; PS 18; PS 21; PS 24	PS 1; PS 15; PS 20
(PS12) Gwanak	3	PS 12	1	PS 10; PS 16; PS 20
(PS13) Gwangjin	5 PS 10; PS 11; PS 27	PS 1; PS 6; PS 16; PS 17; PS 20	1	PS 15; PS 19; PS 23; PS 24
(PS14) Hehwa	PS 1; PS 10; PS 13; PS 15	2 PS 1; PS 5; PS 14; PS 17; PS 31	4	1
(PS15) Jongam	3 PS 1; PS 10; PS 15	PS 1; PS 5; PS 14; PS 17	4	16
(PS16) Jongno	PS 16	1 PS 1; PS 14; PS 17; PS 22	7	1
(PS17) Jungbu	PS 1; PS 10; PS 15; PS 22; PS 23	10	3	2
(PS18) Jungnang	PS 10; PS 11; PS 12; PS 27	PS 2; PS 5; PS 17; PS 31	3	PS 5; PS 15; PS 20; PS 23
(PS19) Mapo	PS 10; PS 11; PS 22; PS 27	PS 19	1 PS 6; PS 18; PS 30	PS 1; PS 15; PS 20; PS 23
(PS20) Namdaemun	PS 2; PS 10; PS 22	4	4	18

Table A2. *Benchmarks Determined by DN-DEA for 2014 (Continued)*

2014	Financial	Learning	Business	Customer
(PS21) Nowon	1 PS 10; PS 11; PS 22; PS 27	PS 21	4	3
(PS22) Seobu	10	5	5	PS 15; PS 20; PS 23
(PS23) Seocho	7	PS 1; PS 20; PS 22	PS 9; PS 15; PS 22; PS 28	9
(PS24) Seodaemun	PS 10; PS 22; PS 23; PS 27	PS 20; PS 22	4 PS 6; PS 8; PS 14; PS 15	3 PS 1; PS 20; PS 21; PS 24
(PS25) Seongbuk	PS 1; PS 10; PS 11; PS 22; PS 23	PS 1; PS 2; PS 5; PS 17; PS 31	1 PS 1; PS 17; PS 19; PS 20; PS 26; PS 27	PS 1; PS 20; PS 21
(PS26) Seongdong	PS 8; PS 10; PS 21; PS 22; PS 27	PS 2; PS 6; PS 17; PS 20	2 PS 6; PS 14; PS 15; PS 20; PS 22	PS 10; PS 15; PS 17; PS 20
(PS27) Songpa	PS 27	PS 27	1	PS 27
(PS28) Suseo	PS 10; PS 22; PS 23; PS 27	1 PS 5; PS 22; PS 29; PS 31	1 PS 13; PS 16; PS 18; PS 20; PS 22	PS 1; PS 15; PS 20
(PS29) Yangcheon	PS 29	2	PS 6; PS 16; PS 24	PS 5; PS 15; PS 20; PS 23
(PS30)Yeongdeungpo	PS 12; PS 13; PS 27; PS 30	PS 1; PS 8; PS 11; PS 28	5	PS 15; PS 23; PS 24
(PS31) Yongsan	PS 10; PS 11; PS 22; PS 23; PS 27	8	PS 31	PS 1; PS 15; PS 20; PS 23; PS 24

Table A3

Benchmarks Determined by DN-DEA for 2015

2015	Financial	Learning	Business	Customer
(PS1) Bangbae	PS 1	8	2	10
(PS2) Dobong	1	7	7	PS 1; PS 16; PS 20
(PS3) Dongdaemun	PS 12; PS 13; PS 27	1	2	PS 1; PS 16; PS 20
(PS4) Dongjak	PS 10; PS 12; PS 13	PS 1; PS 17; PS 20; PS 31	PS 1; PS 2; PS 6; PS 10; PS 15	PS 1; PS 8; PS 20
(PS5) Eunpyeong	PS 13; PS 22; PS 27	PS 1; PS 2; PS 20; PS 22	PS 2; PS 6; PS 9; PS 15; PS 16	PS 20
(PS6) Gangbuk	PS 10; PS 13; PS 22	PS 16; PS 17; PS 20; PS 31	9	PS 14; PS 16; PS 20
(PS7) Gangdong	1 PS 12; PS 13; PS 27	PS 16; PS 17; PS 20; PS 22	PS 6; PS 9; PS 16; PS 21; PS 27	PS 1; PS 8; PS 20
(PS8) Gangnam	PS 7; PS 11; PS 23; PS 27	PS 2; PS 16; PS 22; PS 29	PS 2; PS 9; PS 25; PS 27	12 PS 1; PS 8; PS 20
(PS9) Gangseo	PS 13; PS 22; PS 23; PS 27	PS 9	9	PS 8; PS 19
(PS10) Geumcheon	5	PS 10	4	2
(PS11) Guro	1	PS 2; PS 3; PS 13	PS 3; PS 9; PS 13	PS 8; PS 20; PS 22
(PS12) Gwanak	7	PS 1; PS 13; PS 31	PS 12	PS 12
(PS13) Gwangjin	14	2	1	PS 10; PS 20; PS 22; PS 23
(PS14) Hehwa	PS 10; PS 13; PS 22	PS 1; PS 16; PS 31	PS 14	PS 14
(PS15) Jongam	PS 10; PS 13; PS 22	PS 1; PS 20; PS 31	4	PS 1; PS 16; PS 20; PS 23
(PS16) Jongno	1	5	7	PS 16
(PS17) Jungbu	PS 10; PS 13; PS 22	4	PS 17	PS 1; PS 16; PS 20
(PS18) Jungnang	PS 13; PS 22; PS 27	PS 1; PS 2; PS 20; PS 31	PS 2; PS 3; PS 6; PS 10; PS 27	PS 20
(PS19) Mapo	PS 12; PS 13; PS 27	PS 19	PS 6; PS 9; PS 16; PS 21	2 PS 20; PS 22; PS 23
(PS20) Namdaemun	PS 2; PS 16; PS 22	7	PS 1; PS 2; PS 16; PS 25; PS 27	21

Table A3. *Benchmarks Determined by DN-DEA for 2015 (Continued)*

2015	Financial	Learning	Business	Customer
(PS21) Nowon	PS 21	1	2	PS 21
(PS22) Seobu	12	5	PS 22	4 PS 8; PS 20; PS 22
(PS23) Seocho	3	PS 1; PS 16; PS 17; PS 22	PS 2; PS 9; PS 16; PS 27	6 PS 1; PS 8; PS 20; PS 23
(PS24) Seodaemun	PS 22; PS 27; PS 29	1 PS 2; PS 21	PS 2; PS 6; PS 9; PS 10	PS 1; PS 8; PS 20
(PS25) Seongbuk	PS 12; PS 13; PS 22; PS 27	PS 25	2	PS 8; PS 20; PS 22; PS 23
(PS26) Seongdong	PS 13; PS 22; PS 27	PS 1; PS 2; PS 20; PS 31	PS 6; PS 10; PS 15	PS 20
(PS27) Songpa	12	1	6	PS 1; PS 8; PS 20
(PS28) Suseo	PS 22; PS 27; PS 29	PS 28	PS 6; PS 9; PS 15; PS 16	PS 28
(PS29) Yangcheon	2	2	PS 6; PS 9; PS 16; PS 27	PS 8; PS 20
(PS30)Yeongdeungpo	PS 12; PS 23; PS 27; PS 30	PS 2; PS 22; PS 24; PS 27; PS 29	PS 30	PS 8; PS 10; PS 19; PS 23
(PS31) Yongsan	PS 12; PS 13	PS 31	PS 31	PS 8; PS 20; PS 23

APPENDIX B. TARGET LEVELS DETERMINED BY DNSBM DEA

(2013-2015)

Table B1

Target Input Levels Determined by DN-DEA

DMU	Score	Target_Iuput_2013				Target_Input_2014				Target_Input_2015			
		①	⑤	⑨	⑬	①	⑤	⑨	⑬	①	⑤	⑨	⑬
Bangbae	1.00	823253.97	14.00	1470.00	13.00	844613.93	10.00	1480.00	13.00	805623.75	9.00	1261.00	8.00
Dobong	0.96	1092305.01	25.00	3067.88	21.00	1118834.05	21.00	3124.00	20.53	1152778.62	19.00	2664.00	8.14
Dongdaemun	0.83	1347478.67	50.00	4352.65	9.64	1325300.62	39.00	4363.00	10.19	1201989.53	39.00	4720.00	7.15
Dongjak	0.65	1168006.95	17.43	3671.11	12.00	1105912.57	18.14	3419.10	13.04	1012805.31	15.91	3338.06	7.12
Eunpyeong	0.86	946306.77	22.00	3111.70	9.00	908714.45	21.00	2741.00	9.00	828801.48	14.70	2565.57	5.00
Gangbuk	0.80	1144055.15	15.53	4020.00	12.94	1083306.37	27.85	4030.00	10.64	989018.62	27.63	4257.00	7.22
Gangdong	0.70	1252448.30	21.23	5142.02	9.00	1223745.80	19.04	5192.04	10.53	1094485.41	17.14	4746.48	5.60
Gangnam	0.80	1310491.12	15.27	5492.61	10.47	1344968.38	28.98	5303.02	14.76	1233790.01	16.94	4704.30	7.99
Gangseo	0.85	1408964.30	28.15	5852.41	10.40	1428588.28	33.00	5289.00	12.09	1340739.97	37.00	5585.00	15.07
Geumcheon	1.00	1027982.70	26.00	3960.00	19.00	942433.35	25.00	3842.00	19.00	887546.15	25.00	3781.00	12.00
Guro	0.76	1227089.40	20.89	4895.91	21.00	1254243.21	20.11	4786.38	11.31	1213857.00	27.06	5093.63	6.21
Gwanak	0.89	1499092.66	46.00	7237.00	15.28	1383543.26	40.00	6781.00	15.67	1314603.00	16.85	6345.00	24.00
Gwangjin	0.83	1180685.46	20.60	6135.00	13.36	1136475.63	16.58	6268.00	23.10	1010346.90	32.00	5909.00	8.71
Hehwa	0.97	957116.15	19.00	2628.00	21.00	953681.07	17.00	2632.00	21.00	813830.02	16.76	2233.00	12.00
Jongam	0.90	865371.19	16.68	2019.00	15.00	901360.35	16.84	2144.00	15.00	740870.37	17.21	1917.00	9.85
Jongno	0.93	1168553.08	14.31	2473.00	26.00	1319113.85	12.31	2389.00	26.00	1223207.92	18.00	2472.00	12.00
Jungbu	0.95	1000729.57	25.00	3128.00	18.00	936447.45	18.00	2932.00	18.00	837503.86	18.00	2860.00	9.63
Jungnang	0.77	1286847.88	26.62	5168.00	17.28	1253300.28	26.10	5353.00	11.71	1146978.36	23.63	5033.49	5.00
Mapo	0.79	1347260.61	35.00	4721.38	9.00	1325969.22	30.00	4868.76	11.43	1275174.84	34.00	4943.79	7.38
Namdaemun	0.97	961907.80	19.00	2112.00	9.00	1053858.52	18.00	2299.00	9.00	992236.49	18.00	1884.51	5.00
Nowon	0.97	1339779.12	34.00	5256.00	17.19	1321463.95	31.00	5312.00	24.00	1263559.00	31.00	5130.00	17.00
Seobu	0.98	848316.20	16.00	2244.00	15.00	800509.09	12.00	2400.00	13.93	730710.47	12.00	2085.00	6.76
Seocho	0.85	1202191.78	15.38	4000.75	15.00	1117631.53	12.51	3829.90	15.00	1025801.45	13.99	3808.82	8.25
Seodaemun	0.81	1183905.75	15.28	4156.13	22.00	1158021.15	17.31	3887.24	21.16	1105701.83	23.71	3533.39	5.92
Seongbuk	0.83	1031811.75	24.77	2405.42	10.00	958748.72	17.48	2462.00	12.97	893543.81	22.00	2237.00	6.09
Seongdong	0.73	1164990.30	21.47	2815.62	39.00	1127667.04	21.11	2946.28	16.05	1023620.81	18.08	3024.77	5.00
Songpa	0.89	1605631.68	53.00	7252.00	12.95	1680112.14	51.00	8004.00	32.00	1555658.68	49.00	6778.00	6.76
Suseo	0.85	1090449.29	19.35	3142.57	12.10	1067913.82	21.56	3083.15	12.58	1053236.05	25.00	3331.10	14.00

Table B1. *Target Input Levels Determined by DN-DEA (Continued)*

DMU	Score	Target_Iuput_2013				Target_Input_2014				Target_Input_2015			
		①	⑤	⑨	⑬	①	⑤	⑨	⑬	①	⑤	⑨	⑬
Yangcheon	0.80	1340131.56	34.00	4010.60	9.00	1313212.37	32.00	4100.57	10.67	1310173.40	29.00	4010.68	5.60
Yeongdeungpo	0.77	1542562.07	30.84	6863.00	11.51	1537017.90	31.96	7003.00	18.97	1478003.83	25.51	6867.00	9.94
Yongsan	0.86	1199269.49	38.00	3881.00	12.60	1158440.63	34.00	3799.00	18.03	1051348.62	33.00	3820.00	6.59

Table B2

Target Output Levels Determined by DN-DEA

DMU	Score	Target_Output_2013				Target_Output_2014				Target_Output_2015			
		②	⑥	⑩	⑭	②	⑥	⑩	⑭	②	⑥	⑩	⑭
Bangbae	1.00	347.23	71.60	792.00	74.60	348.97	75.60	813.00	76.00	355.60	75.00	736.00	74.10
Dobong	1.00	510.26	65.00	2033.00	80.90	525.27	80.10	1986.00	77.40	535.98	76.90	1900.00	78.73
Dongdaemun	0.99	745.84	82.20	2891.00	76.40	765.15	81.00	2882.00	75.83	779.27	87.70	3205.00	78.19
Dongjak	0.96	618.01	74.76	2564.00	77.30	635.48	79.30	2312.00	75.92	649.30	80.10	2397.00	74.65
Eunpyeong	0.96	452.07	81.50	2028.00	76.10	466.95	78.60	1827.00	75.80	473.29	75.51	1774.00	78.50
Gangbuk	0.96	604.06	73.56	2845.00	77.60	621.51	85.20	2806.00	75.85	638.12	81.10	3113.00	78.99
Gangdong	0.95	676.28	79.08	3234.00	76.10	704.04	78.69	3285.00	75.16	710.91	75.20	3171.00	77.25
Gangnam	0.95	706.81	65.93	3372.00	75.55	767.09	74.65	3329.00	73.95	778.37	71.80	3370.00	68.84
Gangseo	0.93	782.48	71.79	3604.00	76.33	809.32	71.60	3225.00	74.84	860.92	69.20	4190.00	71.91
Geumcheon	0.92	518.24	67.00	2677.00	76.50	534.23	81.40	2434.00	76.00	554.90	80.90	2796.00	75.30
Guro	0.90	670.17	78.85	3180.00	80.90	731.41	78.85	2985.00	75.83	745.57	83.50	3502.00	74.06
Gwanak	0.90	749.25	68.80	3718.00	77.50	775.59	78.40	3701.00	75.87	829.88	79.36	3712.00	68.90
Gwangjin	0.89	632.96	68.90	3704.00	77.60	656.75	77.38	3531.00	76.30	658.38	89.00	3707.00	74.20
Hehwa	0.89	457.69	78.00	1710.00	80.40	466.61	78.02	1848.00	77.00	475.17	72.51	1354.00	74.00
Jongam	0.89	389.17	74.80	1194.00	77.90	403.50	78.78	1548.00	75.80	402.71	75.06	1231.00	78.10
Jongno	0.88	609.66	70.24	1599.00	79.20	620.79	75.54	1580.00	75.20	633.18	67.70	1589.00	80.20
Jungbu	0.87	496.93	75.50	1839.00	78.90	497.99	79.80	1752.00	76.10	502.74	82.40	1716.00	79.07
Jungnang	0.86	706.53	83.55	3407.00	79.30	728.05	83.16	3259.00	75.40	736.38	79.96	3405.00	78.50
Mapo	0.85	744.19	69.70	3087.00	76.10	765.70	81.20	3136.00	75.69	820.52	74.10	3597.00	74.80
Namdaemun	0.85	455.18	66.00	1376.00	76.10	516.92	77.50	1436.00	75.80	502.21	74.50	1226.00	78.50
Nowon	0.84	737.69	72.50	3227.00	78.50	762.86	78.70	3300.00	78.40	773.29	79.70	3268.00	74.00
Seobu	0.84	380.67	62.70	1419.00	78.90	388.32	71.50	1498.00	74.81	392.18	74.60	1393.00	71.81
Seocho	0.83	619.18	68.33	2326.00	77.10	640.62	75.06	2309.00	72.30	651.21	71.40	2714.00	74.10
Seodaemun	0.83	627.73	65.88	2665.00	80.30	648.40	76.81	2588.00	77.60	656.94	78.00	2579.00	75.74
Seongbuk	0.82	503.33	75.99	1528.00	75.73	518.23	78.82	1480.00	76.40	524.77	69.20	1498.00	75.40
Seongdong	0.82	608.16	66.88	1857.00	81.00	624.39	80.45	2048.00	75.93	634.38	77.84	2123.00	78.50
Songpa	0.80	911.45	75.70	3403.00	76.10	990.62	71.90	3982.00	77.90	1004.35	74.80	4113.00	75.11
Suseo	0.80	557.17	73.21	2092.00	76.10	583.22	76.07	1933.00	75.96	599.10	71.80	2362.00	75.30

Table B2. *Target Output Levels Determined by DN-DEA (Continued)*

DMU	Score	Target_Output_2013				Target_Output_2014				Target_Output_2015			
		②	⑥	⑩	⑭	②	⑥	⑩	⑭	②	⑥	⑩	⑭
Yangcheon	0.75	701.24	72.90	2572.00	76.10	722.97	74.90	2654.00	75.05	733.70	71.00	2815.00	76.14
Yeongdeungpo	0.75	867.75	67.70	3862.00	74.88	890.08	69.98	4072.00	76.00	941.68	76.80	4154.00	71.00
Yongsan	0.69	643.13	91.10	2427.00	76.10	662.47	89.30	2340.00	75.60	681.49	86.80	2483.00	73.60

Table B3

Target Link and Carry-over Levels Determined by DN-DEA

DMU	Score	Target_Link_2013			Target_Link_2014			Target_Link_15			Target_Carry over_2013				Target_Carry over_14			
		③	⑦	⑪	③	⑦	⑪	③	⑦	⑪	④	⑧	⑫	⑮	④	⑧	⑫	⑮
Bangbae	1.00	10.00	72.60	200.00	10.00	70.50	238.00	10.00	78.20	231.00	8793.54	77.30	678.00	71.00	1073.83	77.20	667.00	73.30
Dobong	1.00	14.00	64.50	187.00	13.00	78.50	225.67	12.00	78.40	226.00	2175.31	70.35	1034.88	66.80	3040.70	79.94	1138.00	66.90
Dongdaemun	0.99	11.00	81.70	203.82	11.00	79.60	232.06	11.00	88.50	229.00	3754.36	80.67	1461.65	68.27	2953.83	81.16	1481.00	67.20
Dongjak	0.96	11.00	74.35	198.25	10.00	77.40	244.99	9.39	81.40	239.00	3252.43	78.24	1107.11	68.15	5616.14	80.68	1107.10	71.10
Eunpyeong	0.96	12.00	79.20	202.00	11.00	75.90	226.00	11.00	77.70	234.00	2919.60	81.06	1083.70	68.60	2197.95	78.65	914.00	65.80
Gangbuk	0.96	9.69	74.10	207.00	10.25	83.70	224.00	10.07	82.70	228.00	2465.74	78.09	1175.00	68.09	3405.91	87.08	1224.00	66.04
Gangdong	0.95	12.00	77.10	202.00	10.34	76.30	247.40	10.40	76.80	237.00	2759.67	79.78	1908.02	68.60	3393.98	79.46	1907.04	67.30
Gangnam	0.95	11.27	66.42	201.27	13.00	73.35	268.67	11.00	73.50	282.38	453.63	73.38	2120.61	69.48	522.23	76.14	1974.02	73.20
Gangseo	0.93	14.00	70.87	216.75	13.00	70.50	261.17	12.00	70.70	321.00	96.31	74.78	2248.41	69.28	1959.29	73.98	2064.00	69.66
Geumcheon	0.92	7.00	64.80	150.00	9.00	79.70	218.00	8.00	83.20	276.00	2339.56	74.12	1283.00	62.00	8610.00	82.54	1408.00	60.40
Guro	0.90	9.00	78.37	187.00	11.00	76.15	240.52	11.00	84.74	257.21	6402.57	80.79	1715.91	66.80	0.00	80.03	1801.38	67.60
Gwanak	0.90	9.00	64.80	173.00	10.00	75.50	220.00	10.00	81.90	201.00	1324.96	72.38	3519.00	65.32	7747.36	79.87	3080.00	62.90
Gwangjin	0.89	9.59	67.90	228.00	10.25	74.70	286.00	10.00	90.10	279.00	4682.50	74.98	2431.00	68.13	3496.00	79.82	2737.00	67.51
Hehwa	0.89	9.00	77.50	163.00	8.40	75.60	216.00	10.00	74.90	217.00	2660.88	79.87	918.00	64.00	4125.32	79.11	784.00	64.20
Jongam	0.89	10.66	74.70	214.00	8.94	77.10	270.00	10.81	77.30	238.00	4997.56	78.49	825.00	61.30	2452.56	80.45	596.00	68.40
Jongno	0.88	10.31	71.11	176.00	10.00	72.00	203.00	10.00	69.80	215.00	1219.76	76.36	874.00	73.40	8781.67	77.45	809.00	66.60
Jungbu	0.87	9.00	76.40	152.00	9.00	79.00	215.00	9.00	82.90	222.00	2670.53	78.25	1289.00	65.80	6409.88	81.79	1180.00	67.10
Jungnang	0.86	10.00	82.60	188.00	11.00	81.40	252.00	11.00	81.60	234.00	4372.08	84.55	1761.00	65.33	2739.11	84.31	2094.00	65.80
Mapo	0.85	12.00	70.70	202.00	12.00	80.20	240.90	11.00	75.60	272.96	2298.67	73.44	1634.38	68.60	3976.44	80.74	1732.76	69.01
Namdaemun	0.85	12.00	66.00	202.00	12.00	76.20	226.00	11.00	76.60	234.00	2309.46	73.64	736.00	68.60	4250.18	80.51	863.00	65.80
Nowon	0.84	10.00	72.40	205.00	12.00	75.40	232.00	12.00	82.50	206.00	2958.95	74.80	2029.00	69.40	3075.04	79.67	2012.00	69.80
Seobu	0.84	12.00	62.90	219.00	12.00	69.20	278.00	11.00	75.70	271.00	3439.26	71.15	825.00	65.50	2031.12	74.59	902.00	69.24
Seocho	0.83	11.00	68.78	265.00	11.00	71.61	326.00	10.00	73.90	256.38	456.44	74.85	1674.75	71.50	81.61	77.38	1520.90	73.00
Seodaemun	0.83	11.28	66.37	233.00	12.00	75.40	253.59	12.00	80.01	246.88	1935.32	73.35	1491.13	69.60	1671.00	79.83	1299.24	68.07
Seongbuk	0.82	13.00	74.42	201.50	11.00	75.70	229.43	11.00	73.20	256.00	2220.39	77.50	877.42	69.20	2403.75	79.80	982.00	68.00
Seongdong	0.82	13.00	66.33	179.00	12.00	79.00	237.00	11.00	79.88	234.00	1706.66	72.35	958.62	63.60	2529.31	81.03	898.28	65.80
Songpa	0.80	12.00	71.60	220.00	14.00	69.60	262.00	13.00	76.90	240.00	1907.88	76.22	3849.00	64.90	1455.66	74.36	4022.00	70.20
Suseo	0.80	12.00	72.01	217.25	12.00	74.20	238.29	12.00	74.40	264.00	2240.87	76.69	1050.57	70.10	1569.68	78.49	1150.15	72.00

Table B3. *Target Link and Carry-over Levels Determined by DN-DEA (Continued)*

DMU	Score	Target_Link_2013			Target_Link_2014			Target_Link_15			Target_Carry over_2013				Target_Carry over_14			
		③	⑦	⑪	③	⑦	⑪	③	⑦	⑪	④	⑧	⑫	⑮	④	⑧	⑫	⑮
Yangcheon	0.75	15.00	72.30	202.00	13.00	73.40	250.09	13.00	73.70	247.82	56.90	75.54	1438.60	68.60	981.57	76.53	1446.57	67.28
Yeongdeungpo	0.75	12.00	65.30	229.00	12.12	66.70	282.00	12.00	78.60	295.00	1373.90	73.11	3001.00	68.40	2827.94	76.63	2931.00	68.57
Yongsan	0.69	10.00	90.40	240.00	11.00	88.40	282.00	10.00	88.20	268.00	1947.80	91.93	1454.00	69.82	4068.92	91.76	1459.00	69.41

APPENDIX C. SLACKS (INEFFICIENCIES) DETERMINED BY DNSBM DEA

(2013-2015)

Table C1

Input Slacks Determined by DN-DEA

DMU	Score	Slack_Iuput_2013				Slack_Input_2014				Slack_Input_2015			
		①	⑤	⑨	⑬	①	⑤	⑨	⑬	①	⑤	⑨	⑬
Bangbae	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Dobong	0.96	0.00	0.00	-162.12	0.00	0.00	0.00	0.00	-0.47	0.00	0.00	0.00	-5.86
Dongdaemun	0.83	-115422.98	0.00	-537.35	-15.36	-80467.53	0.00	0.00	-14.81	-150236.43	0.00	0.00	-7.85
Dongjak	0.65	-68381.90	-22.57	-753.89	-22.00	-92636.45	-16.86	-449.90	-20.96	-144790.11	-19.09	-735.94	-10.88
Eunpyeong	0.86	-56534.72	0.00	-58.30	-4.00	-81383.40	0.00	-290.00	-4.00	-154983.12	-4.30	-94.43	-4.00
Gangbuk	0.80	-136176.75	-17.47	0.00	-8.06	-133897.67	-0.15	0.00	-10.36	-204135.23	-3.37	0.00	-6.78
Gangdong	0.70	-52044.80	-18.77	-391.98	-17.00	0.00	-11.96	-199.96	-15.47	-125077.59	-13.86	-497.52	-11.40
Gangnam	0.80	-14658.34	-19.73	-520.39	-7.53	0.00	-13.02	-121.98	-3.24	-125822.99	-25.06	-489.70	-0.01
Gangseo	0.85	0.00	-7.85	-123.59	-19.60	0.00	0.00	0.00	-17.91	-56798.10	0.00	0.00	-4.93
Geumcheon	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Guro	0.76	0.00	-14.11	-769.09	-9.00	0.00	-13.89	-602.62	-18.69	0.00	-5.94	-552.37	-9.79
Gwanak	0.89	0.00	0.00	0.00	-9.72	0.00	0.00	0.00	-9.33	0.00	-27.15	0.00	0.00
Gwangjin	0.83	0.00	-14.40	0.00	-13.64	-52291.75	-16.42	0.00	-3.90	0.00	0.00	0.00	-6.29
Hehwa	0.97	-106729.65	0.00	0.00	0.00	-82046.27	0.00	0.00	0.00	-131805.28	-0.24	0.00	0.00
Jongam	0.90	-145090.89	-5.32	0.00	0.00	-68015.47	-3.16	0.00	0.00	-132400.92	-6.79	0.00	-1.15
Jongno	0.93	-264086.22	-6.69	0.00	0.00	0.00	-5.69	0.00	0.00	0.00	0.00	0.00	0.00
Jungbu	0.95	-72739.28	0.00	0.00	0.00	-94604.49	0.00	0.00	0.00	-212146.91	0.00	0.00	-2.37
Jungnang	0.77	-112997.76	-7.38	0.00	-17.72	-174024.26	-3.90	0.00	-23.29	-125861.58	-5.37	-159.51	-12.00
Mapo	0.79	-7050.43	0.00	-690.62	-20.00	-92796.56	0.00	-871.24	-17.57	-118091.72	0.00	-910.21	-11.62
Namdaemun	0.97	-180663.70	0.00	0.00	0.00	-45888.65	0.00	0.00	0.00	-118700.66	0.00	-209.49	0.00
Nowon	0.97	-25658.48	0.00	0.00	-6.81	-940.21	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Seobu	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-1.07	0.00	0.00	0.00	-1.24
Seocho	0.85	-69801.60	-11.62	-272.25	0.00	0.00	-13.49	-248.10	0.00	0.00	-12.01	-374.18	-0.75
Seodaemun	0.81	-30818.96	-20.72	-130.87	0.00	-41587.99	-17.69	-306.76	-0.84	-46529.62	-11.29	-495.61	-7.08
Seongbuk	0.83	-214278.65	-0.23	-342.58	-7.00	-248961.63	-6.52	0.00	-4.03	-115155.07	0.00	0.00	-4.91
Seongdong	0.73	-115093.10	-16.53	-391.38	0.00	-147838.56	-10.89	-635.72	-22.95	-192882.99	-13.92	-333.23	-13.00
Songpa	0.89	0.00	0.00	0.00	-19.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-14.24
Suseo	0.85	-120260.61	-9.65	-521.43	-9.90	-99654.48	-1.44	-342.85	-9.42	-77524.23	0.00	-91.90	0.00

Table C1. *Input Slacks Determined by DN-DEA (Continued)*

DMU	Score	Slack_Iuput_2013				Slack_Input_2014				Slack_Input_2015			
		①	⑤	⑨	⑬	①	⑤	⑨	⑬	①	⑤	⑨	⑬
Yangcheon	0.80	0.00	0.00	-607.40	-18.00	0.00	0.00	-541.43	-16.33	0.00	0.00	-517.32	-14.40
Yeongdeungpo	0.77	-17083.93	-20.16	0.00	-22.49	0.00	-11.04	0.00	-15.03	-25704.33	-20.49	0.00	-13.06
Yongsan	0.86	-153462.07	0.00	0.00	-11.40	-89670.41	0.00	0.00	-5.97	-231254.10	0.00	0.00	-10.41

Table C2

Output Slacks Determined by DN-DEA

DMU	Score	Slack_Output_2013				Slack_Output_2014				Slack_Output_2015			
		②	⑥	⑩	⑭	②	⑥	⑩	⑭	②	⑥	⑩	⑭
Bangbae	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Dobong	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.73
Dongdaemun	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.23	0.00	0.00	0.00	1.39
Dongjak	0.96	0.00	1.96	0.00	0.00	0.00	0.00	0.00	1.52	0.00	0.00	0.00	5.15
Eunpyeong	0.96	0.00	0.00	0.00	3.40	0.00	0.00	0.00	0.40	0.00	0.61	0.00	3.30
Gangbuk	0.96	0.00	0.76	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	4.39
Gangdong	0.95	0.00	4.88	0.00	1.10	0.00	0.79	0.00	0.46	0.00	0.00	0.00	7.05
Gangnam	0.95	0.00	3.63	0.00	4.85	0.00	7.55	0.00	3.85	0.00	1.20	0.00	2.14
Gangseo	0.93	0.00	3.19	0.00	1.03	0.00	0.00	0.00	2.44	0.00	0.00	0.00	1.01
Geumcheon	0.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Guro	0.90	0.00	2.15	0.00	0.00	0.00	7.95	0.00	2.83	0.00	0.00	0.00	0.56
Gwanak	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.07	0.00	0.46	0.00	0.00
Gwangjin	0.89	0.00	0.00	0.00	0.00	0.00	0.48	0.00	0.00	0.00	0.00	0.00	0.00
Hehwa	0.89	0.00	0.00	0.00	0.00	0.00	0.52	0.00	0.00	0.00	1.61	0.00	0.00
Jongam	0.89	0.00	0.00	0.00	0.00	7.25	0.78	0.00	0.00	0.00	0.66	0.00	0.00
Jongno	0.88	0.00	5.64	0.00	0.00	0.00	2.24	0.00	0.00	0.00	0.00	0.00	0.00
Jungbu	0.87	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.17
Jungnang	0.86	0.00	0.25	0.00	0.00	0.00	0.46	0.00	0.00	0.00	0.36	0.00	2.70
Mapo	0.85	0.00	0.00	0.00	3.60	0.00	0.00	0.00	1.49	0.00	0.00	0.00	0.00
Namdaemun	0.85	0.00	0.00	0.00	0.00	43.75	0.00	0.00	0.00	14.90	0.00	0.00	0.00
Nowon	0.84	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Seobu	0.84	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00	2.91
Seocho	0.83	0.00	5.33	0.00	0.00	0.00	5.46	0.00	0.00	0.00	0.00	0.00	0.00
Seodaemun	0.83	0.00	4.28	0.00	0.00	0.00	7.81	0.00	0.00	0.00	0.00	0.00	2.44
Seongbuk	0.82	0.00	1.59	0.00	2.83	0.00	0.82	0.00	0.00	0.00	0.00	0.00	0.00
Seongdong	0.82	0.00	1.28	0.00	0.00	0.00	0.15	0.00	2.03	0.00	0.24	0.00	3.70
Songpa	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.01
Suseo	0.80	0.00	1.31	0.00	0.00	0.00	1.87	0.00	5.56	0.00	0.00	0.00	0.00

Table C2. *Output Slacks Determined by DN-DEA (Continued)*

DMU	Score	Slack_Output_2013				Slack_Output_2014				Slack_Output_2015			
		②	⑥	⑩	⑭	②	⑥	⑩	⑭	②	⑥	⑩	⑭
Yangcheon	0.75	0.00	0.00	0.00	0.30	0.00	0.00	0.00	0.45	0.00	0.00	0.00	2.34
Yeongdeungpo	0.75	0.00	0.00	0.00	6.18	0.00	0.48	0.00	0.00	0.00	0.00	0.00	0.00
Yongsan	0.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table C3

Link and Carry-over Slacks Determined by DN-DEA

DMU	Score	Slack_Link_2013			Slack_Link_2014			Slack_Link_15			Slack_Carry over_2013				Slack_Carry over_14			
		③	⑦	⑪	③	⑦	⑪	③	⑦	⑪	④	⑧	⑫	⑮	④	⑧	⑫	⑮
Bangbae	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	78.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Dobong	1.00	0.00	0.00	0.00	0.00	0.00	-4.33	0.00	78.40	0.00	0.00	0.00	-162.12	0.00	0.00	0.00	0.00	0.00
Dongdaemun	0.99	0.00	0.00	-16.18	0.00	0.00	-43.94	0.00	88.50	0.00	3588.89	0.00	-537.35	6.97	2265.35	0.00	0.00	0.00
Dongjak	0.96	0.00	2.65	-28.75	0.00	0.00	-25.01	0.39	81.40	0.00	2525.61	0.00	-753.89	0.25	4306.57	0.49	-449.90	0.00
Eunpyeong	0.96	0.00	0.00	-16.00	0.00	0.00	-35.00	0.00	77.70	-24.00	2808.80	0.00	-58.30	6.40	1841.92	0.00	-290.00	1.80
Gangbuk	0.96	0.69	0.00	0.00	0.25	0.00	0.00	0.07	82.70	0.00	863.11	2.32	0.00	3.99	1381.59	0.00	0.00	3.54
Gangdong	0.95	0.00	3.20	-12.00	0.34	0.00	-10.60	0.40	76.80	0.00	2665.38	0.00	-391.98	1.90	3392.50	0.00	-199.96	0.00
Gangnam	0.95	0.27	7.32	-97.73	0.00	9.05	-56.33	0.00	73.50	-20.62	0.00	2.61	-520.39	2.48	0.00	0.00	-121.98	0.00
Gangseo	0.93	0.00	3.37	-11.25	0.00	0.00	-81.83	0.00	70.70	0.00	0.00	0.00	-123.59	2.28	0.00	0.00	0.00	6.66
Geumcheon	0.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	83.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Guro	0.90	0.00	1.37	-15.00	0.00	7.45	-63.48	0.00	84.74	-49.79	0.00	0.00	-769.09	3.10	0.00	4.37	-602.62	0.00
Gwanak	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	81.90	0.00	0.00	0.00	0.00	4.02	0.00	0.00	0.00	0.00
Gwangjin	0.89	0.59	0.00	0.00	2.25	0.00	0.00	0.00	90.10	0.00	2451.28	2.58	0.00	1.23	0.00	0.00	0.00	3.51
Hehwa	0.89	0.00	0.00	0.00	1.40	0.00	0.00	0.00	74.90	0.00	2520.96	0.00	0.00	0.00	2450.26	0.76	0.00	0.00
Jongam	0.89	2.66	0.00	0.00	1.94	0.00	0.00	2.81	77.30	0.00	4621.04	0.00	0.00	0.00	2452.07	0.51	0.00	0.00
Jongno	0.88	1.31	7.21	0.00	0.00	0.00	0.00	0.00	69.80	0.00	0.00	6.82	0.00	0.00	0.00	0.00	0.00	0.00
Jungbu	0.87	0.00	0.00	0.00	0.00	0.00	0.00	0.00	82.90	0.00	585.94	0.00	0.00	0.00	6113.44	0.00	0.00	0.00
Jungnang	0.86	0.00	0.00	0.00	0.00	0.00	0.00	0.00	81.60	-14.00	3454.05	0.32	0.00	5.33	1564.69	0.57	0.00	1.20
Mapo	0.85	0.00	0.00	-11.00	0.00	0.00	-81.10	0.00	75.60	-58.04	1228.83	0.00	-690.62	4.20	2671.35	0.00	-871.24	1.31
Namdaemun	0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	76.60	0.00	2029.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nowon	0.84	0.00	0.00	0.00	0.00	0.00	0.00	0.00	82.50	0.00	2938.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Seobu	0.84	0.00	0.00	0.00	0.00	0.00	0.00	0.00	75.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.24
Seocho	0.83	0.00	6.48	0.00	0.00	4.11	0.00	0.00	73.90	-47.62	0.00	3.71	-272.25	0.00	0.00	3.41	-248.10	0.00
Seodaemun	0.83	1.28	6.27	0.00	0.00	8.50	-14.41	0.00	80.01	-14.12	355.46	5.56	-130.87	0.00	1294.21	7.47	-306.76	1.27
Seongbuk	0.82	0.00	0.12	-24.50	0.00	0.00	-29.57	0.00	73.20	0.00	788.82	0.00	-342.58	0.00	0.00	0.00	0.00	5.40
Seongdong	0.82	0.00	1.03	0.00	0.00	0.00	0.00	0.00	79.88	-19.00	552.51	0.92	-391.38	0.00	2523.31	0.00	-635.72	1.80
Songpa	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	76.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table C3. *Link and Carry-over Slacks Determined by DN-DEA (Continued)*

DMU	Score	Slack_Link_2013			Slack_Link_2014			Slack_Link_15			Slack_Carry over_2013				Slack_Carry over_14			
		③	⑦	⑪	③	⑦	⑪	③	⑦	⑪	④	⑧	⑫	⑮	④	⑧	⑫	⑮
Suseo	0.80	0.00	0.91	-65.75	0.00	1.80	-107.71	0.00	74.40	0.00	2064.63	0.00	-521.43	0.00	1156.83	0.00	-342.85	0.00
Yangcheon	0.75	0.00	0.00	-15.00	0.00	0.00	-43.91	0.00	73.70	-46.18	0.00	0.00	-607.40	2.50	0.00	0.00	-541.43	0.28
Yeongdeungpo	0.75	0.00	0.00	0.00	0.12	0.00	0.00	0.00	78.60	0.00	1294.51	1.26	0.00	6.70	1367.04	2.02	0.00	3.87
Yongsan	0.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00	88.20	0.00	715.30	0.00	0.00	2.02	2845.06	0.00	0.00	3.81