DYSFUNCTIONAL CONSEQUENCES OF THE KOREAN PERFORMANCE BUDGETING SYSTEM AND THEIR POLICY IMPLICATIONS

By

SANG HOON SHIN

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Institute of Local Government Studies
School of Government and Society
The University of Birmingham
March 2013

ABSTRACT

In 2005, the South Korean government implemented a system of performance budgeting: Self-Assessment of Budgetary Programmes (SABP). Most studies on this system have focused on the relationship between SABP results and subsequent budget allocations. These studies are based on the premise that the SABP system itself is operating well, and consequently SABP results are reliable. However, this thesis questions that premise and analyses the process for arriving at SABP results, especially focusing on differences in views on the merit of programmes between the spending ministries and the Ministry of Strategy and Finance (MOSF), which controls SABP. The thesis addresses four key research questions: Which factors affect differences in views in the SABP process? What are the dysfunctional consequences of SABP? What is the impact of these dysfunctional consequences? And, lastly, what feasible policy alternatives can be proposed? The study suggests that there is a tendency to optimism bias by spending ministries in their self-assessment programmes, often leading to a subsequent drastic downward review of such assessments by the MOSF. These results are established by both quantitative and qualitative analysis. This thesis also provides evidence of dysfunctional effects arising from the SABP process, some of which are "unintended" by both spending ministries and the MOSF, while others are "unintended" by the designers of the SABP system but are likely to be "intended" by the spending ministries, as "agents" in the principal-agent relationship. The thesis concludes that both the unintended and intended dysfunctional consequences of SABP are sufficiently important to suggest that the performance budgeting system needs to be carefully re-designed, and proposals are made for feasible refinements to the SABP process.

DEDICATION

I WOULD LIKE TO DEDICATE MY THESIS TO MY BELOVED FAMILY

TO MY DEVOTED WIFE, HYE-SEONG MUN

WHO GAVE ME HER WHOLEHEARTED SUPPORT

DURING MY LONG STUDY JOURNEY

TO MY DAUGHTER, JU-JIN AND SON, DONG-YUN

WHO HAVE BEEN A GREAT SOURCE OF PLEASURE AND MOTIVATION

TO MY PARENTS, YOUNGER SISTER AND BROTHER

WHO GAVE ME UNCONDITIONAL LOVE AND SACRIFICE

AND

TO MY PARENTS-IN-LAW, SISTER-IN LAW AND BROTHER-IN-LAW

WHO GAVE ME TREMENDOUS ENCOURAGEMENT AND SUPPORT

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my supervisors, Professor Tony Bovaird and Reader Peter A. Watt, for helping me to complete this PhD thesis successfully. They always conveyed their expertise, inspiration and a spirit of adventure with regard to my research. Without their outstanding insight and guidance, this thesis would not have been possible.

I would like to thank all the staff at INLOGOV, especially Professor John W. Raine and Dr Phillip Whiteman, for giving me valuable comments. Also, I am very thankful to Professor Andy Neely at Cambridge University, my external examiner, for great comments at final stage. I am extremely grateful for Helen Hancock's suggestions on the language throughout my thesis. I am also thankful to all my colleagues, especially to Cheol-eon Lim, Jun-hyung Park, Sung-won Lee, Dumisani Jantjies and Dr Zahira Latif.

In addition, I thank Professor Hyung-koo Mun at Korea University, former Vice-minister Su-won Lee, and Dong-sun Kim for words of encouragement. I am especially indebted to Dr Sang-hun Shin, Dr Young-sun Ko, Professor Jae-duck Keum, and Dr No-wook Park who contributed greatly to enabling me to complete the thesis. I am also grateful to all the people who gave their time and experiences during the interviews.

Finally, I would like to thank the Korean Government and the KIAT, which gave me financial support to study in the UK. In particular, I am thankful to Pastor Min Lee, Sung-bin Han, Sung-il Ha and Sang-yun Lee, who have always prayed for my family, and to all the Korean families at Birmingham Jesus Vision Church who shared the time and place with my family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BAI	Board of Audit and Inspection of Korea
DR	Disagreement Ratio
GAO	The United States Government Accountability Office
GPA	Government Performance Assessment Act
GPRA	Government Performance and Results Act
GPRAMA	GPRA Modernization Act
IES	In-depth Evaluation System
KDI	Korea Development Institute
KIPF	Korea Institute of Public Finance
MBO	Management by Objectives
MEST	Ministry of Educations, Science and Technology
MKE	Ministry of Knowledge and Economy
MLTM	Ministry of Land, Transport and Maritime Affairs
MOEL	Ministry of Employment and Labour
MOSF	Ministry of Strategy and Finance
MW	Ministry of Health and Welfare
NABO	National Assembly Budget Office
NFA	National Finance Act
NIE	New Institutional Economics
NPM	New Public Management
NSTC	National Science and Technology Commission
OECD	Organisation for Economic Co-operation and Development
OMB	Office of Management and Budget
PART	Program Assessment Rating Tool
PMS	Performance Management System of Programmes
PPB	Programme Budgeting
ROR	Relative Optimism Ratio
SABP	Self-Assessment of Budgetary Programmes
ZBB	Zero Base Budgeting

CHAPTER 1

INTRODUCTION

1.1 Background

Performance management of public services has spread widely across national boundaries in recent years. It arose widely around the late 1980s or early 1990s with the advent of New Public Management (NPM); however, performance management is not a new concept and its antecedents can be traced back to long before the recent era (Talbot, 2010; Johnsen, 2005; Behn, 2003). The US Congress established the Bureau of Efficiency to increase the efficiency of the US Federal Government's expenditure in 1916 (Lee, 2006); and performance management was used as the measurement method for US metropolitan cities' activities in the 1940s (Ridley and Simon, 1943). More recently, performance management has been developed in the form of Programme and Performance Budgeting (PPB) in the 1960s, and Management by Objectives (MBO) and Zero Base Budgeting (ZBB) in the 1970s. The Government Performance and Results Act (GPRA) in the 1990s and the Program Assessment Rating Tool (PART) in the 2000s have also been used as measures for improving the efficiency of government expenditure (Schick, 1992; Ko, 2004).

Performance management in the public sector is a form of results-oriented management, focusing on the output, outcome and influence per input. Performance management can be seen as including the budgeting system as well as the organisational, personnel and accounting systems of an organisation, while performance budgeting[1] has the most important role in these processes (Ko, 2004). Many OECD (Organisation for Economic Co-operation and Development) countries which had suffered from fiscal deficits and subsequent high central government debts[2] caused by the excessive requirements of welfare expenditures introduced performance budgeting from the 1980s onwards. NPM argues for the necessity of making a link between performance and budgeting in a process of controlling public expenditures. Under the influence of this view, advanced countries such as the UK, the US, Australia and New Zealand started to institutionalise performance budgeting (Shah, 2007). In the late 1990s, many developing countries also accepted performance budgeting as an important part of their national agenda (Mackay, 2007). The reason why so many countries from the OECD to South America, Africa and the developing regions have introduced performance budgeting systems relates to the positive effects attributed to such systems. It is argued that a government can increase the efficiency, effectiveness, accountability, transparency, and democratic control of policies by applying performance information to strategic planning, organisational and personnel management, budgeting, service delivery improvement, and evaluation (Shah, 2007; Epstein, 1992; Wholey and Hatry, 1992; Wang, 2002; Ammons, 1995a; Mackay, 2007).

[1] Performance budgeting has been implemented under various names, such as performance-based budgeting, programme budgeting, result-oriented management, performance evaluation and management, monitoring and evaluation, results-based management, etc. This study uses the phrase performance budgeting as a concept that includes these various terms.

[2] Predicted average central government debt as percentage of GDP in 2011 in OECD countries is 103.0% (OECD, 2012)

After the Asian financial crisis in the late 1990s, the South Korean government also launched a performance budgeting system to overcome the problems it faced in its efforts to reform the public sector and to improve the effectiveness and efficiency of budgetary programmes. The Budget Office of Korea[3] adopted performance budgeting on a trial basis in 1999 and expanded it into the Performance Management System of Programmes (PMS) in 2003. Later, in 2005, Self-Assessment of Budgetary Programmes (SABP) was implemented in order to forge a link between the performance results of a programme and its budgeting. The performance budgeting system took its legal basis from the National Finance Act (NFA) of 2006. In the same year, the Prime Minister's Office (PMO) enacted the Government Performance Assessment Act (GPA) which specifies the concepts, principles and objectives of government performance management more broadly than the NFA, which treats Korean performance budgeting more directly and specifically (Cho, 2010). The NFA will be discussed in more detail later. SABP, which has played the most important role among the various performance management systems in Korea, was modelled on the PART employed by the United States and has been used in Korea since 2005 along with other fiscal reforms[4]. As so-called the "top-down budgeting" has been introduced, every spending ministry has been able to propose a budget with greater autonomy than in the past, whilst adhering to the firm annual spending ceiling set by the Ministry of Strategy and Finance (MOSF); so the stricter control of spending ministries' expenditures has become more important.

[3] This has been transformed into the Planning and Budget Commission (1998.2 – 1999.5), the Ministry of Planning and Budget (1999.5 – 2008.2) and the Ministry of Strategy and Finance (2008.2 – present) as a result of government reorganisations.

[4] The Korean Government's four major fiscal reforms at that time were "a medium-term expenditure framework", "top-down budgeting", "a performance budgeting system", and "a digital budget information system" (Bang, 2009).

SABP has been used to assess a third of all the programmes of each spending ministry every year and has included mechanisms to reflect these assessment results in the ministries' budget for the next year. From 2005 to 2010, a total of 2,920 programmes were assessed by SABP and the budgets for them totalled about 220,000 billion Won[5].

1.2 Research Objectives and Research Questions

Performance budgeting has been the subject of considerable analysis, most of which has focused on the relationship between performance results and subsequent budget allocations (Gilmour and Lewis, 2005, 2006; GAO, 2001; Melkers and Willoughby, 2001; Park, 2005, 2008; Bang, 2009; Cho, 2010) or the factors that affect budgeting and performance assessment results (Kang, 2007; Chang and Yoon, 2002; Yoon, 2001). These approaches are based on the premise that the performance budgeting system itself is operating well, and consequently its results are reliable and accurately reflect a programme's performance. However, in this thesis we ask whether performance budgeting systems in the public sector do genuinely indicate the success or failure of a programme. Are these systems measuring the actual performance of a programme accurately? Are they achieving their purposes successfully? Unfortunately, it is very hard to answer 'Yes' to these questions because many previous studies on performance management or budgeting in the public sector have suggested that there might be various kinds of dysfunctional consequences stemming from the restraints and limitations involved when performance is measured in this way (Ridgway, 1956; Weitzman, 1980; Smith, 1995; Bevan and Hood, 2006; de Bruijn, 2002; Kelman and

[5] The Won is the unit of Korean currency. One Korean Won is equivalent to nearly 0.0006 UK pounds, as of February 2013.

Friedman, 2009; Bouckaert and Balk, 1991; LeGrand and Bartlett, 1993; Radnor, 2008; Thiel and Leeuw, 2002; Kong et al., 2009; Shin, 2010). These dysfunctional consequences can take many different forms, such as problems stemming from focusing on short-term performance, from focusing on only measured performance, from deliberately setting a lower performance target level, and from manipulating or interpreting performance in a favourable way. Despite the importance of performance management, the approach is not a panacea. Indeed, performance management may raise new problems or worsen existing ones. It can be a major obstacle to the successful modernisation of the public sector if it is inappropriately established and operated (Bouckaert and Peters, 2002). Ignoring these dysfunctional consequences or treating them as trivial phenomena can not only undermine the objectivity and reliability of performance measurement, but can also weaken the degree of acceptance of the results of performance assessment, and may, therefore, prevent performance budgeting from fully achieving its aims in the long term. Exploring the dysfunctional consequences of performance budgeting may therefore enable policy makers to identify the cause of these consequences systematically and design feasible improved policy alternatives.

Against this overall background, this study examines differences of view between spending ministries and the MOSF that emerge in the process of reaching SABP results. As SABP consists of two steps – the spending ministries' self-assessment stage and the MOSF's review stage – differences between the SABP scores allocated by the spending ministries and the scores allocated by the MOSF are, in a sense, inevitable. In addition, considering the nature of performance budgeting, which gives incentives or penalties according to the performance of a programme, it can be said that this difference

between spending ministries and the MOSF is not an unexpected phenomenon. However, if particular differences have been significant, persistent and widespread between ministries[6], it is important to analyse how this has arisen. The existence of and reasons for such differences are examined on the basis of two different approaches. The first focuses on explaining the factors which may affect the differences between SABP scores allocated by spending ministries and scores allocated by the MOSF by using numerical data drawn from SABP results covering 2005 to 2010. The second approach explores a range of dysfunctional consequences of performance budgeting in the SABP context and the impact of these consequences on the behaviours of spending ministries and the MOSF. It does so by analysing qualitative data obtained from various documents, archival records and in-depth interviews. After analysing the existence and reasons for differences between spending ministries and the MOSF in the process of SABP, the study will suggest feasible policy alternatives for improving SABP.

This is the background to the following research objectives.

- To explain the factors which affect the differences in the SABP score between spending ministries and the MOSF
- To examine the types of, extent of, and reasons for dysfunctional consequences of SABP

[6] As regards this, the SABP score gap between spending ministries' self-assessment and the MOSF's review was 26.1, 26.6, 25.7, 18.9, 26.3, and 25.8 points on the basis of 100 points in the years 2005, 2006, 2007, 2008, 2009, and 2010 respectively. Furthermore, the average of spending ministries' self-assessment score for the six years is nearly 90 points (MOSF, 2012a).

- To explore the influences of these dysfunctional consequences of SABP on spending ministries' self-assessment scores and the MOSF's review scores

- To suggest the role of spending ministries and the MOSF in improving the SABP system

These research objectives will be addressed through the four research questions.

(1) What are the factors which affect the differences in SABP results between spending ministries and the MOSF?

(2) What are the types of, extent of, and reasons for dysfunctional consequences of SABP?

(3) How do these dysfunctional consequences impact on both spending ministries' self-assessment scores and the MOSF's review scores?

(4) How can the SABP system in Korea be improved and refined?

These questions are based on the relevant SABP results and on previous studies of dysfunctional consequences of performance budgeting in the public sector. The first research question gives rise to much of the quantitative analysis and leads to an examination of whether factors within the SABP system, such as the MOSF's review score, the budget percentage change of a programme, the programme type, the organisation type and the programme purpose can impact on the differences between the two scores by employing numerical data drawn from SABP results covering six years.

Answers to this question will be linked with the answers to the following research questions.

The second research question aims to explore whether various dysfunctional consequences actually occur in the SABP system, and if so, to what extent and for what reasons. This question is investigated over eight programmes in three case study areas. In particular, the thesis categorises diverse dysfunctional consequences in spending ministries into "unintended" and "intended", based on a structure informed by the principal-agent relationship.

The third research question is designed to investigate the impact of these dysfunctional consequences, and focuses particularly on spending ministries' "optimism bias" as manifested by spending ministries' high self-assessment scores and the MOSF's subsequent drastic cutting of these. High, significant, persistent and widespread differences during the implementation of SABP can be successfully explained by examining the impact of dysfunctional consequences; and they also indicate the need for feasible policy alternatives to prevent or reduce them.

The last research question concerns ways in which dysfunctional consequences of performance budgeting might be reduced by feasible changes to the SABP system. This question stems from the fact that the most important thing in performance budgeting is not the existence of dysfunctional consequences but the question of what feasible policies can be devised to solve such problems. Also, the thesis suggests that policy makers should be cautious both about intended and unintended dysfunctional

consequences which are already evident in SABP and about those that subsequent policy changes might bring.

1.3 The Scope of the Research

This thesis will discuss dysfunctional consequences of performance budgeting in the public sector broadly, but its empirical analyses will focus on the SABP system in Korea. Since SABP has both characteristics that relate to managing budgetary programmes' performance and ones that relate to assessing government performance, the GPA and the NFA are both relevant to the SABP system. The GPA comprehensively stipulates the concepts, principles and objectives of "government performance management", changing from input- and process-oriented to outcome-oriented management to increase the efficiency and effectiveness of government (Art. 2). As government performance management includes financial, personnel and organisational tasks, SABP is one part of this integrated government performance management in the GPA. At the same time, the performance budgeting system includes three sub-systems: the Performance Management System for Programmes (PMS), Self-Assessment of Budgetary Programmes (SABP), and the In-depth Evaluation System (IES). Consequently, SABP, the research object, is broadly related to government performance management in the GPA, and at the same time it is one of the performance budgeting systems in the NFA.

There are two types of data analysed in this thesis: quantitative and qualitative. The quantitative data are analysed in order to explain the factors which affect the differences

in SABP score between spending ministries and the MOSF, as well as to give some useful indications for policy alternatives; while the qualitative data are examined in order to explore the types of, extent of and reasons for various dysfunctional consequences of the SABP system, their influences, and feasible policy alternatives. For the quantitative data, 2,920 budgetary programmes' SABP results in 49 central government departments from 2005 to 2010 are used. These SABP scores include both spending ministries' self-assessment scores and the MOSF's review scores for each programme. For the qualitative data, documents, archival data, and two sets of in-depth interviews with Korean government officials and experts in eight budgetary programmes covering three case study areas (public housing, youth employment and social overhead capital) are used. Interviewees were chosen because they all had experience of SABP and could give useful insights for analysing dysfunctional consequences of the SABP system and the policy implications for it.

1.4 Methodology

The objectives of the study are both to explain factors which affect the differences in SABP scores through a quantitative approach and to explore dysfunctional consequences of the SABP system and the impact of these on SABP scores through a qualitative approach. In other words, these aims include not only verifying the relationships between some factors and differences in SABP scores but also understanding the subjective meanings of government officials' perceptions and behaviours. This approach does not "mix" different philosophical worldviews for study purposes, but uses an umbrella paradigm of pragmatism (Creswell and Plano Clark,

2011). Following the assumptions of pragmatism, the focus of the research is on the consequences of study, on the use of multiple methods as well as different forms of data collection and analysis, and on the "what" and "how" of research (Creswell, 2009; Cherryholmes, 1992; Morgan, 2007). Pragmatism as a worldview (Creswell, 2009) or a paradigm (Kuhn, 1970) for a study is closely associated with the mixed approach, employing both quantitative and qualitative analysis (Creswell and Plano Clark, 2011). Taking pragmatism as a philosophical worldview, the study collects both quantitative and qualitative data concurrently but separately. The main sources for quantitative data are the numerical results of SABP over six years, while the sources for qualitative data are document analysis, archival data analysis and in-depth interviews. The two data sets are analysed separately and independently by employing quantitative and qualitative analytic procedures. After summarising the two sets of data analysis, the study discusses issues raised by both approaches, interpreting and merging the findings from them in order to produce a more complete understanding of the differences between spending ministries and the MOSF in the process of arriving at the SABP results.

1.5 The Structure of the Thesis

The organisation of this thesis can be seen in Figure 1-1. The following is a brief overview of the structure of the thesis. As already noted, the thesis seeks to examine the differences between spending ministries and the MOSF by using both quantitative and qualitative approaches. Thus, it starts by reviewing the relevant literature on the differences between spending ministries and the MOSF, based on the principal-agent relationship and focusing on the dysfunctional consequences of performance budgeting

(Chapter 2). At the same time, it also reviews studies of SABP results, focusing on factors which may affect differences in these results (Chapter 3). Drawing on these reviews of relevant literature, the thesis presents a methodological framework, including four key research questions, designs for both quantitative and qualitative approaches, and a data collection strategy (Chapter 4). It then explains the relationship between various factors and differences in SABP results through quantitative analysis, and explores differences between spending ministries and the MOSF by employing qualitative analysis (Chapters 5 and 6). Finally, the thesis discusses issues raised by both analyses and suggests further research (Chapters 7 and 8).

The following is a more detailed description of the eight chapters.

Chapter 1 introduces the background, research objectives and research questions, and the scope, methodology and structure of the research.

Chapter 2 gives definitions of performance and performance budgeting, and then extensively explores previous studies on dysfunctional consequences of performance budgeting based on the principal-agent theory, which is one of the main streams of new institutional economics, in order to develop a conceptual framework for exploring the differences between spending ministries and the MOSF through a qualitative approach.

Chapter 3 reviews trends in performance budgeting in many countries, including Korea, and then describes the SABP system in detail, focusing on factors which may affect differences in SABP results.

Chapter 4 presents a methodological framework for the research. It begins by suggesting four research questions and provides grounds for a mixed approach. It then explains both the quantitative and qualitative approaches in detail.

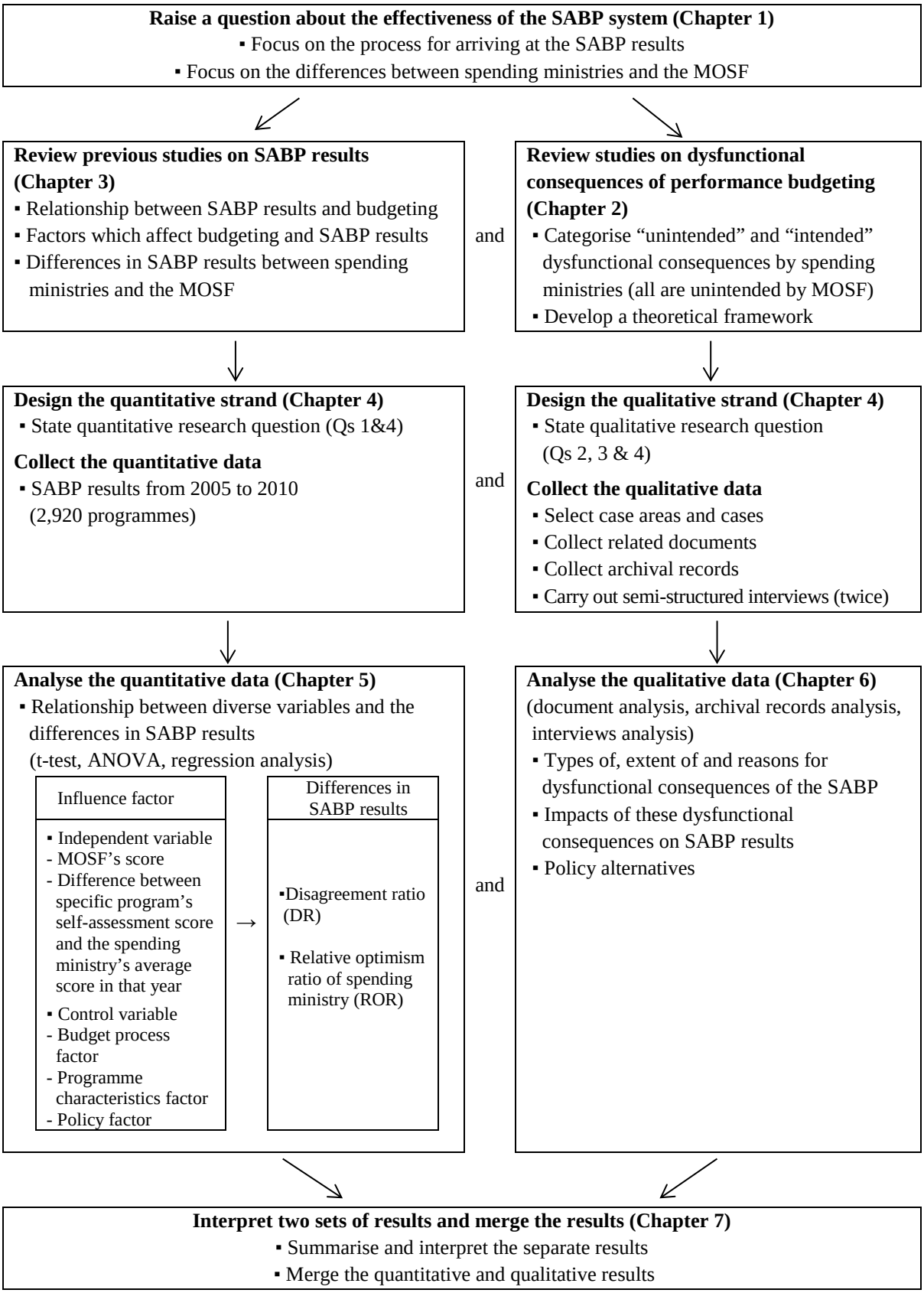
Chapter 5 analyses the relationship between diverse factors and differences in SABP results by employing various kinds of statistical methods such as correlation analysis, t-test, ANOVA and multiple regression analysis. The independent and control variables are drawn from the relevant literature reviews in Chapter 3.

Chapter 6 explores the types of, extent of and reasons for dysfunctional consequences of SABP; the impact of dysfunctional consequences on SABP results; and initial ideas for policy alternatives to produce a better SABP system through various kinds of qualitative data.

Chapter 7 discusses issues raised by both analyses and confirms the findings of these. It interprets and merges the findings from both analyses, and suggests feasible policy alternatives for minimising or preventing dysfunctional consequences. Also, it discusses the contribution of the thesis to three key areas.

Chapter 8 concludes the thesis by addressing the research questions and summarising the main findings of the study; and it considers the limitations of the research, while suggesting further research areas.

Figure 1-1 The Structure of the Thesis



Note: modified Creswell and Plano Clark (2011)

CHAPTER 2

THEORY OF PERFORMANCE BUDGETING

2.1 Introduction

It is not easy to define public sector performance as a unitary concept, because performance can have different meanings according to who is doing the defining, at what point measurement of performance takes place, and the priority given to various aspects of performance (Stewart and Walsh, 1994). For example, performance can be diversely described in the short-term, in the long-term, in a quantitative way, in a qualitative way, according to the point in time reached or the characteristics of performance being measured. Talbot (2005) outlines various dimensions of performance – e.g., accountability, user choice, customer service, efficiency, effectiveness and results, resource allocation, and public value. He also raises some problematic issues of performance, such as: the "unit of analysis problem", which relates to boundaries within the public domain; "conceptual problems", which relate to the agreed, operable definitions of performance, "technical problems", which are associated with measuring various aspects of performance; and "political and values problems" which relate to what is measured (Talbot, 2010). Bovaird (1996, p.147) indicates that performance "is not a unitary concept, within an unambiguous meaning. Rather, it must be viewed as a

set of information about achievements of varying significance to different stakeholders". Smith (1995, p.278) also mentions that, "The definition of performance in a public sector organisation is often elusive". These views give a useful insight into various aspects of performance, including process and results (Lee, 2007). For example, if the results of economic activities are focused on, then performance can be measured by cost-effectiveness, productivity and effectiveness; while if non-economic aspects are emphasised, then performance can be measured by satisfaction and service quality (Poister, 2003). However, this chapter will not focus on arguments about the concept of performance. Rather, it will apply the concept of "the three Es" (economy, efficiency, and effectiveness, based on the input into, process by, and output of organisations), which have been widely used in public sector performance management literature to explain performance (Carter et al., 1995; Flynn, 1997; Rouse, 1999; Diamond, 2005; Johnsen, 2005; Bouckaert and Halligan, 2008).

In addition, the definition of performance management in the public sector is often elusive and means different things at different administrative levels. However, many scholars agree that performance management is one of the managerial approaches that emphasises the autonomy and responsibility of organisations by measuring outcomes among employees and providing them with rewards or penalties based on their results, rather than controlling input or procedure (Joyce and Sieg, 2000; Epstein, 1992; Andrews and Hill, 2003). Smith (2002, p.105) also defines performance management in the National Health Service (NHS) in the UK as "a set of managerial instruments designed to secure optimal performance of the health care system over time, in line with policy objectives", suggesting that the three main aspects are: an emphasis on

managerial aspects, a longer time dimension for performance, and given policy objectives. These basic concepts of performance management have been adopted in the financial and budgetary areas: what is called performance budgeting in the public sector. Since the definition of performance budgeting also varies greatly, according to the scholars who have studied it, the process can reasonably be viewed by drawing common characteristics from various approaches to it.

With regard to managing performance in the public sector, previous studies suggest both a necessity for and limitations to performance budgeting. In particular, limitations can lead to unexpected dysfunctional consequences when it comes to measuring performance, and that is why many scholars have indicated these and tried to define and categorise them, drawing extensively on empirical evidence. However, these dysfunctional consequences of performance budgeting have not been considered seriously, or even regarded as exceptional phenomena, because many researchers have considered them as a sort of “cost”, and have decided that in the long-term the benefits of measuring performance can exceed these costs (Johnsen, 2005; Pidd, 2005). However, ignoring dysfunctional consequences or treating them as trivial matters can undermine the objectivity and reliability of performance budgeting and weaken the degree of acceptance by organisation members, leading to a failure in achieving the intentions and purposes of the introduction of performance budgeting in the long-term. Consequently, it is very important to understand the nature of dysfunctional consequences, and to analyse the reasons for them explicitly, in order to implement a performance budgeting system successfully (Shin, 2010). SABP in Korea, as a performance budgeting system, is exposed to various dysfunctional consequences not

only because spending ministries and the MOSF behave under different incentive structures but also because there is an asymmetry in the amount of information available to them about budgetary programmes. Also, these dysfunctional consequences can incur diverse types of transaction costs, such as target setting costs, programme implementation monitoring costs, performance measuring costs and performance confirming costs.

Against this background, this chapter firstly describes relevant previous studies related to the definition of performance and performance budgeting, elements and types of performance budgeting, and dysfunctional consequences of performance budgeting. Secondly, it presents the principal-agent theory as a theoretical lens through which to analyse the complex characteristics of dysfunctional consequences of SABP, focusing on the relationship between spending ministries and the MOSF. Thirdly, it develops a typology of dysfunctional consequences of SABP on the basis both of previous studies on dysfunctional consequences of performance budgeting and the principal-agent theory. Finally, it suggests a conceptual framework within which to explore dysfunctional consequences of SABP based on the study's own typology of dysfunctional consequences. The chapter does not focus on making an attempt to identify similar nuances of meaning that surround the definition of performance and performance budgeting. Instead, it focuses on drawing out general dysfunctional consequences of performance budgeting, and developing a useful typology to explore those of the SABP system, based on the principal-agent relationship between spending ministries and the MOSF.

2.2 Performance Budgeting in the Public Sector

2.2.1 Definition of Performance

This section describes the concept of performance, focusing on "the three Es", that is, economy, efficiency and effectiveness based on input into, process by, and output or outcome of organisations (Carter et al., 1995; Flynn, 1997; Rouse, 1999; Diamond, 2005; Johnsen, 2005; Bouckaert and Halligan, 2008, Boland and Fowler, 2000). An "input" refers to resources, such as personnel, finance and physical assets (Boland and Fowler, 2000). A "process" refers to activities or throughputs in which inputs are used to produce outputs. "Process" in the public sector means various activities undertaken to produce public goods and services. An "output" is the result of a process, and means the direct goods and services obtained by the process. An "outcome" is usually expressed in terms of final impacts on the society and beneficiaries resulting from the output. Based on a simple input, process and output model of organisations, "economy" can be defined as "the cost divided by the input (e.g., the cost per employee, the costs per office)" (Bovaird and Loffler, 2003, p.131; Diamond, 2005, p.5) and traditional budget systems put emphasis on inputs or resources, so the key to the performance of these systems is economy. Many measures of performance in the public sector can be defined on the basis of this economy or input-oriented perspective (Boland and Fowler, 2000). "Efficiency" can be defined as the ratio of input to output, and output-oriented budget systems concentrate on producing indicators of efficiency. Outputs can easily be measured in quantifiable terms: for example, number of crimes solved per police officer, patients treated per doctor, closures per inspection, and so on. It is in this way that

governments tend to measure organisational efficiency when it comes to producing operation results. Finally, "effectiveness" is related to the extent to which outputs achieve an organisation's goals and requirements, defined as "outcome divided by output", and outcome-oriented budget systems are concerned with producing indicators of effectiveness (Diamond, 2005, p.5). Outcome tells us more about the real success or failure of an organisation or a programme, and therefore it is much more difficult to measure or assess. According to the US Government Accountability Office (GAO) (2003, p.488), "high performing organisations have recognised that an effective performance management system can be a strategic tool to drive internal change and achieve desired results". The Office of Management and Budget in the US (OMB) (2003; re-cited GAO, 2004, p.21) also states that a performance budget is "a performance-oriented framework in which strategic goals are paired with related long-term performance goals (outcomes) and annual performance goals (mainly outputs)".

Although performance can be still conceptualised in various ways when presenting it in relation to tasks or activities undertaken by individuals or organisations, the focus of budgeting systems in the public sector has been changed from input to output or outcome; and therefore the focus of performance in the public sector tends to lie in the measurability of the efficiency and effectiveness of programmes (Miller et al., 2001; Bouckaert and Halligan, 2008). Consequently, performance in the public sector, especially in government activities, can be defined as the outputs, outcomes and impacts which can be achieved by the use of both tangible and intangible assets derived from the citizens' taxes in order to attain the intended and desirable purposes of government or government officials, and therefore this concept includes efficiency as well as

effectiveness. It is against this background that the next section will discuss the concept of performance budgeting in the public sector.

2.2.2 Definition of Performance Budgeting

Performance budgeting has been expressed in various terms, such as “performance-based budgeting”, “result-oriented management”, “programme budgeting”, “performance evaluation and management”, and “result-based management”. Defining performance budgeting as a unitary concept is not easy, because each government has its own approaches, definitions and methods for transforming performance data into budget decisions. Thus, the rest of this section describes diverse definitions of performance budgeting and draws out common characteristics from them in order to define performance budgeting in the public sector.

Shick (2007, p.123) states that “loosely defined, any system that provides information on the volume of outputs, the activities of government agencies, their workload, indicators of demand or need for public services, or the impact of expenditure qualifies as a performance budget. Strictly defined, only the budget systems which formally link increments in spending to increments in results would qualify”. Jordan and Hackbart (1999, p.69) describe performance budgeting, in a rather broad way, as “preparing the budget document with identified performance measures”. However, most academic researchers and international organisations define performance budgeting as a moderate version of the above definitions. Robinson and Brumby (2005, P.5) focus on the use of performance information, stating that it is “procedures or mechanisms intended to

strengthen links between the funds provided to public sector entities and their outcomes and/or outputs through the use of formal performance information in source allocation decision-making". Mackay (2007) indicates the importance of formal performance information, pointing out that performance budgeting can be understood as a budget system which pursues fiscal efficiency and the priority of public expenditures on the basis of performance information. Lauth (1985), Melkers and Willoughby (2001), and McGill (2001) focus on a process of linking the performance of programmes and activities with their budget allocations. The GAO (1999) defines performance budgeting as a system which links performance information and budgeting, and the OMB (2006) states that it is a budget system which explains clearly the relationship between performance goals and the cost of achieving targeted performance. The OECD (2005a, 2005b) states that it is a type of budgeting which links measurable results to budget allocation. The Ministry of Planning and Budget in Korea (2003) states that performance budgeting is an outcome-oriented budgeting system which measures the outputs and outcomes of public programmes and gives penalties or incentives on the basis of these. Finally, Park (2008) defines performance budgeting as a kind of financial reform which increases the links between performance management and budgeting in the public sector.

From these diverse definitions, we can see that the main characteristics of performance budgeting are the clarification of performance goals and objectives, the development of performance indicators and measurement techniques, and the feeding of results back into the process of implementing budgetary programmes (Pollitt, 2001). Consequently, performance budgeting can be defined as a kind of budgeting system which is intended

to improve the performance and increase the efficiency of budgetary programmes by clarifying their goals and objectives, developing performance indicators in line with them, measuring and analysing performance, and reporting and reflecting the results of analysis.

2.2.3 Elements of Performance Budgeting

Performance budgeting undertakes several stages in order to ascertain the level to which a performance target has been achieved by measuring the performance of a programme. The Department of Energy (DOE) in the US (1995) divides performance management and the budgeting process into 11 stages, such as identifying a member's understanding and analysing an organisation's environment, selecting core programmes in line with the organisation's mission and goals, clarifying the performance objectives of core programmes, defining performance indicators, determining the method of colleting performance data, measuring and analysing the performance, and reporting the results of analysis. The GAO (1996) simply categorises them into three steps: setting the objectives and desired outcome, measuring performance, and utilising performance indicators. Winstanley and Smith (1996) point out three core stages of performance budgeting: setting the goals and objectives, managing the performance, and measuring the performance. Similarly, Shin (2010) indicates four stages of performance budgeting: clarifying performance indicators and a desired outcome, implementing main programmes, collecting and analysing performance data, and evaluating and interpreting the analysis. Considering these previous studies and the characteristics of performance budgeting mentioned in Section 2.2.2, this study suggests three main stages of

performance budgeting: performance planning, performance measurement, and performance reporting.

Performance Planning

Performance planning includes clarifying the mission and goal for an activity, developing performance indicators, and setting the desired outcome. Of these, developing appropriate performance indicators is the core process of performance planning, because it is through measuring performance indicators that performance can be assessed. In general, a performance indicator is defined as a criterion against which the mission, strategic goal and performance goal of an organisation or a programme can be measured; and performance indicators can be divided into four types – input, process, output and outcome indicators – according to the progress of an activity (KIPF, 2006; Lee, 2003; Bang, 2009). Ammons (1995b) presents validity, reliability, timeliness, cost-efficiency and controllability as the appropriate characteristics of performance indicators. The Korea Institute of Public Finance (KIPF) (2012) also suggests five criteria for relevant performance indicators in SABP: they should be specific, measurable, attributable, reliable and timely (SMART). According to the GPA, performance indicators should be outcome-oriented as much as possible, in order to measure the ultimate impact of a programme, while output as well as process performance indicators must be used supportively when creating outcome performance indicators proves difficult (Art. 6). Taken together, these criteria indicate that developing appropriate performance indicators for a programme is very important for successful performance budgeting, because if they do not measure performance

objectively and accurately, performance budgeting cannot achieve its purposes effectively. Thus, key performance indicators should be made in accordance with the mission, and the strategic and performance goals of an organisation or programme.

Performance Measurement

Wang (2000, p.104) defines performance measurement as "the measuring of levels of activities and achievements through a range of indicators". Radnor and Barnes (2007, p.393) and Radnor (2008, p.317) define performance measurement more specifically as "a quantitative or qualitative value of the input, output, outcome or level of activity of an event or process", while differentiating it from performance management and performance reporting. According to Wholey and Newcomer (1997), performance measurement goes beyond the technical methodology of applying performance data and includes measuring all the processes of decision-making and management using performance data. Interestingly, Lebas (1995) indicates that performance measurement and management are not separable, although there are distinct differences between them. If this is the case, performance measurement can be used as a similar concept to performance management and budgeting. However, in general, performance measurement plays a key role in performance management and budgeting, as one of the core components of these processes (Lee, 2007); and performance management and budgeting include all processes of performance planning, performance measurement and performance reporting (Shin, 2010). Radnor and Barnes (2007) also state that performance measurement is related to efficiency, productivity and utilisation, while performance management is concerned with effectiveness and a more holistic view of

operations. Furthermore, differentiating the concept of performance measurement from that of performance management will provide more meaningful ways to analyse various dysfunctional consequences of performance budgeting systematically.

The purposes of performance measurement are described in different ways by different people, according to their different intentions. Behn (2003, pp. 593-598) points out eight managerial purposes of measuring performance, focusing on public managers: “to evaluate the outcomes, to control the inputs, to budget efficient measures, to motivate almost-real-time outputs, to promote easily understood aspects of performance about which citizens really care, to celebrate periodic and significant performance targets, to learn disaggregated data that can reveal deviances from the expected and to improve performance”. In order to achieve useful performance information, he indicates, public managers firstly need to consider carefully the purposes of measurement and then to select the best measures available for each purpose (Behn, 2003). Also, performance is measured from various aspects at different levels, so public managers need to consider seriously the differences in the purposes of, and requirements for, performance information at each level, e.g., that of a project or a team, a programme or department, or a strategy, in order to choose relevant performance measures (Osborne et al., 1995). The GAO (1996, pp.24-28) identifies the characteristics of appropriate performance measures at each organisational level, such as “demonstrating results, telling each organisational level how it is achieving its goals, restricting to the vital few performance indicators per goal, responding to multiple priorities and linking to offices that have the responsibility for making programs work”. Also, Wholey (1999) indicates that public

managers should balance the cost of data collection against the benefits to be drawn from useful performance information.

Consequently, performance measurement is not an end in itself but a means of performance budgeting (Osborne et al., 1995), and it is one of the key elements of performance management, as well as a useful tool for performance reporting, as will be seen in the next section.

Performance Reporting

Radnor and Barnes (2007, p.393) and Radnor (2008, p.317) define performance reporting as "providing an account, and often some analysis, of the level of input, activity, output or outcome of an event or process usually against some form of target". As the final stage of performance budgeting, not only the results of performance measurement but also the evaluation opinions arrived at in the process of measuring performance, such as the reason for not achieving a performance target level, are reported and processed publicly following the end of a fiscal year (GAO, 1996; Kong, 2008). In other words, performance reporting needs to provide specific information about the level achieved by a programme: whether this level attains the planned performance target level; to what extent performance has been improved; and how much it should be improved in the future (Kong, 2008). The GPA in Korea also deals with the open reporting of the results of performance measurement, the subject and method of performance reporting, and the feedback of performance results. Individuals and organisations, as well as programmes, can improve their performance by publishing

and utilising performance information openly. That is why performance reporting as the final component of performance management has an important meaning. The main contents of each stage of performance management are summarised in Table 2-1 below.

Table 2- 1 Elements of Performance Management and their Main Contents

Element		Main contents
Performance management	Performance planning	- Clarifying a vision, mission or goal - Developing performance indicators - Setting a desired outcome (performance target level)
	Performance measurement	- Collecting performance data - Measuring a quantitative or qualitative value
	Performance reporting	- Evaluating performance results - Providing an account of performance results - Feedback of performance results

2.2.4 Types of Performance Budgeting

Mackay (2007) divides performance budgeting into three types, according to the extent of the use of information provided by performance monitoring: direct performance budgeting, indirect performance budgeting and presentational performance budgeting. In direct performance budgeting, budget allocation for a programme is directly related to the performance results achieved, in order to eliminate political influence on budgets. In other words, budget allocation is determined by a formula or automatic process based on specific performance results or activity indicators. This type of performance budgeting has been applied to specific sectors in a small number of OECD countries, such as the education sectors of Denmark, Finland and Sweden (e.g., linking university funding to the number of students graduating), because direct performance budgeting is possible only if the outputs or outcomes of a programme are expressed accurately as a quantity (OECD, 2005c). With respect to indirect performance budgeting, performance

results are just one kind of information associated with budgeting for a programme, and other kinds of information, such as policy priorities, can also impact on budget allocation. Indirect performance budgeting is the most common type of performance budgeting in countries such as the UK, the US, Australia, New Zealand and Korea. Under presentational performance budgeting, monitoring and evaluation of information is related not to budget allocation but to reporting past or future performance information for the Congress or Parliament. Performance information is presented as background information to increase the accountability of spending ministries and their affiliated institutes (Jordan & Hackbart, 1999). Countries such as Denmark and Sweden, which have no formal or systematic budgeting system, have adopted this kind of performance budgeting. Similarly, Shah (2007) divides performance budgeting into four types, according to the criteria of how performance information is used in budget allocation: performance-reported budgeting, performance-informed budgeting, performance-based budgeting and performance-determined budgeting. In performance-reported budgeting, performance information is included as part of budget documentation, but it is not used for budget allocation. Performance-informed budgeting implies that performance information is used as a minor factor when it comes to budget decision-making. However, under performance-based budgeting performance results have a key role in budget allocation, along with many other factors. The UK, the US, New Zealand and Korea have adopted this kind of performance budgeting (Bang, 2009). In performance-determined budgeting, performance information is directly and explicitly linked to budget allocation for a programme. In a policy context, budget allocation for a programme is determined by various economic, political and social factors including performance results, so it is difficult to link performance information

directly to resource allocation (Ko, 2004). Consequently, according to the criteria of Mackay (2007) and Shah (2007), SABP in Korea can be regarded as indirect performance budgeting as well as performance-based budgeting (Bang, 2009; Jung, 2012).

2.2.5 Necessities and Limitations of Performance Budgeting in the Public Sector

Ferris and Graddy (1998) indicate that issues of public sector reform, including performance budgeting, have transcended national boundaries and have shed light on diverse policy areas for the last three decades; and they mention that the efforts made to reinvent government in the US (Osborne and Gaebler, 1992) and to introduce new public management (NPM) in the UK (Hood, 1991) have been the most representative of these initiatives. The reason why so many countries from the OECD to South America, Africa and the developing regions have introduced the performance budgeting system is its diverse and positive effects. Governments can increase the efficiency, effectiveness, accountability, transparency, and democracy of policies by applying performance information to strategic planning, organisational and personnel management, budgeting, service delivery improvement, and evaluation (Shah, 2007; Epstein, 1992; Wholey and Hatry, 1992; Wang, 2002; Ammons, 1995a; Mackay, 2007). First of all, performance budgeting can solve the problem of lack of incentives for increasing the efficiency and effectiveness of the public sector. The public sector generally offers fewer incentives for making an effort to enhance the efficiency and effectiveness of a programme than the private sector, because it is less exposed to competition and the risk of being wiped out. By evaluating the performance of a

programme and using it as one of the key factors in budget allocation, the public sector can increase the efficiency and effectiveness of a programme, its personnel, its organisation and its budget. Secondly, performance budgeting will lead to increasing accountability in the public sector, because it focuses on allocating not only more discretion but also more responsibility to a programme (Robinson and Brumby, 2005). By presenting performance targets explicitly and enabling public managers to focus on core tasks, it gives them more autonomy in their programmes while making them more responsible for performance. Thirdly, performance budgeting can improve the transparency of, and confidence in, the public sector by making it easier for people to understand its visions, missions, goals, strategies, programmes and activities. By reviewing performance results, citizens are able to identify more easily the core tasks of the public sector and the extent to which these have been achieved. Finally, performance budgeting can improve the democracy of policies by increasing opportunities for citizens' participation in public programmes (Mackay, 2007).

However, despite the importance of performance budgeting, as mentioned above, the approach is not a panacea. Indeed, performance budgeting may raise a new problem or worsen an existing one. It can be a big obstacle to the successful modernisation of the public sector if it is inappropriately set up and operated (Smith, 1995, 2005, 2008, 2011; Bouckaert and Peters, 2002). Performance budgeting in the public sector is based on the premise that performance objectives are quite obvious and can be measured objectively. However, studies indicate that in many cases performance targets in the public sector are unclear. Holmstrom and Milgrom (1991) and Baker (1992) raise the multi-tasking problem in the public sector when it comes to designing an economic incentive system.

By contrast with private companies, it is difficult for the public sector to have clear objectives, because a public organisation's value is determined by multiple and competing objectives (Heinrich and Marschke, 2010). In other words, in general, the public sector has to be responsible to various groups such as taxpayers, local government, other public organisations and ordinary citizens, so that in many cases the public sector's objectives cannot be clear and public programmes are complex by their very nature (Moore, 2002; Hood, 2006). Thus, policy makers who want to design economic incentive systems based on performance assessment in the public sector should consider carefully the multitasking problem. Dixit (1997, p.378) also describes this phenomenon using the phrase "common agencies with several principals".

In addition, although the objectives of public programmes are clear, it is difficult to measure and evaluate them accurately and objectively, due to their intangibility, indirectness, externalities and complexities (Wilson, 1989; Thiel and Leeuw, 2002). In many cases, there is no equivalent of a public service in the market, so it is hard to measure performance on the basis of the market or competition. Another reason is that it takes a long time for public programmes to impact on, or influence, society. Performance measurement carried out too late will mean that incentives or penalties based on performance results will lose their relevance; while measurement carried out too early makes it impossible to calculate performance meaningfully. Also, in cases where several programmes or organisations contribute simultaneously to producing a performance, the problem of how to measure the joint inputs and outputs can arise. Downs and Larkey (1986) state that evaluating performance in the public sector is a very difficult task because uncertain or uncontrollable external factors can affect it. In

such cases, poorly evaluated individuals or organisations can shift their responsibilities onto the external factors; and, furthermore, co-operation and communication among them can be undermined in the long run.

These limitations of performance measurement in the public sector can lead to various kinds of dysfunctional consequences for performance budgeting, and Smith (2005, p.1) indicates that these predicted and actual negative effects need to be monitored as "an intrinsic parts of the performance monitoring process". In this sense, various dysfunctional consequences need to be explained in detail.

2.3 Dysfunctional Consequences of Performance Budgeting

2.3.1 Previous Studies on Dysfunctional Consequences of Performance Budgeting

This section explores various dysfunctional consequences identified in previous studies based on the elements of performance budgeting mentioned in 2.2.3, and does so from the following angles. Firstly, the severity and extent of the impact of dysfunctional consequences may be different according to the elements of performance budgeting. For example, dysfunctional consequences in performance planning may produce the most negative impacts, because if there are dysfunctional consequences in performance planning, then the later stages cannot help but be exposed to various problems. Indeed, since the questions in SABP are logically linked with each other, dysfunctional consequences in performance planning are particularly likely lead to perverse effects in the later stages of assessment. Similarly, if there are some problems in performance

measurement, then performance reporting cannot produce meaningful information for people. Thus, by identifying dysfunctional consequences based on the elements of performance budgeting, policy makers can extract more useful performance information. Secondly, the possibility of identifying or controlling dysfunctional consequences may be different according to the elements of performance budgeting. In general, the Ministry of Finance and other related organisations, such as the National Assembly and the Board of Audit and Inspection, can more easily identify or control dysfunctional consequences in performance reporting than those in performance planning or measurement, because performance reports are published openly and can be evaluated objectively by several experts, whilst performance planning and measurement are more dependent on spending ministries' discretion. Thirdly, identifying dysfunctional consequences based on the elements of performance budgeting may produce useful insights that will contribute to my typology of dysfunctional consequences, which will be explained later. Many previous studies deal with similar phenomena using different terminologies, so it is difficult to work out the exact meaning of, and the relationship between, other studies. By categorising dysfunctional consequences under the same element of performance budgeting, diverse terminologies reflecting similar phenomena can be presented by one terminology.

Dysfunctional Consequences in Performance Planning

Dysfunctional consequences at this stage are mainly related to the development of performance indicators and the setting of performance target levels. Firstly, as for the development of performance indicators, four out of the eight types of unintended

consequences suggested by Smith (1995) are associated with this stage. Smith suggests the concepts of tunnel vision, sub-optimisation, myopia and ossification arising in the process of monitoring and investigating performance in the UK public sector. Tunnel vision occurs when managers choose the most easily quantifiable performance indicators and ignore the rest (Smith, 1995). This phenomenon is likely to happen in the public sector because public managers have a tendency to avoid the uncertainty and externality caused by the ambiguous and unclear objectives of public programmes when they assess their performance. Sub-optimisation occurs when public managers pursue narrow objectives, damaging the performance of the overall system (Smith, 1995). In particular, in many cases public organisations' final objectives and overall performance are made up of the joint outputs of several sub-divisions; so if they are not expressed in relation to these divisions' objectives or performance, then it is possible not to achieve the overall performance of the organisations. Myopia arises when public managers focus on short-term targets at the expense of longer-term objectives (Smith, 1995). In a recent study on the intended and unintended effects of the British health care sector's targets system for general practitioners (GPs) introduced in 2004, Smith (2008) states that if incentives are not carefully designed, GPs tend to focus on short-term objectives at the expense of long-term ones. In general, the results of public programmes can be shown after a long time, so performance indicators have some limitations when it comes to reflecting medium or long-term objectives well. Consequently, if public managers put excessive emphasis on this performance indicator, the long-term and strategic aspects of programmes may be ignored. Also, Smith (1995) refers to ossification to indicate that no one has the interest or capacity to revise or remove performance indicators which have already lost their purposes, due to ossification of the system. In particular, he

thinks that the public sector has not been faced with competition, in comparison with private companies, so it has less incentive to revise performance indicators at regular intervals. De Bruijn (2002) also mentions that performance measurement in the public sector may raise some negative effects as well as positive ones. He states that excessive emphasis on quantifiable performance will ignore the diversity and complexity of public programmes, and this can lead to the killing off of a professional attitude (Smith, 1993; Goddard et al., 2000). He also indicates that performance indicators focus on measuring quantities, that is clearly definable aspects of performance, and this leads to the blocking of professionalism. Similarly to Smith discussing ossification, he comments that measuring performance in the public sector can block innovations because the organisation faced with measuring performance tries to achieve its performance as efficiently as possible, not accepting the risk of exploring the unknown, which may reduce the organisation's output (de Bruijn, 2002). About 55 years ago, Ridgway (1956) pointed out that indiscriminate use of quantitative performance indicators relating to job performance regardless of single, multiple and composite indicators might lead to unanticipated side effects and reactions if there was no better understanding of organisational behaviours for public managers, because of the complexity of large organisations. He mentioned possible dysfunctional consequences of quantitative performance indicators based on organisational theory. For example, a US federal law enforcement agency's priority in the cases it investigated was the length of the case rather than its urgency (Blau, 1955; re-cited Ridgway, 1956); and contradictory managerial decisions might arise under multiple quantitative performance indicators such as profits and production.

Secondly, with respect to the setting of performance target levels, Weitzman (1980) verifies the ratchet effect on the basis of an economic theory: planners of an enterprise (e.g., individual workers, an intermediate sized department, or a giant sector) have a tendency to use a current performance target level as a criterion for setting future targets, especially under a repetitive contract relationship between principal and agents, in order to maximise the compensation. Smith (1995, p.297) refers to this phenomenon as “gaming”, which occurs when a manager deliberately under-achieves in order to secure a lower target in the next round of activity. More recently, Smith has also stated that quantitative performance targets need to have an appropriate basis which takes account of variations in key sources (Smith, 2005). Following Weitzeman’s conceptualisation of the ratchet effect based on an economic theory, Hood (2006) indicates that the ratchet effect is the most common dysfunctional consequence of managing public services by targets. Bevan and Hood (2006, p. 521) discuss three types of gaming problem that arose in the 60-year history of the Soviet target system – the ratchet effect, the threshold effect, and opportunistic output distortions – defining this gaming as “reactive subversion such as ‘hitting the target and missing the point’ or reducing performance where targets do not apply”. Of these three types of gaming, the ratchet effect and the threshold effect are closely related to the dysfunctional consequences of setting performance target levels. After referring to the ratchet effect as the tendency for public managers to set lower targets which they can easily exceed in preparation for the next year’s performance, and the threshold effect as a situation in which performance is aimed at a point near, but not reaching, the target, they discover the existence of these dysfunctional consequences in the UK’s NHS governance by

targets (star rating from 0 to 3) introduced in 2001 by the Department of Health in England.

Dysfunctional Consequences in Performance Measurement

Dysfunctional consequences in this stage are associated with collecting performance data and measuring the quantitative or qualitative value of a programme or activity. Smith (1995, p.290) indicates that when outcomes are difficult to measure, public managers have a natural tendency to focus on measurable outputs rather than the desired outcomes, and calls this phenomenon "measure fixation". For example, he states that the quality of work in an important area that is not covered by performance measurement, such as mental health, could fall under a new contract system of quality targets and incentives for GPs introduced in 2004 (Smith, 2008). Bevan and Hood (2006, p.521) point out similar phenomena as opportunistic output distortions, which are "attempts to achieve targets at the cost of significant but unmeasured aspects of performance". Hood (2006, p.517) points to evidence of "widespread storming which drafts in other medical staff and cancels other operations" in order to achieve measured performance indicators, in surveys by the British Medical Association. Kelman and Friedman (2009, p.917) conceptualise the dysfunctional consequences of performance measurement as "effort substitution (reducing effort on non-measurable performance dimensions)" and "gaming (making performance on the measurable dimension appear better, when in fact it is not)" when it comes to measuring performance. "Effort substitution" seems to be closely related to measure fixation, and it may easily happen in the public sector, which has diverse but simultaneous objectives (Holmstrom and

Milgrom, 1991), because public managers have a tendency to focus more on easily measured or achievable objectives. They state that "effort substitution" improves measured performance and causes non-measured performance to fall, whilst "gaming" does not improve any performance in terms of consumption of resources. They also indicate that although gaming includes direct data falsification or cheating, in many cases it contains indirect data falsification. Friedman and Kelman's gaming is differentiated from Smith's misrepresentation in the way that their concept of gaming is related to the manipulation of actual behaviour when measuring performance whilst misrepresentation, which will be explained later, deals with distortions in reported behaviour (Smith, 1995).

In addition, LeGrand and Bartlett (1993, pp.31-34) refer to the concept of "cherry picking", which is "the tendency of executive agents to discriminate against inefficient aspects of the policies to be implemented by providing services or goods only to those who make the least or least expensive use of them". De Bruijn (2002) describes this phenomenon as blocking an organisation's ambitions through optimising its inputs. Studies for the Job Training Partnership Act in the US showed that job training centre employees tried to select more likely trainees to increase their performance, and focused on methods which could increase the employment rate of trainees, not their long-term capability (Friedlander, 1998; Zornitsky et al., 1988). Bouckaert and Ulens (1998; re-cited Thiel and Leeuw, 2002, p.273) state that some Belgian welfare programmes dropped more challenging clients in favour of ones who were easier to deal with.

More broadly, Bouckaert and Balk (1991) suggest 13 “measurement diseases” of public productivity measurement under three headings: measurement assumptions; perceived numbers and volumes; and the content, position and amount of measures – e.g., measuring performance in the public sector is impossible due to the complex and collective nature of public services (impossibility disease), perceiving larger or smaller volumes than are in reality the case (convex or concave disease), mixing input activities with output quantities and effects (pollution disease), and taking measurements that are different from those management wants (mirage disease).

In addition to discussing these dysfunctional consequences, some studies focus on gaming based on the principal-agent relationship. In their studies for the US Federal Job Training Program, Courty and Marschke (2004, 2007) found that local job training centre mangers tried to adjust the timing of trainees’ graduation from the job centre in order to obtain more economic incentives, and that this behaviour impacted negatively on the true goals of the organisation. Also, they indicated the development of gaming relationships between the Federal Government and local managers, as local managers developed new strategies to counter the changes made by Federal Government in performance measurement methods. Similarly, in their study on economic incentives and their dynamics in public sector performance measurement, Heinrich and Marschke (2010) state that although some performance measurement seems to be very successful at some periods, agents adjust new performance measurement methods over time, and so gaming relationships between the principal and agents are possible. These studies imply that there is an information asymmetry between principal and agents in performance measurement systems in the public sector.

Dysfunctional Consequences in Performance Reporting

Dysfunctional consequences in this stage are associated with evaluating performance results and providing an account of, and feedback on, performance results. In many cases the production and delivery system for public services is very complex; so developing performance management systems for assessing exact and genuine performance is also very difficult. This is the reason why various circumstances surrounding organisations should be considered when it comes to assessing performance. However, although public managers have all the performance information related to the object they are measuring, they are sometimes likely to suggest false policy alternatives, due to the bounded rationality of interpreting performance data. Smith (1995, p.294) calls this phenomenon a “misinterpretation”. Moreover, managers sometimes deliberately manipulate performance data when reporting them, leading to a weak relationship between actual and reported performance. In particular, if public managers have wide discretion regarding the collection, recording and analysis of the reported performance, then they can be tempted to report advantageous data rather than disadvantageous ones, or even intentionally report distorted performance data. Smith (1995, p. 292) calls this a “misrepresentation”. He also finds that if the new GP contract introduced in 2004 is dependent on their self-reported performance data, there may be a great risk of misrepresentation (Smith, 2008). In addition, while mentioning performance paradox as a phenomenon of low correlation between performance indicators and performance itself (Meyer and Gupta, 1994), Thiel and Leeuw (2002) state that performance paradox is more closely related to performance reporting than to

performance itself. They refer to the cases of the Dutch police and the British National Health Service [7] to show underrepresentation and overrepresentation of actual performance respectively. Also, they mention that misrepresentation or misinterpretation is a deliberate performance paradox, and they differentiate this from an unintended performance paradox, such as tunnel vision, measure fixation, suboptimisation, myopia and cherry picking. Heinrich (2007) also empirically analyses dysfunctional consequences in performance reporting through the high performance bonus systems of the US Government under the Workforce Investment Act. She finds that performance bonus systems encourage misrepresentation of performance rather than actual performance improvement, and this implies that economic rewards are too often false rather than fitting performance improvement or efforts to achieve organisational goals. Poister (2003) mentions the possibility of non-comparability of performance data and indicates that comparing performance data drawn from different organisations and collected according to different criteria can lead to distortions in performance interpretation. In particular, misinterpretation can occur in the process of benchmarking using other organisations' performance. He also refers to underreporting and overreporting to explain the phenomenon in which actual performance is underrepresented or overrepresented. He argues that misrepresentation can sometimes occur because of problems of performance measurement, but in most cases it reflects the deliberate intention of individuals or organisations. Interestingly, Jones and Euske

[7] The case of the Dutch police (Wiebrens and Essers, 1999; re-cited Thiel and Leeuw, 2002) illustrates under-representation of actual performance. Although the percentage of crimes solved is decreasing, that of arrest, prosecution and the imposition of penalties is increasing, compared to the past. The former means that the police's performance is deteriorating. However, the latter indicates an improvement in performance. The case of the British National Health Service (Smith, 1995) is an example of overrepresentation of actual performance. Patients' waiting time for an operation seemed to have dropped to less than two years. However, the waiting time was calculated after the first hospital consultation, so hospitals postponed consultations in order to reduce the waiting time following them.

(1991, p.437) discuss the concept of "strategic misrepresentation" in public budgeting and identify 13 budget process factors which may encourage strategic misrepresentation. They define strategic misrepresentation in public budgeting as "the planned, systematic distortion or misstatement of fact – lying – in response to incentives in the budget process". The 13 budget process factors that may stimulate strategic misrepresentation include: "uncertainty regarding cause-and-effect relationships, information asymmetry, absolute constraints on revenues and expenditures, rigidity of control, excessive budget control, evidence that budgetary misrepresentation has worked" (Jones and Euske, 1991, pp. 443-454). Although their use of the term "misrepresentation" seems to be more related to the budgeting process than to the performance reporting itself, the nature of the misreporting or distortion, such as inflating budget requests, overestimating revenues or underestimating the expenditures of a programme, may be similar to other studies' explanations of misrepresentation. A summary of previous studies on dysfunctional consequences of performance budgeting in the public sector can be seen in Table 2-2 below.

The next section will describe the construction of a theoretical lens through which to explore these dysfunctional consequences more effectively.

Table 2- 2 Previous Studies on Dysfunctional Consequences of Performance Budgeting

Author	Research strategies	Findings	Dysfunctional consequences
[Dysfunctional consequences in performance planning]			
Smith (1995, 2005, 2008)	Documentary analysis	Identify unintended consequences in relation to publishing performance data and targets systems, taking examples from the UK health service sector	Tunnel vision Sub-optimisation Myopia Gaming (ratchet) Ossification
De Bruijn (2002)	Documentary analysis	Indicate various perverse effects of performance measurement in the public sector and suggest five strategies for preventing these * Blocking innovations, professionalism and systems responsibility are similar to ossification, tunnel vision and suboptimisation respectively.	Blocking innovations, Blocking professionalism, Killing systems responsibility
Ridgway (1956)	Documentary analysis	Indiscriminate use of single, multiple, and composite quantitative performance measures may lead to side effects and reactions from managers who overestimate the benefits	No balanced stress on objectives, Unanticipated consequences
Weitzman (1980)	Multiperiod stochastic optimisation model	Planners of enterprises have a tendency to use current performance as a criterion for setting future targets, especially under the dynamic incentive system	Ratchet effect
Bevan & Hood (2006)	Documentary analysis	Examine the reported successes and problems of measurement and gaming in the English public health service	Ratchet effect Threshold effect
Hood (2006)	Documentary analysis Interviews	Identify the existence of classic types of gaming in the target regime by employing both documentary analysis and interviews	Ratchet effect Threshold effect
[Dysfunctional consequences in performance measurement]			
Smith (1995, 2008)		-	Measure fixation
Bevan & Hood (2006), Hood(2006)		-	Output distortions
Kelman & Friedman (2009)	Documentary analysis Regression analysis	Conceptualise dysfunctional consequences such as effort substitution or gaming and find there were dramatic performance improvement in the England hospital trust	Effort substitution Gaming
LeGrand & Bartlett (1993)	Documentary analysis	Indicate executive agents' tendency to make organisations appear to be more successful than they actually are	Cherry picking (Cream skimming)
De Bruijn (2002)		-	Blocking ambitions

Bouckaert & Balk (1991)	Documentary analysis	Suggest 13 dysfunctional effects of productivity measurement in the public sector	13 diseases
Courty & Marschke (2004 / 2007)	Experiment Questionnaire Archival records / Documentary analysis Questionnaire	Training agencies time graduation dates to maximise their awards and this behaviour has a negative impact on the organisation. / Local training managers learn how to game and develop new strategies in response to the federal agency's new performance measurements	Gaming
Heinrich & Marschke (2010)	Documentary analysis	On the basis of the principal-agent model, indicate strategic behaviour of individuals over time in a dynamic framework for performance measurement and incentive systems	Strategic gaming
[Dysfunctional consequences in performance reporting]			
Smith (1995, 2008)		-	Misrepresentation Misinterpretation
Thiel & Leeuw (2002)	Documentary analysis	Performance paradox is more related to performance reporting than performance itself and can be divided into unintended and deliberate performance paradox	Deliberate paradox (Misrepresentation Misinterpretation)
Heinrich (2007)	Documentary analysis Regression analysis	High performance bonus systems do not motivate improvements in performance but encourage misrepresentation of performance	Misrepresentation
Poister (2003)	Documentary analysis	Non-comparability of performance data can lead to misinterpretation, and performance is under- or over-reported through the intention of individuals or organisation	Misinterpretation Under- or overreporting
Jones & Euske (1991)	Documentary analysis	Define the concept of strategic misrepresentation in public budgeting and identify 13 budget process factors which stimulate misrepresentation	Strategic misrepresentation

2.3.2 New Institutional Economics and Principal-Agent Theory

This section explains the theoretical usefulness of new institutional economics (NIE) for analysing the various dysfunctional consequences of performance budgeting mentioned above. In particular, it shows how the principal-agent theory, as one of the main streams

of NIE, can provide a useful theoretical basis for examining dysfunctional consequences in SABP, the Korean performance budgeting system.

Hood (1991) regards the NIE approach, including principal-agent theory and transaction cost analysis, as one of the major origins of NPM, which is the background for performance budgeting. In addition, Heinrich and Marschke (2010), Baker (1992) and Holmstrom and Milgrom (1991) indicate the usefulness of the "principal-agent theory" for analysing the appropriateness of linking performance management to incentive systems. They argue that it provides important insights for policy makers interested in the design and implementation of an economic incentive system when it comes not only to analysing how the characteristics of performance measurement systems affect agents' behaviours towards economic incentives but also to tackling a situation of information asymmetry. Ferris and Graddy (1998, p.226) also state that principal-agent theory, as one of the strands of NIE, provides a useful basis for "both a formal contractual relationships relating to the production of public services and a more implicit contract such as public budgeting and intergovernmental grants". NIE puts emphasis on the importance of context in relation to the behaviour of individuals and the interactions among them, and it makes it possible to explain budgetary reforms within that context (Hall and Taylor, 1996; Ha, 2002).

In the strict sense, it is possible to distinguish NIE from rational choice new institutionalism. However, this study makes no attempt to distinguish between the two approaches. Instead, it regards them as the same, because most arguments on

institutions are identical to each other[8] (Hall and Taylor, 1996; Kang, 2000; Kim, 1997). NIE starts with the basic assumptions of neo-classical economic theory, such as rationality, complete information, utility maximisation and complete competition, which are the basis of micro-economics (Harris et al., 1995). However, NIE modifies and extends neo-classical economics' rational individuals and perfect information assumptions to take in limited rationality and incomplete information (North, 1995). The starting point for this theory is "collective action dilemmas" due to the bounded rationality of individuals and imperfect information, which means that the combined rationalities of individuals can lead to irrational or sub-optimal results, as each individual seeks to pursue their own interest. Thus, institutions are created to solve the problem of collective action dilemma by providing stable rules for exchange (Ha, 2002). North (1992, p.4) states that "institutions are the rules of the game of a society or more formally are the humanly devised constraints that structure human interaction. They are composed of formal rules (statute law, common law, regulations), informal constraints (conventions, norms of behaviour, and self- imposed rules of behaviour); and the enforcement characteristics of both". Bates (1995) also argues that the core logic of new institutionalism is that rational individuals, who are faced with limited rationality, create institutions as a way of providing incentives or imposing new constraints, in order to overcome these limitations. Hall and Taylor (1996) indicate that NIE has explained distinctively how institutions originated, and they argue that the institution which

[8] In addition, Eggertsson (1990) divides the major components of neo-classical economics into Lakatos' (1974) hard core (rational choice, stable preferences and equilibrium) and a protective belt (full information, zero transaction costs, homogeneous goods and individual property rights/free competition). Then, he defines the perspective of denying hard core as "new institutional economics", contrasting it with "neo-institutional economics", which focuses on the protective belt while accepting the hard core. On the other hand, Lane (1993) treats neo-institutional economics as the same thing as Eggertsson's new institutional economics.This study too does not distinguish between new- and neo-institutional economics.

provides more benefits to its actors can better survive the process of competitive selection. Although it is very difficult to understand the diverse studies on NIE as a unitary framework, and NIE has been criticised for the way that it focuses too much on the perspective of efficiency, has too narrow a perspective, based on the economic and the formal side of institutions, and has no interest in the influence of unbalanced power relations and cultures on the formation of institutions (Lowndes, 1996), the logical framework and core components of NIE, such as bounded rationality, asymmetric information and subsequent opportunism, can provide very useful insights for exploring the dysfunctional consequences of SABP, focusing on the differences between spending ministries and the MOSF.

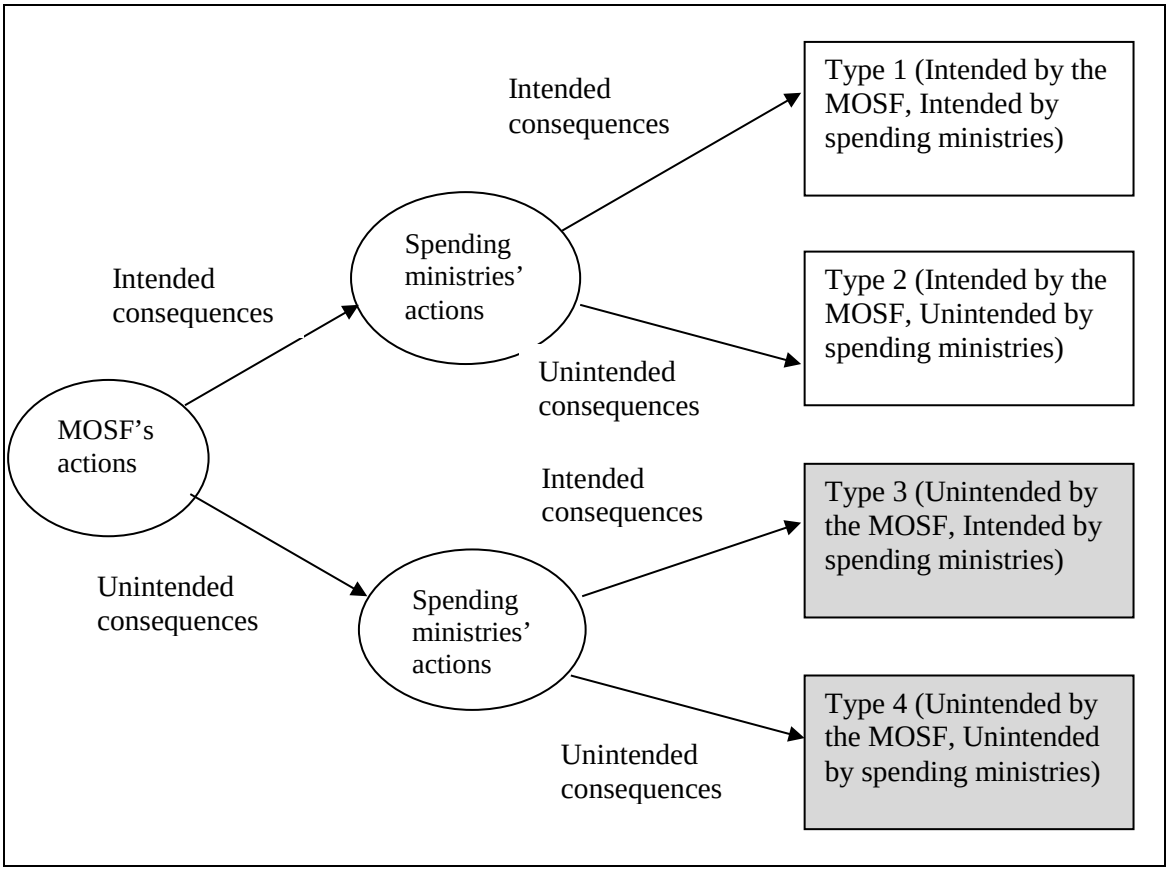
In particular, principal-agent theory is the main branch of NIE (Nabli and Nugent, 1989) and this theory provides a powerful tool with which to analyse the dysfunctional consequences of SABP which arise within the relationship between spending ministries and the MOSF. Principal-agent theory starts by analysing an asymmetric information situation, when one party does not have access to information about the other party's behaviour. The basic framework of this theory is that the principal employs the agent, who works for the benefit of the principal; but the principal does not have complete information about the agent's behaviour. As a result, there are differences of opinion as to desirable behaviour between the principal and the agent, and so agency problems can arise (Jensen and Meckling, 1976; Williamson, 1975, 1985; Lee, 1993). "Adverse selection" and "moral hazard" are identified in the context of insurance markets as two main agency problems relating to the information asymmetry of hidden action (Pauly, 1974). Adverse selection occurs when an insurance company does not know the

characteristics of the people they insure. That is, someone who is high-risk in terms of driving is more likely to take out driving insurance than a low-risk person, and the insurance company cannot obtain complete information about the person's driving habits due to information asymmetry (Pauly, 1974). While adverse selection is related to pre-contractual information asymmetry, moral hazard is associated with post-contractual information asymmetry. In the car insurance market, moral hazard arises when people who are insured have the ability or economic incentive to do things which affect insurance benefits (Pauly, 1968, 1974). For example, an insured driver has a higher tendency to be careless when driving than a non-insured one. The core of the principal-agent problems lies in the information asymmetry, and the agent tends to engage in opportunistic behaviour in order to increase benefits at the expense of the principal. Asymmetric information leads to various agency costs (Yoon and Lee, 1992). Firstly, the principal has the burden of monitoring the agent's behaviour (monitoring costs). Secondly, the agent also has to try to obtain the principal's confidence (bonding costs). Finally, there can be a minimal cost level which cannot be reduced anymore and is caused by the difference between the preferences of the principal and those of the agent (residual loss). This theory puts emphasis on how to devise an incentive system to control agents' hidden behaviour under conditions of information asymmetry.

The framework of NIE and the principal-agent theory, and their assumptions about incomplete information, bounded rationality, opportunism and self-interest, are suitable for use in analysing the dysfunctional consequences of SABP. To do this, let's start with the more general policy situations surrounding the MOSF and spending ministries, the two major actors in SABP. Simply put, the MOSF is the designer and controller of

SABP, and spending ministries are assessed through SABP; so the MOSF can be called "the principal", whilst the spending ministries can be seen as "the agents" in the principal-agent relationship. Based on this principal-agent relationship, the actions of the MOSF and spending ministries can be divided into the following four types, as shown in Figure 2-1: type 1 (intended by the MOSF, intended by spending ministries), type 2 (intended by the MOSF, unintended by spending ministries), type 3 (unintended by the MOSF, intended by spending ministries), type 4 (unintended by the MOSF, unintended by spending ministries).

Figure 2-1 Four Types of Consequences Based On the Principal-Agent Relationship



In the case of SABP, the consequences intended by the MOSF are to increase the efficiency and effectiveness of budget allocation and to improve spending ministries'

performance management capability. With these intended consequences of the MOSF, if spending ministries follow them well, then type 1 results can be expected. However, if spending ministries do not understand and do not follow the intentions of the MOSF, then type 2 results are possible. On the other hand, in the process of SABP various kinds of unintended consequences by the MOSF may arise, such as the dysfunctional consequences mentioned in previous studies. All these dysfunctional consequences can also be regarded as unintended side effects by the MOSF. In addition, some consequences regarded as unintended by the MOSF may be intended and deliberate on the part of the spending ministries, while others may be unintended by the spending ministries as well. The former case can lead to type 3 consequences, while the latter can give type 4 consequences. Among four types of consequences of actions by both spending ministries and the MOSF, this study focuses on type 3 and 4, which are unintended consequences by the MOSF, because analysing unintended side effects of actions by the MOSF can provide an opportunity not only to indicate the potential problems of SABP in a practical way but also to create a new academic theory. More specifically, the unintended consequences of a policy have their own contexts or logics which have led to these consequences, so a new policy devised without careful understanding or consideration of these may make the problems more complex or the effort useless when the policy is implemented. That is, SABP cannot be a panacea without cautious consideration of potential unintended side effects by the MOSF. In addition, a study of intended consequences can only justify or confirm existing theories, whilst a study of unintended ones can suggest a possible new theory which cannot be drawn from the existing theories (Lee, 2009). In other words, a study of side effects not intended by the MOSF may lead a researcher to disprove existing theories, rather than

to verify them, thus giving the researcher a chance to create a new hypothesis or theory (Magee, 1985). In fact, many previous studies, such as those by Park (2005, 2008), Bang (2009) and Cho (2010), have focused on analysing consequences intended by the MOSF (type 1 and 2), which will be explained in detail in Chapter 3, while there have been almost no studies on consequences unintended by the MOSF (type 3 and 4). That is why the present study pays attention to consequences unintended by the MOSF, that is, types 3 and 4. Furthermore, it notes that NIE and the principal-agent theory can provide a useful insight for analysing type 3 and 4 consequences for the following reasons. Firstly, as mentioned earlier, NIE focuses on institutional factors and incentive structures for interested groups. Institutions affect the incentive system, which has an influence on actors' behaviour; and actors seek to maximise their interests within the constraints of institutions. Thus, NIE can provide fundamental explanations for the unintended consequences of SABP, considering the behaviours of both spending ministries and the MOSF in relation to an economic incentive system. Secondly, the principal-agent theory is suitable for explaining spending ministries' opportunistic behaviour within an asymmetric information situation, and so it can give a powerful insight into dysfunctional consequences, especially type 3 consequences, of activities by the MOSF and spending ministries. Moreover, these theories also suggest that policy makers need to consider unintended side effects seriously when designing and refining performance budgeting systems. The next section will specify the type 3 and 4 consequences of SABP in detail.

2.3.3 Development of a Typology of Dysfunctional Consequences of SABP

There are many previous studies focusing on the conceptualisation, categorisation or development of a typology of dysfunctional consequences of performance management and budgeting in the public sector. Smith (1990) indicates that the use of performance indicators in the public sector has not enhanced accountability as was expected, and he raises some possible issues and problems from the perspective of the principal-agent relationship. He also states that agents have continuous incentives to suggest modest performance target levels because the principal cannot help but set agents' target levels on the basis of previous performance, and that there are difficulties in measuring efficiency and effectiveness by using performance indicators due to the diverse aspects of public services. On the basis of these findings, Smith (1995) identifies and categorises eight unintended dysfunctional consequences of publishing performance data in the public sector, such as tunnel vision, suboptimisation, myopia, measure fixation, misrepresentation, misinterpretation, gaming and ossification, all of which have been mentioned earlier. He indicates that the reason common to all eight dysfunctional consequences is "the lack of congruence between the goals of the agent and the actual goals of the principal" (Smith, 1995, p.283). In particular, "the divergence between organisational objectives and the measurement scheme" can produce the first three phenomena; "an inability to measure complex phenomena with precision or fidelity" is the reason for the next two dysfunctional consequences; "an inability to process performance data correctly" produces the sixth and seventh phenomena; and "an inability to respond to new circumstances" leads to ossification (Smith, 1995, p.283).

In addition, Pidd (2005) tries to categorise Smith's eight unintended dysfunctionalities based on the concept of "grid and group". He mentions that "grid indicates the degree to which people's action are governed by externally imposed rules and conventions, and group indicates the degree to which actions are governed by group choice, that is the social collectivity to which someone is committed" (Pidd, 2005, p.490). Based on these two criteria, he suggests four types of people: hierarchists (high grid and high group type), individualists (low grid and low group type), fatalists (high grid and low group type) and egalitarians (low grid and high group type). He states that this typology will be helpful to understand the reason why dysfunctional consequences may occur. For example, if a fatalist is treated as a hierarchist, then people tend to focus on the easiest performance measures, and this leads to tunnel vision, suboptimisation, myopia, measure fixation and even misrepresentation undertaken to produce a good impression.

De Bruijn (2002, pp.581-583) suggests seven negative effects of performance measurement in the public sector, referring to the work of Smith (1995), such as "prompting game playing, adding to internal bureaucracy, blocking innovations, blocking ambitions, killing professionalism, killing system responsibility and punishing good performance". As mentioned earlier, among these seven negative effects, blocking innovations is closely related to ossification and blocking professionalism is related to the tunnel vision of Smith's work (1995). Also, killing system responsibility seems to be related to suboptimisation because this effect refers to less willingness to share best practices among organisations.

Hood (2006, p.519) suggests that target gaming is based on two axes: one is performance data, "creatively interpreted/contrived/spun or invented/dropped/not provided"; and the other is whether "the underlying performance in provision of service alters or remains unchanged". Within the table he divides gaming into being acceptable, generally seen as an unacceptable, seen as a fact of life in bureaucratic politics, or as generally seen as cheating. Hood (2007) also categorises dysfunctional consequences according to three types of performance management system (target systems, ranking systems and intelligence systems)[9], identifying ratchet and threshold effects in target systems, indeterminacy and volatility in ranking systems, and output distortion in both systems.

Radnor (2008, pp.324-325) suggests four types of organisational gaming – muddled, massaging, manoeuvring and manipulated – based on the ideas of Hood (2006) and others who have considered the extent, range and result of the impacts of gaming. She employs two axes: "level of gaming within the organisation" and "impact of gaming". If the level of gaming is low, it just changes the output data; but if the level becomes high, the actual activities may be changed in order to achieve the target. The impact of gaming is related to whether the gaming happens within a department or organisation, or whether it affects citizens, consumers, customers or patients. She describes a situation where the level of gaming is low and has an effect only within a department as "muddled"; while if the level of gaming is low but affects citizens or consumers, she

[9] Target systems measure how much actual performance has been achieved during the specified period. Taylor's scientific management and the former Soviet Unions' target system are good examples of this type. Under ranking systems, the ranking of individuals or organisations is determined by the results of comparing performance against that of others. Intelligence ranking systems comprehensively consider various kinds of background information, such as credit ranking, health care records, even rumour, as well as performance (Hood, 2007).

describes it as “massaging”. “Manoeuvring” and “manipulated” are labels for when the level of gaming is high, so that the gaming changes the actual activities in order to achieve the target.

Thiel and Leeuw (2002) divide performance paradox into unintended performance paradox and deliberate performance paradox, mainly referring to Smith’s work (1995). They regard tunnel vision, measure fixation, myopia, suboptimisation and cherry picking as unintended performance paradoxes, and these paradoxes can occur as a result of the minimal accountability requirements of the public sector, the elusiveness of policy objectives, the non-quantifiability and non-measurability of policy goals, and an excessive emphasis on monitoring and efficiency within the public sector. In addition, they argue that public sector organisations can attempt to “sabotage an audit” and “hide ill performance by misrepresenting and misinterpreting performance indicators”, and they call these phenomena a “deliberate performance paradox”, contaminating the relationship between reported and actual performance (Thiel and Leeuw, 2002, p.274). Moreover, interestingly, they state that, “a deliberate performance paradox can occur only if the conditions for an unintended performance paradox are present as well.” They consider that the paradox is related not to performance itself but to reports on performance, and they seem to consider a deliberate paradox as an extension of an unintended paradox.

Keum and Lee (2009) divide the dysfunctional effects of performance management systems into unintended and intended ones, while investigating whether those effects appear in the work of the central government of South Korea as it implements

performance management. They suggest tunnel vision, myopia, reinforcing internal bureaucracy/blocking innovations and blocking cooperation/irrational compensation systems as unintended dysfunctional effects, while measure fixation, suboptimisation, misrepresentation/misinterpretation, gaming (includes cherry picking and ratchet effect) are the intended ones. But they do not explain the reason why they assign some dysfunctional consequences to the unintended category, while others are assigned to the intended one. Table 2-3 summarises a typology of dysfunctional consequences of performance budgeting in the public sector suggested by previous studies.

Table 2- 3 A Typology of Dysfunctional Consequences Suggested by Previous Studies

Author	Criterion or Reason		Typology
Smith (1995)	Lack of congruence between the goals of the agent and the goals of the principal		Eight dysfunctional consequences
	Divergence between organisational objectives and the measurement scheme		Tunnel vision, suboptimisation, myopia
	Inability to measure complex phenomena		Measure fixation, misrepresentation
	Inability to process performance data correctly		Misinterpretation, gaming
	Inability to respond to new circumstances		Ossification
Pidd (2005)	High grid and high group		Hierarchist
	High grid and low group		Fatalist
	Low grid and high group		Egalitarian
	Low grid and low group		Individualist
De Bruijn (2002)	There is no specific criterion. Prompting game playing, adding to internal bureaucracy, blocking innovations, blocking ambitions, killing professionalism, killing system responsibility, and punishing good performance are all mentioned.		
Hood (2006)	Creatively interpreted, contrived or spun (performance data)	Alters (underlying performance)	Acceptable
		Remains unchanged	Generally seen as an unacceptable form of gaming
	Invented, dropped or not provided	Alters	Fact of life in bureaucratic politics
		Remains unchanged	Generally seen as a cheating
Hood (2007)	Target systems		Ratchet and threshold effects
	Ranking systems		Indeterminacy and volatility
	Both systems		Output distortion
Radnor (2008)	Level of gaming within the organisation & Impact of gaming		Muddled, massaging, manoeuvring, manipulated

Thiel &Leeuw (2002)	Minimal accountability requirement, elusiveness of policy objectives, non-quantifiability and non-measurability of goals, excessive monitoring and efficiency	Unintended performance paradox (tunnel vision, measure fixation, myopia, suboptimisation, cherry-picking)
	Sabotage an audit, hide poor performance + condition of unintended performance paradox	Deliberate performance paradox (misrepresentation,misinterpretation)
Keum & Lee (2009)	There is no specific criterion or reason for unintended or intended dysfunctional effects.	Unintended dysfunctional effects (tunnel vision, myopia, reinforcing internal bureaucracy, blocking innovations, blocking cooperation, irrational compensation system)
		Intended dysfunctional effects (measure fixation, suboptimisation, misrepresentation, misinterpretation, gaming (includes cherry-picking, ratchet effect))

From the previous studies attempting to make a typology of dysfunctional consequences of performance budgeting, I have drawn three useful implications for my study. First of all, the studies on dysfunctional consequences mentioned in Section 2.3.1 can give a meaningful starting point for developing my own typology of dysfunctional consequences in order to explore the differing perspectives of spending ministries and the MOSF, because most of the works mentioned in Section 2.3.1 also started to make a typology on the basis of previous findings on the dysfunctional consequences of performance management systems. For example, the works of Pidd (2005), de Bruijn (2002), Thiel and Leeuw (2002) and Keum and Lee (2009) develop a typology while referring to Smith's (1995) or other previous findings. Radnor (2008) also suggests a typology mainly based on the work of Smith (1995) and Hood (2006). Consequently, various dysfunctional consequences related to the elements of performance budgeting mentioned in Section 2.3.1 could be the basis of my work.

Secondly, considering the overlapping of various terminologies relating to dysfunctional consequences, the comprehensiveness of interesting ones, the importance

of analysis, and the possibility of application to SABP, I decided that the potential dysfunctional consequences of SABP included the following seven phenomena: tunnel vision, myopic management, suboptimisation, measure fixation, ratchet effect, misrepresentation and cherry picking. Researchers explore or explain the same phenomena by using different terminology, so the meaning of a term sometimes overlaps that of another term, or is divided among different dysfunctional consequences, and as a result it was very difficult to find clear criteria for developing a typology of dysfunctional consequences. For instance, regarding the concept of gaming, Smith (1995) indicates Weitzman's ratchet effect (1980), but Kelman and Friedman (2009) refer to it as an active manipulation of performance in measuring performance, and Courty and Marschke (2004, 2007) and Heinrich and Marschke (2010) identify it as the strategic behaviours of agents in response to a new performance measurement introduced by the principal. With respect to Smith's measure fixation, Bevan and Hood (2006) and Hood (2006) explain similar phenomenon as "output distortions" and Kelman and Friedman (2009) as "effort substitution". De Bruijn (2002) also uses different terminologies to indicate what Smith (1995) calls ossification and tunnel vision, and LeGrand and Bartlett call (1993) cherry picking phenomenon. He describes these phenomena as blocking innovations, killing professionalism and blocking ambitions respectively. Consequently, in order to analyse the complex phenomena surrounding SABP, I needed to draw out some dysfunctional consequences suitable for applying to SABP, and then define the interesting dysfunctional consequences on the basis of previous studies. According to Smith (1995, pp.284-290), Bevan and Hood (2006, p.521) and LeGrand and Bartlett (1993, p.31-34), my suggested seven dysfunctional consequences of SABP can be defined under the following headings:

- *Tunnel vision*: an emphasis on phenomena that are quantified in the performance measurement scheme at the expense of unquantifiable aspects of performance

- *Myopic management*: concentration on short-term issues, to the exclusion of long-term criteria that may only show up in performance measurements in many years' time

- *Measure fixation*: an emphasis on measures of success rather than the underlying objectives

- *Suboptimisation*: pursuit of narrow local objectives by managers, at the expense of the objectives of the organisation as a whole.

- *Ratchet effect*: a tendency for managers to set lower targets, which they can easily exceed, in preparation for the following year's performance measurement.

- *Misrepresentation*: a deliberate manipulation of data, including "creative" accounting and fraud, so that reported behaviour differs from actual behaviour.

- *Cherry picking*: a tendency by executive agents to counter inefficient aspects of the policies to be implemented by providing services or goods only to those who make the least or least expensive use of them.

Thirdly, this study divides the seven dysfunctional consequences into unintended dysfunctional consequences and intended dysfunctional consequences by spending ministries, whilst all dysfunctional consequences are unintended by the MOSF. That is, some dysfunctional consequences are unintended by both spending ministries and the

MOSF, while others are unintended by the designers of the SABP system but are likely to be deliberate responses by spending ministries, as agents in the principal-agent relationship. My suggested typology of dysfunctional consequences of SABP is shown in Table 2-4.

Table 2- 4 Types of Dysfunctional Consequences of SABP Exhibited by Spending Ministries

Unintended dysfunctional consequences	Intended dysfunctional consequences
- Tunnel vision - Myopic management - Measure fixation - Suboptimisation	- Ratchet effect - Misrepresentation - Cherry picking

* N.B. All dysfunctional consequences are “unintended” by the MOSF.

The first reason for this typology is that unintended dysfunctional consequences are more closely related to the intrinsic characteristics of performance management in the public sector, such as the difficulty of developing productivity indicators, measurements of quality, and long-term and whole performance indicators, whilst intended dysfunctional consequences include a more active and deliberate intention of gaming by a spending ministry as well as the conditions for unintended dysfunctional consequences (Thiel and Leeuw, 2002; Shin, 2012). For example, many previous studies have pointed out the difficulty of measuring and evaluating the performance of public services objectively and accurately, due to the complexity, multitasking problem, intangibility, indirectness and externality involved (Moore, 2002; Hood, 2006; Dixit, 1997; Downs and Larkey, 1986; Wilson, 1989; Thiel and Leeuw, 2002). These intrinsic characteristics of measuring performance in the public sector can lead to a tendency to focus on the quantifiable, short-term, measurable and narrow performance of a programme in order to achieve the target more easily. In other words, spending

ministries may have an intention to achieve higher performance targets more easily, but it can be said that the tendency is strongly influenced by these intrinsic limitations of performance budgeting. On the other hand, setting performance target levels as low as possible, manipulating performance data in performance reporting, and selecting more favourable groups when implementing a programme are more strongly influenced by spending ministries' deliberate and active intentions to achieve high performance targets, although they are also affected by the intrinsic characteristics of performance budgeting in the public sector. In this sense, the former category includes tunnel vision, myopic management, measure fixation and suboptimisation while the latter contains the ratchet effect, misrepresentation and cherry picking. Although Thiel and Leeuw (2002) use similar terminology (e.g., unintended performance paradox and deliberate performance paradox) and Keum and Lee (2009) use the same terms, they do not suggest a clear criterion for their typologies as this study does. In addition, Thiel and Leeuw (2002) do not even mention the ratchet effect, one of the most probable dysfunctional consequences of performance budgeting.

The second reason is related to policy alternatives for a better SABP system. As unintended dysfunctional consequences are more closely linked with the intrinsic limitations of performance budgeting in the public sector, it might be very difficult to abolish them completely. However, in the case of intended dysfunctional consequences, which are more closely associated with spending ministries' deliberate actions, the MOSF and other related organisations, such as the National Assembly and the Board of Audit and Inspection of Korea (BAI), are likely to identify them more easily, and consequently it might be easier to change them than to change unintended dysfunctional

consequences. It may also mean that the focus of policy alternatives for preventing or minimising the two types of dysfunctional consequences need to be different from each other. Policy makers need to focus more on fundamental approaches, such as educating and increasing the expertise of spending ministries' officials, while understanding the characteristics of performance budgeting in the public sector, if they are to reduce unintended dysfunctional consequences. As for intended dysfunctional consequences, less fundamental approaches, such as stronger audits of performance data, may be necessary. Various kinds of policy alternatives will be discussed in detail in Chapter 7.

The study's typology of dysfunctional consequences is distinguished from those of previous studies in the way that one specific dysfunctional consequence can be categorised as "unintended" and "intended" at the same time, based on the principal-agent relationship between spending ministries and the MOSF. In other words, the ratchet effect, misrepresentation and cherry picking can be regarded as intended by spending ministries while at the same time they are unintended by the MOSF. Although many previous studies, such as those of Weitzman (1980), Smith (1995), Courty and Marschke (2004, 2007) and Heinrich and Marschke (2010), analyse dysfunctional consequences within the framework of a principal-agent relationship, they do not simultaneously divide specific dysfunctional consequences into unintended and intended ones based on this relationship. Also, Pidd (2005) and Radnor (2008) do not categorise dysfunctional consequences on the basis of the principal-agent relationship. Rather, Pidd (2005) focuses on the incongruence between the characteristics of individuals who are working under a performance management system and the way in which that system is being directed by the organisation, and he argues that this

incongruence can lead to various kinds of dysfunctional consequences, as mentioned by Smith. Radnor (2008) tries to categorise the types of organisational gaming based on the strength of gaming and the extent of its impact, not seeking specific types of dysfunctional consequences of performance budgeting.

The next section will develop a conceptual framework to analyse dysfunctional consequences of SABP based on the study's suggested typology and the principal-agent theory.

2.3.4 Conceptual Framework to Analyse Dysfunctional Consequences of SABP

The Nature of SABP

The Korean government introduced SABP with the intention of increasing the efficiency and effectiveness of budget allocation. In particular, the MOSF has tried to achieve these aims by offering some incentives as well as penalties based on SABP results. These incentives and penalties can affect the behaviours of spending ministries and their officials by working as economic incentives as well as constraints. In this sense, SABP can be regarded as a type of "implicit contract" between spending ministries and the MOSF, a contract that operates on the basis of performance targets and economic incentives. In addition to a contractual relationship between spending ministries and the MOSF, a hierarchical control system is at work, because the MOSF, the designer of the SABP system, develops and distributes an SABP manual in order to review the self-assessment results of spending ministries (Spiller, 2008; Lee, 1998). The

existence of both an implicit contractual relationship and a hierarchical relationship between spending ministries and the MOSF can lead to some problems of information asymmetry. In reality, the MOSF and spending ministries basically exist in an asymmetric information situation with regard to a programme's performance. The MOSF tries to increase the efficiency and effectiveness of budget allocation by reflecting the exact performance results of a programme, but the MOSF has less knowledge of, and capacity to use, the performance indicators or performance targets of a programme than spending ministries. Thus, the MOSF hands over discretion on self-assessment of a programme's performance to spending ministries, although the final verdict belongs to the MOSF through its review of spending ministries' self-assessment results. During this process the MOSF and spending ministries can have conflicting interests: that is, the MOSF focuses on overall improvement of budget efficiency and effectiveness based on fair SABP results, while spending ministries pay more attention to simply increasing their programmes' budgets through high SABP results. This kind of information asymmetry can lead to various agency costs, such as the cost of developing appropriate performance indicators and targets, the cost of monitoring programme implementation, the cost of performance measurement and the cost of performance information checks and confirmation. The fact that the MOSF has to pay significant agency costs to review spending ministries self-assessment results, given the situation of information asymmetry, is likely to increase the possibility of moral hazards for spending ministries. High self-assessment SABP scores by spending ministries can be a good example of these moral hazards. Moreover, if moral hazards for spending ministries persist for many years, then the MOSF's agency costs will be incurred continuously and increasingly. These can surely be regarded as dysfunctional

consequences of SABP. Consequently, SABP is a kind of "context" and "institution" that affects the behaviours and decision-making activities of spending ministries and their officials, containing the various potential dysfunctional effects mentioned above at the same time (Hall and Taylor, 1996). In the following components of the conceptual framework the nature of SABP will be explored in detail.

Actors in the SABP System and their Behavioural Assumptions

The main actors participating in the SABP system are the MOSF and 48 spending ministries (including agencies). From the perspective of operating SABP, the MOSF plays a significant role in drawing up SABP guidelines, training spending ministry officials, reviewing spending ministries' self-assessment results, and controlling the SABP system. In addition, the MOSF tends to be risk neutral and tries to maximise the value of SABP results by recommending performance indicators that reflect spending ministries' efforts accurately. Consequently, the MOSF can be said, from the perspective of principal-agent theory, to have the characteristics of a principal. Similarly, spending ministries can be viewed as agents, because they tend to be risk-averse and to try and maximise their compensation whilst minimising their efforts. Moreover, two behavioural assumptions about stakeholders can be developed from traditional transaction cost economics (Williamson, 1985, 1990). This study adopted the idea that the MOSF operates SABP under a condition of bounded rationality, which has been defined as "a semi-strong form of rationality in which economic actors are assumed to be intendedly rational, but only limitedly so" (Williamson, 1985, p.45). Similarly, as mentioned earlier, North (1995) and Bates (1995) have also indicated that rational

individuals are faced with limited rationality and incomplete information. The bounded rationality of the MOSF means that it cannot pursue substantial rationality (Simon, 1978). To be more precise, the distinct characteristic of the bounded rationality of the MOSF is that the Ministry is not able to obtain complete performance information on a programme compared with the spending ministries, because it does not directly implement the programme. The other behavioural assumption is the self-interested opportunism of the spending ministries. Opportunism is defined as "self-interest seeking with guile" (Williamson, 1985, pp. 47-8). Williamson (ibid.) also expands this definition: "Opportunism refers to the incomplete or distorted disclosure of information, especially to calculated efforts to mislead, distort, disguise, obfuscate or otherwise confuse as well as more blatant forms of lying, stealing and cheating". Spending ministries' opportunistic behaviours can be activated by an information asymmetry between the MOSF and themselves, because in general the MOSF has limited access to information about a budgetary programme, whilst spending ministries have a lot of information about their programmes.

Strategies and the Results of Interactions between Strategies

The Korean government, through the MOSF, introduced SABP to measure and evaluate the results of budget spending, because the public sector is less exposed to the market and competition than private companies, and therefore has less incentive for increasing efficiency and effectiveness through cost saving (Robinson and Brumby, 2005). At the same time, SABP focuses the accountability and responsibility of the spending ministries. Given these reasons for introducing SABP and its core objectives, we can

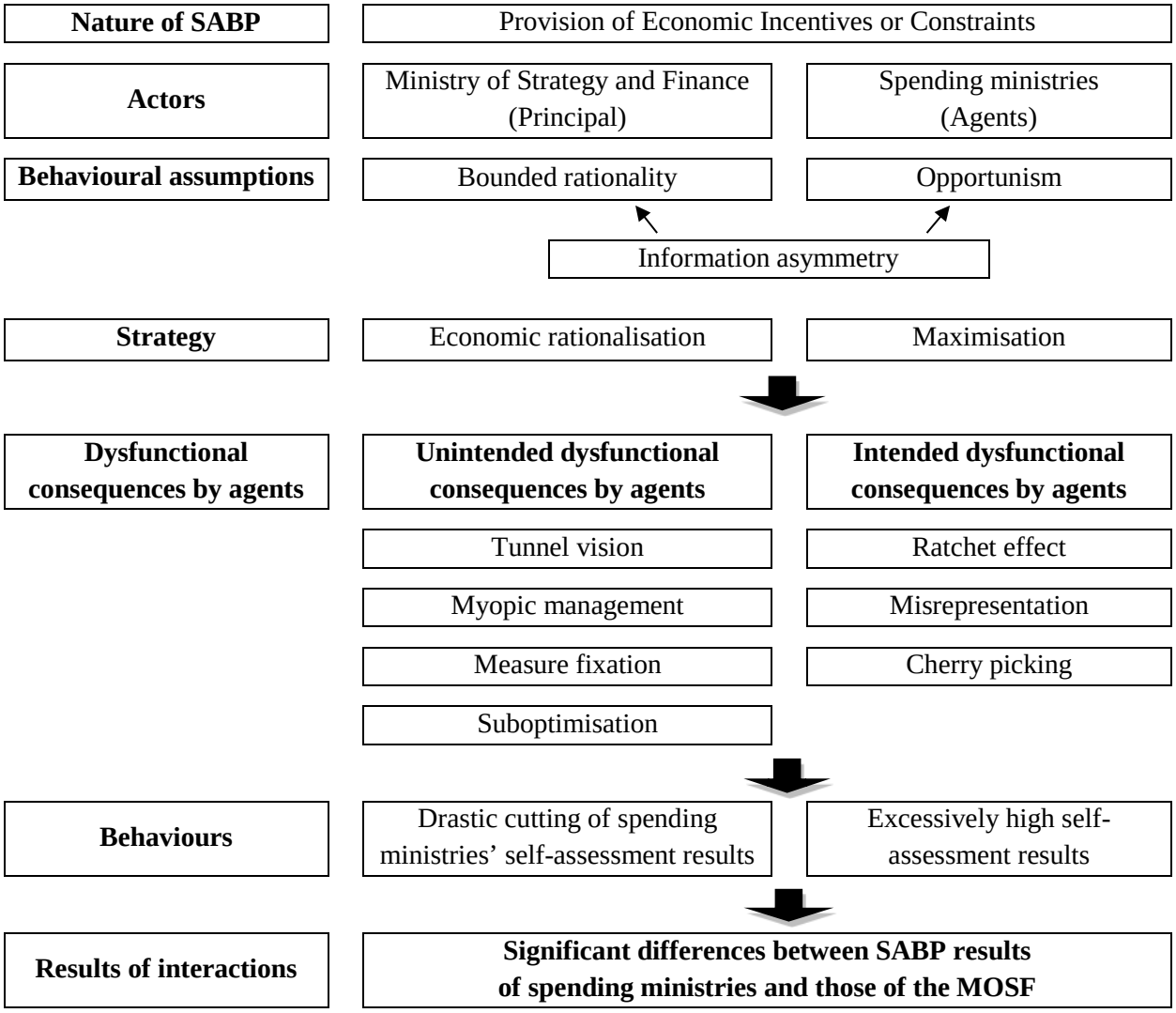
see that the MOSF's behavioural strategy is a kind of "economic rationalisation". On the other hand, the spending ministries' opportunistic behaviours result in a "maximisation" strategy, because these ministries recognise that the MOSF has a bounded rationality, and its reviewing process has to be based on their own self-assessment results. The spending ministries' maximisation strategy can lead to dysfunctional characteristics arising from the SABP process, some of which are unintended by both the spending ministries and the MOSF, while others are unintended by the designers of the SABP system but likely to be deliberate responses by the spending ministries, as agents in the principal-agent relationship. The optimism bias by spending ministries at their self-assessment stage may be caused by these dysfunctional consequences, and is finally expressed in an excessively high SABP score. Conversely, the MOSF's rationalisation strategy leads to drastic cutting of spending ministries' excessively high SABP scores. As a result of interaction between spending ministries and the MOSF, with principal-agent problems being affected by an asymmetric information situation, significant differences between the SABP scores and grades awarded by spending ministries and those awarded by the MOSF can arise.

Implications

On the basis of this conceptual framework, this study will explore what types of dysfunctional consequences exist, the extent to which they exist, and how these dysfunctional consequences have an influence on the behaviours of both spending ministries and the MOSF. Furthermore, the more important thing is to seek useful policy alternatives to prevent or reduce various transaction costs or dysfunctional

consequences arising from the process of SABP. Policy alternatives include both how SABP can reduce dysfunctional consequences and how an incentive system can be devised to control the relationships between spending ministries and the MOSF from the perspective of the principal-agent theory. These topics will mainly be discussed in Chapters 6 and 7. Figure 2-2 shows the conceptual framework for analysing dysfunctional consequences of SABP mainly through a qualitative approach.

Figure 2-2 Conceptual Framework for Analysing the Dysfunctional Consequences of SABP



Note: All dysfunctional consequences are "unintended" by the MOSF, the principal

2.4 Conclusion

This chapter has explored various definitions of performance and performance budgeting in the public sector; it has examined diverse dysfunctional consequences through a review of relevant literature according to three elements of performance management; it has described the principal-agent theory as a useful middle-range theory

to analyse dysfunctional consequences effectively; and it has also developed the thesis's particular typology of dysfunctional consequences of SABP. On the basis of these reviews in particular, the chapter has suggested a useful conceptual framework to explore dysfunctional consequences of the SABP system. Analysing dysfunctional consequences of SABP so as to seek feasible policy alternatives is important not only for securing the objectivity and reliability of performance budgeting but also for making sure that SABP measures genuine performance and whether the ultimate goal of a programme has been achieved. The conceptual framework suggested in this chapter will provide an especially useful basis for the qualitative approach, one of the main approaches of this thesis, to explore the difference between spending ministries and the MOSF in an SABP context. On the basis of these theoretical perspectives, the next chapter will examine a more practical side of performance budgeting. First, it will review the trends of performance budgeting in many countries, particularly focusing on the US case. Then, it will describe the SABP system in detail, including the SABP results from 2005 to 2010 and their meaning. It will also review many previous studies on the practical results of SABP and the PART, including their implications.

CHAPTER 3

PERFORMANCE BUDGETING IN PRACTICE

3.1 Introduction

While Chapter 2 offered a rather theory-oriented approach to performance budgeting, this chapter describes a more practical side of performance budgeting as it takes place in many countries. Performance budgeting has achieved prominence since the 1980s along with the influence of NPM. NPM, which is also called market-oriented public administration, entrepreneurial management of the public sector and government innovation, focuses on the outcomes and efficiency of the public sector by applying marketisation and competition to organisations in the public sector (Hood, 1991). James (2001, p.15) lists the common characteristics of many arguments for NPM as: "more explicit standards and measurement of performance, greater emphasis on output rather than input controls, a shift to disaggregation of units in the public sector, a shift to greater competition and contract based delivery of services, a stress on private sector styles of management practice and greater parsimony in resource use". New Zealand and Australia were the first to initiate performance management and budgeting, in their present form, in the late 1980s, followed by the US, the UK, Canada and Sweden in the early 1990s. Between the late 1990s and the early 2000s, Germany, Switzerland and

Austria introduced various budget reforms (OECD, 2004). In the case of developing countries, Malaysia was the forerunner in performance budgeting, in 1989, followed by Chile and South Africa in the 1990s. In the 2000s, Thailand implemented performance budgeting (Shah, 2007). As mentioned earlier, South Korea introduced SABP in 2005, following a trial run from 1999.

This chapter describes the practical aspects of performance budgeting, including SABP, in various countries, and then analyses previous studies of actual performance budgeting results, focusing on the cases of the US and Korea. To do this, the chapter firstly presents an account of the progress of performance budgeting in the US step by step from the initial stages of the system, such as PPB and ZBB, to recent efforts, such as the GPRA and PART; and it then explains the performance budgeting of other countries, including Korea. Exploring the various experiences of many countries can give useful information to help us understand the history, contextual background and common issues surrounding performance budgeting, and consider the implications for devising a better, but feasible, SABP system. Secondly, it introduces SABP, the main subject of this study, in detail e.g., the system's operating mechanisms, its checklist, the use of its results, the overall results from 2005 to 2010, and comparison between the SABP system in Korea and the PART in the US. Finally, it looks at many previous studies employing actual results of performance budgeting in the US and Korea, and evaluates their contributions and limitations. Limitations of previous studies, especially in relation to analysing the differences between the SABP scores of the spending ministries and those of the MOSF, will lead to meaningful research questions using a quantitative approach.

3.2 Trends of Performance Budgeting

3.2.1 Performance Budgeting in the US

In the US there have been numerous budgetary approaches, such as line item budgeting, activity-based budgeting, and performance-based budgeting, to allocate financial resources and manage public services. However, performance budgeting started to be implemented in earnest when the US Congress enacted the GPRA in 1993. Thus, this section describes the development of US performance budgeting before and after the GPRA.

Antecedents of the GPRA

Performance budgeting in the US incorporated performance information into public budgeting systems. The first trial of performance budgeting was the cost data budget in the New York Bureau of Municipal Research, in 1913, to show the objectives and activities of government; and at the level of the Federal Government, the Department of Agriculture first used programme budgeting in 1934. The Commission on the Organisation of the Executive Branch of the Government, commonly known as the Hoover Commission, recommended a performance budgeting system to President, Truman, as an official function of all Federal Government, in 1949 and the Budget and Accounting Procedures Act was enacted in 1951 (Hyde, 1992). The Commission's recommendations were pursued in an effort to reduce the immense government debt

after World War II, focusing on achieving the objectives of federal agencies. Although this system was not established as a permanent system, due to the difficulty of setting the performance measurement unit and unit cost, it became the origin of the performance budgeting system (Diamond, 2005).

Over the next four decades, several approaches to enhance the efficiency and effectiveness of the Federal Government's expenditures were implemented. Firstly, in 1961, PPB was introduced in the Defense Department's budgeting and management systems by Robert S. McNamara, President John F. Kennedy's Secretary of Defense. PPB was an attempt to forge a link between three-to-five-year long-term planning and resource requirements (programmes), as well as budgeting, in order to provide decision-makers with comparative cost and effectiveness information about various alternatives. Also, it required strategic planning, systematic analysis and cost-effectiveness analysis to make a decision more rationally (Rubin and Willoughby, 2011). Although the Defense Department used PPB for about 30 years more, the other departments did not use this system after 1971, due to the large amount of time and money required for analysing alternative approaches. While budget experts considered PPB a failure (Wildavsky, 1974), some elements of it still remain in the present performance budgeting system. Secondly, in 1973, the Nixon administration replaced PPB with MBO to manage federal agencies' and their sub-organisations' objectives clearly and regularly. However, MBO was considered "not as a new budget system but as a management approach which might come to be an integral part of the budget cycle" (Newland, 1976, p.20). Thirdly, President Carter introduced ZBB into the Federal Government in 1977 to resolve huge government debts. Under ZBB, the Government's

programmes were reorganised on the basis of their priorities. The first criterion for setting priorities was the government's accountability, followed by efficiency, effectiveness, influence and adaptability. However, ZBB was denigrated as decremental budgeting, the opposite concept to incrementalism, and was never implemented successfully by the Carter administration (Rubin and Willoughby, 2011; Hyde, 1992).

The GPRA, the PART and the GPRA Modernisation Act

The Clinton administration enacted the GPRA in 1993 to overcome the inefficiency of, and improve the accountability and effectiveness of, federal governments. This was "primarily a congressional initiative and marked the first time that the US Congress took specific legislative action to link strategic planning and performance budgeting" (Rubin and Willoughby, 2011, p. 9). The GRPA required strategic plans, annual performance plans, and annual programme performance reports (GAO, 1997a, 1997b), and strengthened the Federal Government's autonomy as well as responsibility, linking performance results and budget allocation. Federal government departments were required to submit strategic plans, including plans for at least five years, to the OMB and the Congress. Strategic plans needed to contain a mission statement, general goals, specific objectives, and methods to achieve these; and related federal government departments, local governments, service contractors, interest groups, employees and citizens could participate in drawing up the strategic plans. Annual performance plans included specific contents for implementing each programme, such as objective, quantified and measurable performance goals, and details of the human and physical resources required. Lastly, federal government departments needed to submit annual

reports on programme performance for the previous year in March every year to the President and the Congress. The GPRA had several successful features (GAO, 1997a, 1997b, 1999, 2004b). Firstly, the GAO found that the GPRA represented a great step forward compared to the early performance budgeting system of the 1950s in terms of linking annual performance plans and performance targets, as well as budget requirements. Secondly, the GPRA was implemented as a form of law, which was different from the previous initiatives, which had been a form of Presidential decree, so it could draw on stronger support from Congress. Thirdly, it tried to change the budgetary process gradually, whilst previous initiatives had aimed at a rapid change in policy. It has been revised several times through trial implementations following its introduction in 1993. However, the GAO (1997a, 1997b, 1999, 2004b) also found that the GPRA had some limitations in providing useful information for budget allocation and other fiscal policies. Firstly, in the case of crosscutting programmes, which involved two or more federal governments, there was significant incongruence between the various interest groups, and it was difficult to arrive at the same performance goals and objectives. Secondly, long-term strategic goals were lacking in details of how to link them with performance measures, so it was hard to obtain accurate and consistent performance data for measuring performance. Thirdly, it was difficult to link the achievement of goals with the efforts of federal government because performance could be determined by various external factors, such as the social and economic environment, as by well the efforts of federal governments.

The George W. Bush administration declared Budget and Performance Integration one of the five President's Management Agendas in 2001, and the PART was adopted from

the fiscal year 2003 in order to "explicitly bring performance information into the budget formulation process at the funding-decision level" (Breul, 2007, p.22). The PART was introduced to specify the legislative purposes of the GPRA, because the GPRA had been criticised for not achieving its intended purpose of bringing sufficient use of performance information into budget allocation (Dull, 2006). The PART's main focus was programme target levels, not agency goals, so the first step of the PART was the identification of programmes (Joyce, 2011). The PART could provide performance information consistent with the federal programmes' budget decision-making process, and the budget was redistributed from ineffective to effective grade programmes, based on the PART results of a programme. The PART contributed to increasing the use of performance information in budget allocation, and Shea (2008, p.5) points out that "more and more agencies recognise that the PART can be a useful tool for improving performance". However, the PART could not give effective feedback to a programme because the results were reflected in the next year's budget for the programme, not the current year's. Also, it was difficult to substantially link PART results with a programme's budget, and the Congress ignored the PART results when it came to budgeting, because the PART ratings were just one of the factors which it was necessary to consider (KIPF, 2011; Breul, 2007). Moreover, Gilmour (2007) indicates the learning effects on agencies over time. This showed that an overall increase in the PART score over time might reflect not a genuine improvement in the performance of a programme but the effects of learning how to deal with the PART. The details of the PART will be treated in a later section, with comparison to SABP.

The Obama administration, which began in 2009, did not use the PART any more, but improved the GPRA of 1993 by enacting the GPRA Modernization Act (GPRAMA) in 2010. The main change introduced by the GPRAMA was constant monitoring and performance management, especially on a small number of High Priority Performance Goals (HPPG) and crosscutting policy areas, rather than post-linkage between performance results and budget (OMB, 2010). The PART results were only reflected in the upcoming year's budget, while under the GPRAMA performance results cover two fiscal years (current and upcoming fiscal year). This is a change not only from the comprehensive assessment of all programmes to the selective evaluation of a limited number of important programmes, but also from the production of ex-post performance information to a constant monitoring and management system (Jung, 2012). The federal agencies set three to eight long-term HPPGs and submit annual performance plans, including clear performance goals, in line with these priority goals, to the OMB. The OMB monitors programmes' performance quarterly through the Vice-Minister's meeting, and it links performance information with dashboards containing general information about a programme, such as procedure, budget implementation and personnel, to increase the usefulness and availability of performance information. In addition, the GPRAMA tries to restructure programmes through in-depth programme evaluation. After implementation of programme evaluation by independent and professional institutions, such as the Domestic Policy Council, the National Economic Council, the Council of Economic Advisers and the OMB, the results of evaluation are released online (KIPF, 2011; OMB, 2010). Comparison of the GPRA and the GPRAMA in the strategic planning process is summarised in Table 3-1 below.

Table 3- 1 Strategic Planning and Performance Requirements under the GPRA and the GPRAMA

Topic	GPRA requirements	GPRAMA requirements
Strategic plan content	- Mission statement - Strategic goals - Strategies and identification of resources - Relationship to performance goals	- Relationship to federal government priority goals - Inter-agency coordination and collaboration - Identification of agency priority goals - Description of incorporation of Congressional input
Strategic planning process	- Revised at least every 3 years and covers at least a 5-year period - Consultations with Congress - Submitted to OMB and Congress	- Revised every 4 years - Consultations with relevant congress-ional committees at least every 2 years - Made available online and to President and Congress
Performance plan content	- Performance goals - Cover all programme activities - Strategies and resources - Performance measures - Basis for comparing actual results with performance goals - Means to verify/validate data	- Relationship to agency strategic and priority goals and federal government performance goals - Contributing programmes and activities - Milestones - Data accuracy and reliability - Identification of low-priority programmes
Performance plan process	- Cover the upcoming fiscal year - No set time frame - Submit to OMB	- Cover current and upcoming fiscal year - Timeframe with the President's Budget - Made available online

Source: Adapted from Rubin and Willoughby (2011)

Recent performance budgeting in the US is not a new form. Rather, it is a kind of re-established budgeting system, adopting the advantages and addressing some of the weaknesses of past budgetary reforms in order to reduce the huge amount of government debt accumulated since the 1980s (GAO, 1997a, 1997b). Although the GPRAMA has not yet been adopted completely, and it is difficult to predict its effect, it focuses on performance improvement of a programme rather than linkage between performance results and budgeting, whilst the PART put emphasis on both objectives at the same time. However, after implementing constant monitoring for enhancing

performance, it is expected that the GPRAMA will also seek to link performance information and budget allocation (KIPF, 2011).

3.2.2 Performance Budgeting in Other Countries

The United Kingdom

The UK's outcome-oriented performance budgeting system has been implemented as a strong fiscal reform since the time of the former Prime Minister, Margaret Thatcher, in the 1980s. For example, the Efficiency Scrutiny (1979), the Financial Management Initiative (FMI, 1982), the Next Steps Agency (1988) and the Citizen's Charter (1991) in the Thatcher administration, and Resource Accounting and Budgeting (RAB, 1998) and the Public Service Agreement (PSA, 1998) in the Tony Blair administration, are the main reforms (KIPF, 2008). In particular, in order to improve the quality and cost-effectiveness of public services and maintain sound public finances, the UK government made the carrying out of a three-year Comprehensive Spending Review (CSR) compulsory in 1998, and this has been underpinned by outcome-focused performance targets through the Public Service Agreement (PSA). The main principles of public service performance were "clear long-term strategic goals; independent audit and inspection; maximum local flexibility and discretion to innovate; transparency about what is being achieved" (Noman, 2007, pp. 194-195). Under the PSA, the ministries set the aims, objectives and performance targets, including technical notes about how to measure targets, and they were required to make a Service Delivery Agreement (SDA) to implement the PSA (Noman, 2007). The number of performance targets in the PSA

decreased from 600 in 1998-2001, to 300 in 2001-2004, 110 in 2004-2007 and 30 in 2007-2010, moving the focus of the PSA from output performance indicators to outcome-oriented and crosscutting performance indicators. The proportion of outcome-oriented performance indicators increased from 15% in 1998 to 100% in 2007 (Park, 2008; Jung, 2012). With the arrival of the Coalition Government in 2010, the PSA was changed into the "Public Sector Transparency Framework", focusing on managing individual ministries' performance rather than crosscutting programmes' performance. The 2010 CSR contained a Departmental Business Plan, which included the vision, priority, restructuring plan, expenditures and transparency of each department, in order to reduce government debt while remaining focused on the quality of public services (Jung, 2012). Under the PSA, HM Treasury examines each ministry's spending review and all central government departments make a PSA with HM Treasury. The Delivery Unit in the Prime Minister's Office checks the delivery plan of each ministry, and each ministry reports its implementation of core programmes to Parliament at least twice a year. The National Audit Office inspects the reliability and relevance of performance indicators and targets (KIPF, 2011).

Canada

Since programme evaluation was first implemented in the late of 1970s, the government of Canada has produced performance information continuously for citizen-focused, value-driven, results-oriented and responsible spending. In 1990s the Canadian Government established outcome-oriented management, focusing on linking outcomes and budgeting, and it launched a Public Service initiative in 2000 and a Declaration of

Quality Service (Bang, 2009). The main organisations involved are the Treasury Board Secretariat (TBS) as well as the internal audit and evaluation units within major departments and agencies. Canada has operated two kinds of performance management system: mapping performance information at the level of the government as a whole; and establishing performance management systems for each department. The former was specified as Canada's Performance, announced by the TBS from 2001, while the Management, Resources and Results Structure (MRRS) renewed in 2005, the Programme Evaluation launched in 1977, and the Management Accountability Framework (MAF) of 2003 were the core systems of the latter. The MRRS contains strategic outcomes, performance indicators and targets, so each department needs to develop Programme Activity Architecture (PAA) consisting of strategic outcomes, programme activities and a sub-activities level for hierarchical, results-based management (McCormack, 2007). Periodic (generally once per five years) Programme Evaluation is also being implemented to produce better on-going measurements. Besides Programme Evaluation for each department, the TBS has undertaken strategic reviews for all programmes directly implemented by the Federal Government in a four-year cycle since 2007. The Strategic Review selects the lowest performing five percentage expenditures and then transfers their budgets into other programmes. The World Bank (2010, p.96), interestingly, indicates that the probability of so-called "gaming" by spending ministries is low in the Strategic Review for the following reasons. "Firstly, it is strongly supported by the Prime Minister. Secondly, TBS officials are experts so they are aware of 'gaming' possibility. Thirdly, TBS can influence the performance appraisals and performance pay of Vice ministers. Fourthly, TBS officials thoroughly assess the recommendations of their reviews of Ministers and Vice ministers

in advance. Finally, it requires departments themselves to identify their lowest priority five percent of expenditures." All in all, Canada's performance budgeting system provides useful performance information for budget efficiency as well as general management after a long period of trial and error.

Australia

The government of Australia undertook budgetary reforms in 1983 in order to reduce public spending and improve the efficiency of public services. Programme Management and Budgeting (PMO) introduced the concept of evaluation and performance information, and changed the focus of budgeting from input to outcome. The Department of Finance (DOF) has a key role in coordinating and promoting the reformed system, and each department is required to make a Portfolio Evaluation Plan and submit it to the DOF (Bang, 2008). The current budgetary reforms are associated with the National Commission of Audit (1996), an independent review commissioned by the incoming government (KIPF, 2011). Key features of Australia's performance budgeting and management arrangements are: strong legislative foundations, such as the Financial Management and Accountability Act (FMA Act, 1997), the Commonwealth Authorities and Companies Act (1997), the Charter of Budget Honesty Act (1998) and the Public Services Act (1999, 2000) (Fox, 2007). The FMA Act (1997) suggests effective management guidelines for government assets and empowers individual ministers with both authority and accountability regarding financial management. The second law mainly stipulates requirements for commonwealth authorities and companies when it comes to managing non-public finance. The third law provides a

whole-government framework for the transparent conduct of fiscal policy. The Public Services Act, the fourth law, gives a foundation for accrual budgeting, including accrual appropriation, outcome and output-based management (KIPF, 2011). In addition, the DOF introduced a Strategic Review Unit in 2006 in order to provide useful information about major budgeting concerns for top decision makers. In general, the Strategic Review Unit evaluates high spending and crosscutting programmes with high priority (Jung, 2012).

The main contents of other countries' performance budgeting systems are summarised in Table 3-2 below.

Table 3-2 Main Contents of Other Countries' Performance Budgeting Systems

Country	Main contents
UK	▪ Strong fiscal reforms such as performance budgeting started in 1980s ▪ Compulsory Comprehensive Spending Review was introduced in 1998 (Underpinned by Public Service Agreement and specific Service Delivery Agreement between HM Treasury and Department) ▪ Outcome-oriented performance indicators are 100% in 2007 ▪ Public Sector Transparency Framework was implemented in 2010 to focus on individual ministries' performance rather than crosscutting programmes
Canada	▪ Declaration of Quality Service and Public Service 2000 was implemented in 1990s to establish outcome-oriented management ▪ Whole government level: Canada's Performance announced by TBS from 2001 ▪ Each department level: MRRS was renewed in 2005, Programme Evaluation was introduced from 1977, MAF from 2003 ▪ TBS undertook Strategic Reviews in 2007 with a four-year cycle to select the lowest performing 5 percent of expenditures in each department
Australia	▪ Programme Management and Budgeting were introduced in 1983 ▪ Current reforms are associated with the National Commission of Audit (1996) ▪ Strong legislative foundations: FMA Act in 1997, Commonwealth Authorities

	and Companies Act (1997), Charter of Budget Honesty Act (1998), Public Services Act (1999, 2000) ▪ Strategic Review Unit was implemented in 2006 to provide important performance information with top decision makers

3.2.3 Korean Performance Budgeting

Framework for Performance Management

Performance management in Korea has been implemented on the basis of two related laws: the GPA and the NFA[10]. The PMO manages the overall performance of central government in four functions – major policy, finance, personnel and organisation – on the basis of the GPA; while the MOSF controls performance budgeting on the basis of the NFA. In order to avoid conflicts between the GPA and the NFA, the PMO and the MOSF agreed that the PMO should take performance results by the NFA, like those of the finance sector by the GPA, without further assessment. The GPA stipulates the concepts, principles and objectives of government performance management more comprehensively, and defines performance management as "the activities that develop organisational missions, long and short-term objectives, annual targets and performance indicators and that manage government performance on the basis of economy, efficiency and effectiveness" (Art. 2). According to the GPA, ministries and agencies should develop a "strategic performance plan" for a five-year period, which includes

[10] With the commencement of the National Science and Technology Commission (NSTC) as a new control body for Research and Development (R&D) policies in Korea, in March 2011, responsibility for performance assessment on R&D programmes was transferred from the MOSF to the NSTC, according to the R&D Programme Performance Assessment and Performance Management Act.

their missions and strategic goals, and they should revise this at least every three year, after checking its validity. In addition, they should make an "annual performance plan" that accords with the five-year strategic performance plan (Art 6). The annual performance plan includes the missions, strategic goals, annual performance targets and indicators, and performance results of programmes by three-year period.

The NFA stipulates three sub-systems of the Korean performance budgeting system more directly and concretely: the PMS, SABP and the In-depth Evaluation System (IES). According to the PMS, central government departments should make an "annual performance plan", focusing on the financial aspect of their work, and draw up an "annual performance report" with performance indicators after implementing an annual performance plan (Art 8). Spending ministries (including agencies) which spend public budgets should submit an "annual performance plan" and an "annual performance report" to the MOSF each year with their budget request (Art. 8). Before submitting these to the National Assembly, the MOSF and the KIPF review the appropriateness of the performance indicators in each annual performance plan and advise spending ministries to revise these if they are not relevant. The differences between performance management by the GPA and that by the NFA (PMS) are summarised in Table 3-3 below.

Table 3-3 Comparison of Performance Management by the GPA with Performance Management by the NFA (PMS)

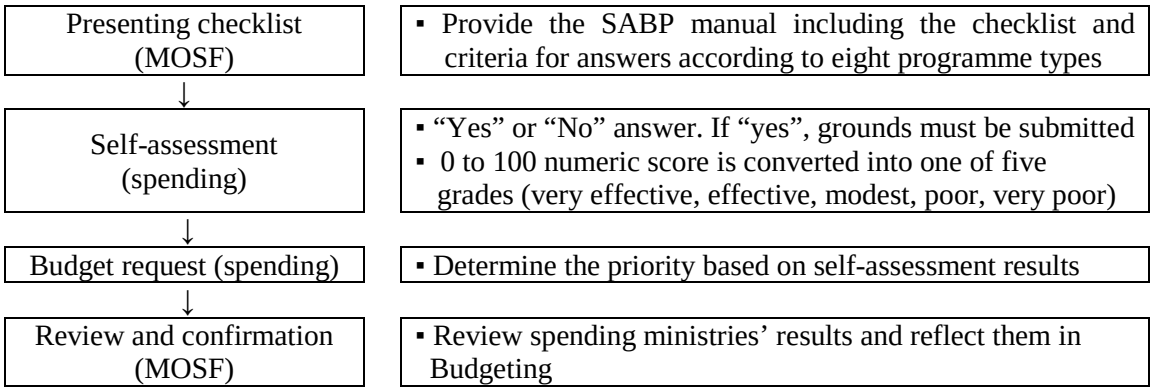
	GPA	NFA (PMS)
Agency in charge	PMO	MOSF
Implementation	1st April 2006 (enactment Mar. 2006)	1st January (enactment Sept. 2006)
Objects of	▪ Central government departments	▪ Central government departments which

performance management	* excludes independent organisations such as the Legislature, the Judiciary, the Board of Audit and Inspection	spend the public budget * includes the Legislature, the Judiciary, the Board of Audit and Inspection ▪ Public fund operators
Scope of performance management	▪ Strategic performance plan(5 years) ▪ Annual performance plan * no annual performance report	▪ Annual performance plan ▪ Annual performance report * no strategic performance plan
Procedure	Develop plans → report to relevant committee of the National Assembly * no guidelines, no submission	Distribute guidelines to spending ministries (MOSF) → develop and submit plan and report to the MOSF (spending ministries) → submit them to the National Assembly with budget proposals (MOSF, spending ministries)

Source: Prime Minister's Office (2008)

While the PMS, in accordance with the NFA, was implemented for managing the performance of central government in a macro way, SABP, in accordance with the NFA, was designed to control the performance of each budgetary programme in a more specific way through the checklist method. SABP is a kind of programme review intended to link SABP results and budget allocation. Since performance monitoring information from the PMS cannot be used directly in budget allocation, the SABP system employs a checklist which includes questions about the planning, management and results of each budgetary programme and produces a numeric score from 0 to 100. SABP has been implemented since 2005 and spending ministries are required to assess a third of all their programmes every year in order to reflect their self-assessment results in their budget request. Then, the MOSF reviews and adjusts the spending ministries' self-assessment results on the basis of the SABP manual, presented to spending ministries in advance, in order to make a budget proposal for the administration. The framework for SABP can be seen in Figure 3-1 below.

Figure 3-1 Framework for SABP



Finally, the IES was introduced to evaluate programmes in depth in 2006. The purpose of the IES is to analyse the specific performance improvement of programmes and describe the factors of success or failure by employing various scientific methods, through objective and independent organisations such as the KDI. The objects of the IES are programmes which need further evaluation after considering the SABP results, crosscutting programmes involving two or more ministries, programmes where it might be possible to waste the budget in an inefficient way, and programmes which are predicted to rapidly increase their budget expenditures in the future. The results of the IES are used in budgeting and improving programme management as well. Even if a programme is under way, if it is judged inefficient by the IES, then the MOSF will integrate it into another programme or reduce its budget or recommend changes to the method of implementation (KDI, 2007).

Although each of the three sub-systems produces rather different performance information, they are all closely linked to improving the performance of a programme and to increasing the efficiency of budget allocation. To sum up, the PMS is used to identify briefly the primary performance of a programme, focusing on performance

goals and indicators; SABP is used to link performance results and the budget of a programme through specific reviews of its planning, management and results; and the IES provides more in-depth evaluation of major programmes by identifying their problems and finding alternative solutions for them. The relationship between the three sub-systems of Korean performance budgeting is summarised in Figure 3-2 below.

Figure 3-2 The Relationship Between the Three Sub-Systems of Korean Performance Budgeting

	PMS (Monitoring)	SABP (Review)	IES (Evaluation)
▪ Object	- Entire programmes including non-budgetary programmes - 49 ministries, 2000 programmes annually	- A third of all budgetary programmes every year - About 450 programmes every year	- Selected budgetary programme - About 10 Programmes
▪Content	- Regular monitoring - No checklist - Monitor performance goal and indicator in annual performance plan - Use generally quantitative way - Implement internal Monitoring	- Regular review - Use checklist - Review performance goal and indicator + budget implementation, results - Quantitative, qualitative - MOSF finally evaluates on the basis of spending ministries' results	- A single evaluation - No checklist - Detailed analysis - Both ways - Internal, external, and self-assessment all possible
▪ Start year	- Trial basis in 2003 - Develop performance indicators in 2004 - Performance annual plan in 2005	- Since 2005	- Since 2006

Besides these post-assessment systems, interestingly, the Korean Government has, since 1999, also implemented Preliminary Feasibility Studies (PFS) before starting large budget programmes (over 50 billion Won). PFS consist of programme overviews, economic feasibility assessments such as benefit-cost analysis and net present value analysis, policy analysis, and general evaluations using the AHP (Analytic Hierarchy

Process) technique. PFS prevent projects with low feasibility from getting started as budgetary programmes. Since the main purpose of PFS is to increase the efficiency of budgetary programmes, it can be said that SABP is closely related to PFS. Also, it can be assumed that the higher the PFS results, the higher the SABP results will be.

The Background to Korean Performance Budgeting

The Korean Government tried to introduce performance management in the financial sector for a long time. However, performance budgeting, which links performance results and budgeting, has only been implemented since the 2000s. In 1962 and 1963, 15 programmes in three ministries, including the Ministry of Agriculture, adopted performance management selectively; but it failed, due to a lack of operational skills. Also, in 1999, after the Asian financial crisis, 16 ministries tried to implement performance management on a trial basis, but it did not spread across the rest of government because of a lack of preparation, which caused error in selecting programmes and poor content in performance plans and performance reports. After some trial and error, the PMS was introduced on a trial basis for 22 ministries in 2003. Thirty per cent of the programmes of these ministries had performance goals and indicators developed in 2004; and annual performance plans for all 22 ministries were completed in 2005. SABP was introduced in 2005 and the IES in 2006. In addition, the NFA was enacted in 2006, so the legal framework for Korean performance budgeting was complete. As SABP is closely related to other financial reforms, it is necessary to explore these in detail in order to understand SABP more completely.

Firstly, the medium-term expenditure framework was implemented in 2004 to change the direction of budgeting from one-year control to five-year strategy and goal. The medium-term budget plan is set in a five-year framework based on economic growth predictions, so it helps to decide the amount of future expenditure and to keep government debts at an acceptable level in the medium or long term. The medium-term expenditure framework is used as a formal framework for annual budget allocation and as an annual budget ceiling for each sector, unlike the informal documents of the MOSF. It is developed by open discussion from its initial stage and submitted to the National Assembly for validation; and it reflects not only the opinions of the MOSF but also the results of participation and consensus among related experts, ministries and interest groups.

Secondly, so-called “top-down” budgeting was introduced while the medium-term expenditure framework was being implemented in 2004. Once the annual budget ceiling for each sector under the medium-term expenditure framework is determined, as a result of open discussion, spending ministries set priorities among their programmes, coordinate opinions among internal bureaux, and reflect the interests of external groups within the budget ceiling. The top-down budgeting system has its own advantage in that it applies spending ministries’ expertise on the subject of their programmes to selecting necessary programmes and the budget size of these. Under the top-down budgeting system spending ministries’ autonomy in setting a budget has been increased compared to the previous bottom-up budgeting system. In this regard, the necessity for performance budgeting has been increased to require spending ministries to have greater

accountability corresponding to their greater discretion by strengthening the links between the performance of and budget for a programme.

Thirdly, the digital budget information system was introduced in 2007 to help the MOSF monitor the financial activities of central government in real time. It managed all kinds of financial information, including revenue, expenditure, budget implementation and settlements. In order to advance the budgeting system, the scientific management of related information was essential and the digital budget information system offered a basis for this. In addition, the digital budget information system provided a basis for changing the budget structure from line item budgeting to programme budgeting, presenting full costs, including direct costs, labour costs and indirect costs. Programme budgeting could increase the accountability and responsibility, as well as the autonomy, of spending ministries by measuring performance on a programme basis. Introducing programme budgeting was a basis for performance budgeting and a top-down budgeting system in Korea. A comparison between line item budgeting and a programme budgeting structure can be seen in Table 3-4 below.

Table 3-4 Comparison Between Line Item and Programme Budgeting Structures

Budget in 2006 (Line item)		Budget in 2007(Programme budgeting)		Annual performance plan
category	number	category	number	
Jang Kwan	28 76	Field (Jang) Sector(Kwan)	16 68	Strategic goals
Hang	963	Programme (Hang)	771	Performance goals
Se-hang	2,413	Sub-programme (Se-hang)	3,431	Task
Se-sehang	8,041	Activities (Se-sehang)	8,831	
Mok Se-mok	49 102	Cost element (Mok) Cost element (Se-mok)	23 90	

The strategic goals, performance goals and task in an annual performance plan are the field/sector, programme and sub-programme respectively in the programme budgeting structure. Under programme budgeting, each budgetary programme can be controlled in the form of “programme” and “sub-programme”. SABP, the main object of this study, has been implemented on a “sub-programme” basis. Since the year 2007, the foundation for performance budgeting has been built by changing from line item budgeting to a programme budgeting system.

3.2.4 Implications

From the trends in performance budgeting seen in the countries mentioned above, this study can draw some useful implications in terms of the use of performance information, the method of performance budgeting, outcome-oriented performance indicators, and the relationship between performance budgeting and other fiscal reforms. The implications mentioned in this section will be fully considered when developing feasible policy alternatives for a better SABP system in Chapter 7.

Firstly, regarding the use of performance information, Shick (2007) divided the contribution of performance budgeting into analytic tools and decision rules. He indicated that the former focused on providing useful information and insights for policy makers, while the latter paid more attention to consistency between increasing expenditures and improving outcomes. He also commented that, “analytic tools empower budget makers, allowing full scope for judgement and subjectivity, whereas decision rules constrain them, making budgeting less judgemental and more objective”

(Shick, 2007, p.111). Many OECD countries use performance budgeting to provide analytic tools for budgeting decision-making, not as decision rules. For example, the PART and SABP have been used as analytic tools because their results are one of the main factors which might influence budget allocation. When policy makers regard performance budgeting as providing analytic tools, this is something to consider cautiously: better integration of performance information into budget allocation is a long term challenge, requiring long term effort from the Ministry of Finance and spending ministries; and if policy makers want to use performance results in budgeting, they need to ensure that links between programmes, outputs and outcomes are clear and measured effectively (Fox, 2007).

Secondly, with respect to methods, performance budgeting can be divided into performance indicator monitoring, programme evaluation and programme review (KIPF, 2008). Performance indicator monitoring is the method which develops indicators related to outcome, output and process and then evaluates performance on the basis of these indicators. The GPRA in the US, the PSA in the UK and PMS in Korea can be regarded as performance indicator monitoring methods. Although these methods are able systematically to produce and utilise performance information within a relatively short time, it is difficult to consider the external factors which influence a programme's budget. Consequently, performance information at the level of performance monitoring cannot be directly applied to the budget decision-making process, despite the fact that it can be a starting point for performance management (KIPF, 2008). In addition, programme evaluation has mainly been used in order to develop policies in spending ministries, universities and academic institutes. The Canadian government has

implemented it for results-oriented performance management and improvement of programmes. Although programme evaluation includes detailed effects, directions for improvement, and the question of whether a programme should be continued or abolished, it takes more time and costs than other methods. Thus, there are some limitations for evaluating programmes comprehensively and providing performance information in a timely fashion. On the other hand, the programme review method has recently been adopted by many OECD countries: for example, the adoption of the PART in the US, Spending Reviews in the UK, a Strategic Programme Review in Canada, a Strategic Review in Australia and SABP in Korea. A programme review can analyse not only the effectiveness and efficiency of current programmes but also the outputs and outcomes of alternative programmes. More importantly, the results of a programme review can be reflected in budget allocation (KIPF, 2008). Due to the advantages of the programme review method, it has spread widely across many countries as an intermediate form between performance monitoring and programme evaluation. That is the reason why, of the three types of performance budgeting systems used in Korea, this research focuses on SABP.

Thirdly, recent performance budgeting systems put an emphasis on outcome-oriented information, because it is hard to understand the real and final effects of a programme just by using process or output information. For example, when evaluating whether a government's health investments are effective or not, examining whether rates of life expectancy or complete recovery (outcome indicator) are rising might be more meaningful than just looking at the increased number of patients treated (output indicator). Although the current trend of many countries is to develop outcome-oriented

performance indicators, measuring outcomes is technically more difficult, due to their complexity and interactions with many other factors. Talbot (2010) indicates some problematic issues of performance budgeting (e.g., unit of analysis, conceptual, technical, and political and values, problems), as mentioned in Chapter 2. Consequently, in order to make good use of outcome information, policy makers should take the following points into serious consideration: there is no end point for results-based management, and persistence over many years is required; strong and central leadership may be needed to build the capacity and culture that produces outcome information; and the Ministry of Finance and spending ministries need to focus continuously on reviewing outcomes and developing a more systematic approach to review existing programmes (McCormack, 2007; Fox, 2007).

Finally, performance budgeting systems in many countries have been developed with other financial reforms such as top-down budgeting, programme budgeting, medium-term budgeting and accrual basis accounting. Thus, a performance budgeting system's success is closely related to the success of other financial reforms. Shick (2007, p.129) impressively indicates that, "one of the most important lessons from half a century of disappointment is that budgeting cannot be transformed in isolation from the management practices and culture in which it is embedded".

3.3 Self-Assessment of Budgetary Programmes (SABP)

3.3.1 Overview and Operating Mechanisms

As mentioned in Section 3.2.3, although the PMS in 2004 controlled the performance of all the programmes of the central government of Korea, based on performance plans and performance indicators, there was a limit to the degree to which it was possible to link the performance results and budget of a programme, because performance information in the PMS was just monitoring level information [11] (Park, 2008). Against this background, SABP was introduced in 2005 in order to increase the linkage between SABP results and budget allocation. SABP was implemented in two major steps: a self-assessment stage for spending ministries (including agencies); and the MOSF's review of the spending ministries' self-assessment results. Generally, a third of all the programmes of each spending ministry are assessed by SABP every year. From 2005 to 2010, a total of 2,920 programmes were assessed by SABP and the budget for them totalled about 220,000 billion Won: 555 programmes (about 35,000bn Won) in 2005, 577 (35,100bn Won) in 2006, 585 (43,300bn Won) in 2007, 384(37,100bn Won) in 2008, 346(24,800bn Won) in 2009, and 473 (45,500bn Won) in 2010, as seen in Table 3-5 below.

[11] This problem was present in a similar way in the GPRA in the US. US federal government departments submitted strategic plans, annual performance plans and annual performance reports to the OMB and the Congress. However, the Congress did not use them well in the process of budget approval. That is why the PART was later adopted.

Table 3-5 Number of Programmes and Budgets Assessed by SABP from 2005 to 2010

	Total	2005	2006	2007	2008	2009	2010
Programme (unit)	2,920	555	577	585	384	346	473
Budget (bn Won)	220,000	35,000	35,100	43,300	37,100	24,800	45,500

SABP Checklist and Programme Types

The questions in the checklist for SABP are divided into two types: 12 common questions, which can be applied to all types of programme; and 1-2 specific questions, which depend on the programme type (MOSF, 2010a). A common question is related to three sections, such as planning (including rationale and design, and performance planning), management, and results, in order to assess a programme's performance. A specific question reflects the specificity of a programme and is included in the planning and management sections. All programmes are classified into one of eight categories: (1) IT System programmes; (2) Social Overhead Capital programmes; (3) Capital Acquisition programmes; (4) Subsidy to Private Sector programmes; (5) Subsidy to Local Government programmes; (6) Loan programmes; (7) Investment programmes ; (8) Other Direct programmes (MOSF, 2010a). In the case of IT System programmes, the current SABP checklist dates from 2008, as a result of government restructuring after the presidential election in 2007. As this study analyses the SABP results from 2005 to 2010, the SABP results of IT System programmes will not be considered in the study in order to obtain data consistency and comparability. An SABP checklist can be seen in Figure 3-3 below.

Figure 3- 3 SABP Checklist

Sections		Questions (Performance Indicators)	Weights Depending on Programme Types							
			IT	SOC	Capital Acquisition	Other Direct	Investment	Loan	Sub. to Private Sector	Sub. to Local Gov.
Planning (30)	Rationale and Design (15)	1-1. Is the purpose of programme clear and valid?	3.75	3.75	3.75	5.0	3.75	3.75	3.75	3.75
		1-2. Is the programme designed so that it is not redundant and does not duplicate any other programmes?	3.75	3.75	3.75	5.0	3.75	3.75	3.75	3.75
		1-3. Are the programme contents and implementation plan efficient and suitable?	3.75	3.75	3.75	5.0	3.75	3.75	3.75	3.75
		<Specific Checklists> 1-IT. Have possible major conditions been considered in planning? 1-SOC. Has the programme gone through objective feasibility studies and was its validity confirmed? 1-CA. Did the programme have any procedures to analyze the relevance of acquisition? 1-SubP. Has the programme reassessed whether the ongoing subsidy is necessary and considered the relevance of the beneficiaries and matching conditions? 1-SubL. Did the programme review and reflect local government ' s business conditions? 1-Loan. Have the terms and conditions been properly set? 1-Inv. Are the investment amounts given in the results of the investigation of the organization's conditions proper?	3.75	3.75	3.75	-	3.75	3.75	3.75	3.75
		Subtotal	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0
	Performance Planning (15)	2-1. Are there clear relationships between the performance indicators and the purpose of a programme?	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5
		2-2. Is the target for performance ambitious?	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5
		Subtotal	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0

Management (20)	3-1. Have funds been allotted in a timely manner?	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
	3-2. Is the programme operating the monitoring system so that it takes action to solve problems?	10.0	5.0	10.0	10.0	10.0	5.0	5.0	5.0
	3-3. Has the programme received the objective and comprehensive evaluation?	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
	<Specific Checklists> 3-SOC. Does the programme manage its aggregate cost properly? 3-Loan. Is the revolving rate sound and does the programme use the credible cost and risk prediction method? 3-SubP. Are the management and supervision over the programme partner sufficient? 3-SubL. Are local governments taking useful actions for improving the efficiency of allotted funds?	-	5.0	-	-	-	5.0	5.0	5.0
	Subtotal	20.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0
Results / Feedback (50)	4-1. Has the programme achieved its planned target of performance indicator?	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0
	4-2. As a result of programme evaluation, is the programme operating effectively?	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
	4-3. Has the programme increased the efficiency in the process of achieving the performance target?	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
	4-4. Have the assessment results and outside recommends used to improve the structure of programme?	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0
	Subtotal	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
Total		100	100	100	100	100	100	100	100

Note: The abbreviations in specific checklists for both the planning and management sections represent the programme types considered in the questions. IT means IT System programmes; SOC means Social Overhead Capital programmes; Sub. to Private Sector means Subsidy to Private Sector programmes; Sub. to Local Government means Subsidy to Local Government programmes

Source: A Guidance for Self-Assessment of Budgetary Programmes in the fiscal year 2009 (MOSF, 2010a)

Calculation of Assessment Results

Most of questions in the SABP checklist are to be answered with a “Yes” or “No”, and spending ministries are required to provide appropriate evidences for a “Yes” answer. The planning, management and results sections are weighted at 30%, 20% and 50% respectively, in order to produce numeric scores from 0 to 100. In particular, the planning section consists of two sub-sections (rationale and design, performance planning) and each sub-section is worth 15%. A “Yes” answer is converted into the maximum score for each question from 3.75 to 30.0 and a “No” scores zero. It should be noted that Question 4-1 has four kinds of answer: “Yes”, “To a large extent”, “To a small extent” and “No”. The four levels are translated into different scores, as shown in Table 3-6.

Table 3- 6 Four Levels of Answers and Scores for Question 4-1

Answer	No	To a small extent	To a large extent	Yes
Score	0	10	20	30

SABP was launched to improve the performance of the budgetary programme, so it emphasises the importance of performance planning by linking this logically with Questions 2-1, 2-2 and 4-1, as shown in Table 3-7. So if the answer to 2-1 is “No”, then 2-2 is automatically “No”, and the answer to Question 4-1 cannot be above “To a small extent”. In other words, a programme for which the answer to Question 2-1 is “No” cannot achieve a score above 65. Consequently, although the weight of performance planning is only 15%, spending ministries make great efforts to achieve a “Yes” answer to Questions 2-1 and 2-2, due to the logical relationship between the questions. This

indicates well why performance planning has greater importance than performance measurement and performance reporting, and why this study categorises various dysfunctional consequences according to three elements of performance budgeting, as mentioned in the Section 2.3.1. A standard for answering Questions 2-1, 2-2 and 4-1 will be explained in detail in the next section.

Table 3- 7 Logical Relationship Between Questions 2-1, 2-2 and 4-1

	Question 2-2	Question 4-1
Q. 2-1 = "No"	Automatically "No"	-
Q. 2-2 = "No"	-	Maximum is "To a small extent"(score = 10)

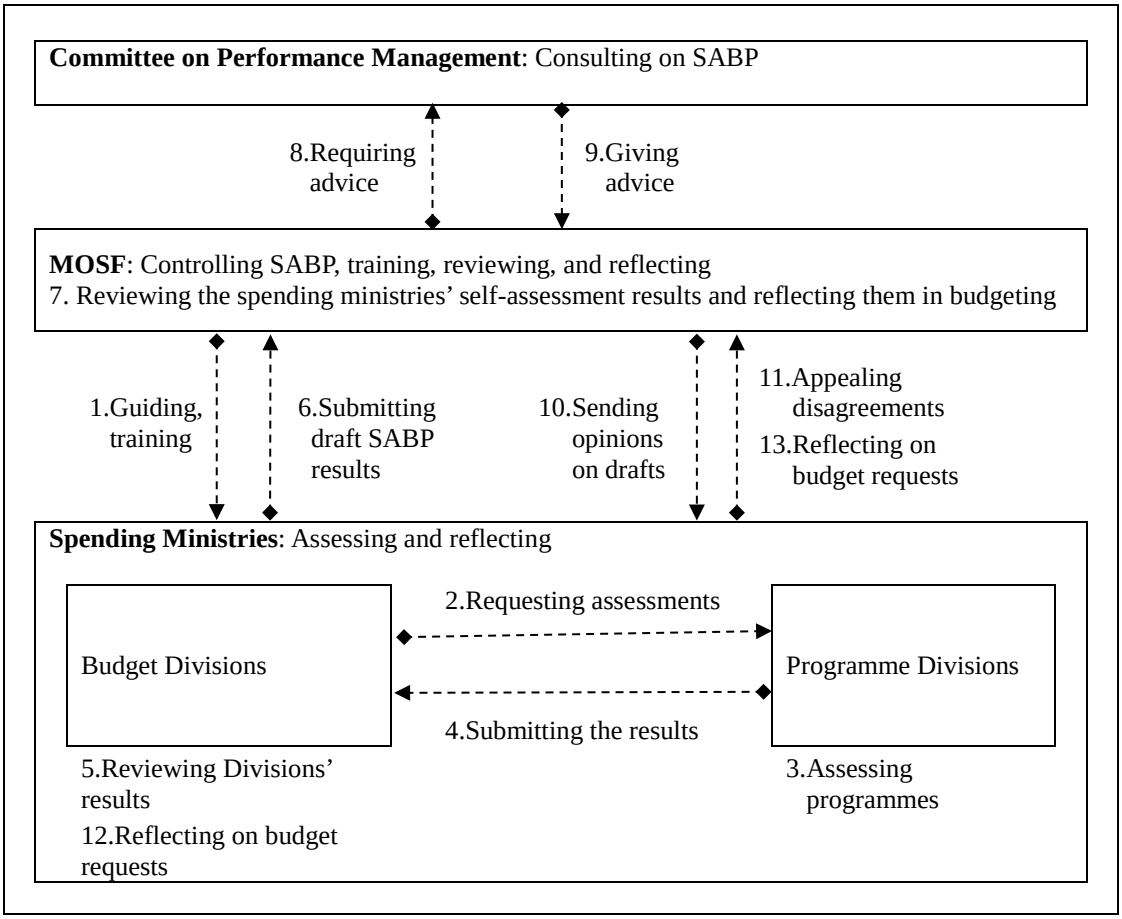
The total numerical score of each programme is converted into one of the following five grades: "Very effective" (100-90); "Effective" (89-80); "Modest" (79-60); "Poor" (59-50) and "Very poor" (less than 50). The SABP grades have been increased from four to five by dividing the "Poor" grade into "Poor" and "Very poor" from 2008 on, in order to enhance the assessment function among poor grade programmes. However, this study follows the four grade criteria ("Effective", "Somewhat effective", "Modest", "Poor") to maintain the consistency and comparability of SABP results. According to the Budgeting Guidelines (MOSF, 2010b), the budgets of a programme which is finally rated "Effective" need to be increased, in principle, the following year, whilst those of a programme which is finally rated "Poor" or below must be reduced by more than ten percentage points the following year.

Procedures of the SABP system

SABP has been applied to all the budgetary programmes of 49 central government departments, and the object of SABP, in principle, is the "sub-programme" in programme budgeting structure. However, programmes which are not considered worth assessing, such as operating costs (including labour costs), comprehensive grants to local government, and internal transfers within governments, are excluded from SABP. There are three bodies involved in the SABP system: 48 spending ministries (including programme divisions and budget divisions); the MOSF (including the Performance Management Bureau and the Budget Office), which controls the SABP system; and the Advisory Committee, which consists of budget directors from the MOSF and outside experts. The Performance Management Bureau in the MOSF distributes guidelines for SABP to spending ministries, and these enable the latter to evaluate their budgetary programmes. Budget divisions in the spending ministries are required to submit the results of self-assessment to the MOSF by March each year, after getting together their programme division's assessment; and they must also submit their budget requirements for the following year, reflecting their self-assessment results, by June. The MOSF reviews the validity of the spending ministries' self-assessment results from March to June. Until 2008, the Budget Office reviewed self-assessment results first; and then the Performance Management Bureau decided the final SABP results. During these periods, the MOSF discusses spending ministries' formal objections and determines final SABP results in consultation with the Advisory Committee. The Performance Management Bureau gives policy recommendations for each programme to the spending ministries, and the Budget Office uses the final SABP results when it comes to making a budget

proposal for submitting to the National Assembly by September. The procedures and roles of the spending ministries and the MOSF can be seen in Figure 3-4 below. As we can see from the figure, in order for SABP to be successful, cooperation and trust between spending ministries and the MOSF are essential. Otherwise, SABP users can be exposed to various unexpected negative effects when presenting performance plans, measuring performance and reporting performance.

Figure 3- 4 The Procedures and Roles of the Spending Ministries and the MOSF



Source: Adapted from Cho (2010)

3.3.2 Specific Criteria for Answering a Question in the Checklist

This section describes specific standards for answering Questions 2-1, 2-2 and 4-1. These questions are closely related to each other and can greatly influence the SABP score of a programme, so spending ministries make every effort to achieve a "Yes" answer, and therefore the possibility of a difference between the SABP result of a spending ministry and the result of the MOSF may be higher than for other questions. Figures 3-5, 3-6 and 3-7 provide specific criteria for answering Questions 2-1, 2-2 and 4-1 in the SABP manual (MOSF, 2010a).

With respect to the appropriateness of performance indicators, Question 2-1 requires that performance indicators should be related clearly and specifically to the purpose of a programme. In particular, in order to achieve a "Yes" answer to the question, firstly, a performance indicator should cover all the contents of a programme, that is, it should be representative of the whole programme; secondly, the performance indicator should be outcome-oriented; and thirdly, the definition of the performance indicator and the performance measurement method should be clear and reasonable. In addition, in a case where it is difficult to set an outcome indicator, due to the characteristics of the programme or the fact that the programme is still on-going, a "quality of output" or "process" indicator can also be accepted. These three criteria should all be accepted at the same point in the process of the MOSF's review in order to achieve a final "Yes" answer. From the specific criteria for answering Question 2-1, we can see that a device for preventing "measure fixation" and "tunnel vision" exists to some extent in the SABP manual.

Figure 3-5 Specific Criteria for Answering Question 2-1

Performance Indicator (Common question)	2-1. Is there a clear relationship between the performance indicators and the purpose of a programme?
Purpose	To determine whether the performance indicators can be used as means for measuring the purpose of the programme
Elements of *Yes* or *No*	□ Elements of *Yes*: A *Yes* answer needs to clearly explain and provide evidence of all the following: o A performance indicator should cover all the contents of a programme * A performance indicator should be representative of the whole programme, not a part of the programme, including programme contents and beneficiaries o An outcome indicator should be set as a performance indicator for a programme * The quality of output can be accepted in a case where there is difficulty in setting an outcome indicator, due to the characteristics of the programme (ex) promotion programme: not "the amount of promotion" measuring simple output, but "the extent of recognition of promotion" measuring the quality of output should be presented as a performance indicator * An on-going programme should set both a process indicator, such as the rate of progress, and an outcome indicator for the completed part * The rate of progress is measured by "total cost over accumulated investment" o Definition of a performance indicator and method of performance measurement should be specific and reasonable * If there is a specific performance measurement method, then the numerator and denominator should be explained in detail. Otherwise, the calculation method should be explained in detail □ Elements of *No*: A *No* answer is possible if a performance indictor is applied to at least one of the following cases: o Where there is a lack of relationship between the performance indicator and the programme's purpose o Where there is a focus on the input or the amount of output o Where the definition of the performance indicator and method of performance measurement are unclear and not reasonable o Where satisfaction with the indicator is suggested without simultaneously setting the quality of output or quantitative outcome indicator

Source: Guidance for Self-Assessment of Budgetary Programmes for the fiscal year 2009 (MOSF, 2010a)

In regard to the relevance of a performance target level, Question 2-2 requires that performance targets should be ambitious. In order to receive a "Yes" answer, spending ministries should set an ambitious performance target level before implementing a programme and consider the extent of automatic achievement without any special

efforts as well as performance improvement with efforts. It means that the grounds for setting performance target levels should be clear and reasonable. If the performance target is set at a level that is attainable without any special efforts, or it just considers past achievement and doesn't reflect efforts for improving performance, then spending ministries cannot obtain a "Yes" answer in the MOSF's review. Furthermore, Question 2-2 is closely related to Question 2-1, so if the answer to Question 2-1 is "No", then the answer to Question 2-2 is also automatically "No".

Figure 3-6 Specific Criteria for Answering Question 2-2

Performance Indicator (Common question)	2-2. Is the target for performance ambitious?
Purpose	To determine whether the performance target is ambitious compared with the achievement of, or the trend in achievement of, performance
Elements of *Yes* or *No*	□ Elements of *Yes*: A *Yes* answer needs to clearly explain and provide evidence of all the following cases: o Where the performance target is set in advance o Where the performance target can automatically be achieved without any special efforts * A predictable change in external environment should be considered before setting the performance target o Where the performance target includes efforts made for improving the performance, such as an improvement in work methods * The extent of the improvement of performance through the efforts made should be considered previously □ Elements of *No*: A *No* answer is possible if a performance indictor is applied to at least one of the following cases: o Where the performance target is set after the programme's implementation o Where the performance target is set at at a level of achievement which can be attained without any special efforts * Where changes in external environment, are not reflected in the target, which is set at the same level as for the previous year or for a lower level than the previous year o Where the performance target is based simply on past achievement, not considering the efforts needed to change things during the past work process ■ **If the answer to Question 2-1 is "No", then that to Question 2-2 is also "No"**

Source: Guidance for Self-Assessment of Budgetary Programmes in the fiscal year 2009 (MOSF, 2010a)

As mentioned above, Question 4-1, which asks whether the programme has achieved its planned performance target, is closely related to Questions 2-1 and 2-2. If the answer to Question 2-1 is "No", the answer to Question 2-2 is automatically "No"; and if the answer to Question 2-2 is "No", then the programme cannot achieve above "To a small extent" (10 points), although the performance target is achieved 100% according to specific criteria for answering Question 4-1. If spending ministries want to achieve a "Yes" answer, then they should achieve both a "Yes" answer to Question 2-2 and a performance target level of 100% or more.

Figure 3- 7 Specific Criteria for Answering Question 4-1

Performance Indicator (Common question)	4-1. Has the programme achieved its planned target of performance indicators?
Purpose	To determine whether the programme has achieved its ambitious target level or not
Elements of *Yes* or *No*	□ Elements of *Yes*: o In a case where the answer to Question 2-2 is "Yes" and the programme also achieves 100% or more of the performance target level □ Elements of *To a large extent*: o In a case where the answer to Question 2-2 is "Yes" and the programme also achieves its performance target level to a large extent o Although the answer to Question 2-2 is "Yes", and the programme achieves its performance target level 100%, at least one of the following is the case: * It is not possible to ascertain the ultimate effect of a programme e.g., it is a new programme or an on-going programme * The budgets of a programme are not being implemented normally □ Elements of *To a small extent*: o In a case where the answer to Question 2-2 is "Yes" and the programme achieves its performance target level to a small extent o Although the answer to Question 2-2 is "Yes", and it achieves its performance target level 100%, at least one of the following is the case: * Unreliable performance data is presented * There is continuous low budget implementation o **In a case where the answer to Question 2-2 is "No" despite the programme achieving a performance target level 100% or more** □ Elements of *No*: o All the other cases

Source: A Guidance for Self-Assessment of Budgetary Programme in the fiscal year 2009 (MOSF,2010a)

3.3.3 The Use of SABP Results

Although SABP results have been continuously applied to budget allocation for programmes or improving the performance management of programmes through the recommendations of the MOSF, the specific method was changed after 2008. From 2005 to 2007, SABP final scores for each programme were converted into one of four types of SABP grade: “Effective” (100-85 points), “Somewhat effective” (84-70), “Modest” (69-50) and “Poor” (less than 50). In order to enhance the discriminatory power of scores, however, the MOSF changed the four grades into five grades – “Very effective” (100-90), “Effective” (89-80), “Modest” (79-60), “Poor” (59-50) and “Very poor” (less than 50) – in 2008, because the proportion of “Effective” grades had been increasing due to the lessons learned by spending ministries. Also, from 2005 to 2007 there was no incentive for effective programmes, but the budgets of poor programmes were reduced by 10 percentage points compared to the previous year. From 2008 on, the budgets of very effective and effective programmes have been increased, in principle, by more than the average budget increase rate for other programmes, whilst those of poor and very poor programmes have been decreased by 10 percentage points compared to the previous year (MOSF, 2010a).

As for the PMO’s performance assessment of major policy, finance, personnel and organisation based on the GPA, the PMO and the MOSF agreed to match the assessed programmes with each other in the finance sector, in order to avoid conflicts between the GPA and the NFA. Consequently, the PMO regards SABP results as the means of assessing the finance sector according to the GPA without any further assessment.

Some spending ministries (including agencies), such as the National Police Agency, apply SABP results to personnel performance evaluation and performance-related salary for officials. In addition, many other spending ministries try to use SABP results as a means of evaluating personnel performance (Cho, 2010). According to a questionnaire survey of the National Assembly Budget Office (2007), 82.5% of central government officials surveyed answered that linking performance information with budgeting was desirable. Interestingly, in a questionnaire survey carried out by Bang (2009) using Likert's five-point scale, the question asking whether it was necessary to link SABP results with budgeting obtained 3.73 points out of five, whilst the question asking whether current SABP results were relevantly linked with budgeting obtained only 3.08 points. This implies that most government officials agreed with the necessity for linking performance results with budgeting, but at the same time they perceived that the current SABP system might not be as appropriate as was necessary. As a result, although SABP results have been applied more widely to budgeting, performance management and personnel performance evaluation, it is necessary to examine whether they are working well, or whether there are any adverse effects.

3.3.4 Differences in SABP Results between Spending Ministries and the MOSF

Overall SABP Results from 2005 to 2010

Overall SABP results from 2005 to 2010 can be seen in Table 3-8 below. Although the number of programmes assessed by SABP from 2005 to 2007 is similar, at around 550-

600, it decreased sharply from 2008, to about 350-450. This was because the MOSF enhanced the assessment object from “activities” to “sub-programme” in the programme budget structure mentioned in Table 3-4 while changing the budget structure from line item to programme budgeting in 2008. The average budget of assessed programmes for six years was about 37,000 billion Won and consisted of about 15% of the total budget for each year. For six years the grade most achieved was “Modest”, at around 65%. The least achieved grade was “Effective”, at around 6%, and the grade “Poor” was the second least achieved, at about 8.2%.

Table 3- 8 Overall SABP Results from 2005 to 2010

Year	Programme (number)	Effective	Somewhat effective	Modest	Poor	Budget (bn Won)	% of total total budget
2005	555	29 (5.2)	102 (18.4)	337 (60.7)	87 (15.7)	35,000	16.7
2006	577	30 (5.2)	94 (16.3)	388 (67.2)	65 (11.3)	35,100	15.7
2007	585	69 (11.8)	143 (24.4)	342 (58.5)	31 (5.3)	43,300	18.2
2008	384	27 (7.0)	108 (28.1)	236 (61.5)	13 (3.4)	37,100	13.9
2009	346	7 (2.0)	115 (33.3)	210 (60.7)	14 (4.0)	24,800	8.7
2010	473	13 (2.7)	52 (11.0)	378 (80.0)	30 (6.3)	45,500	15.6
Total	2,920	175 (6.0)	614 (21.0)	1,891 (64.8)	240 (8.2)	220,000	14.8

Notes: 1. The number in parenthesis is the ratio (%)
2. Although the grade “Poor” has been divided into “Poor”, and “Very poor”, and the grade “Effective has been divided into “Effective” and “Very effective” since 2008, in order to strengthen the criteria for the “Poor” grade, this study employs the former criteria to compare the data

The proportion of “Effective” grades doubled from 5.2% in 2005 to 11.8% in 2007, but it fell again after 2008 to 7%. The proportion of “Poor” grades fell from 15.7% in 2005 to 3.4% in 2008 and then remained at a low level. Interestingly, during the six years most programmes converged near the “Somewhat effective” and “Modest” grades. In the case of these two grades, the Budget Office in the MOSF can increase or decrease the budgets of these programmes at their discretion based on other criteria besides SABP results. The other parts of this section will focus on differences in the SABP

results allocated by spending ministries and the MOSF in order to find some meaningful characteristics when it comes to assessing programmes.

Differences in SABP Scores between Spending Ministries and the MOSF

Differences in SABP scores between spending ministries and the MOSF for six years can be seen in Table 3-9 and Figure 3-8 below. The characteristics related to differences in the SABP scores were as follows. Firstly, there were very high SABP score differences between spending ministries and the MOSF, and these were over 25 points, except for 2008. The average self-assessment score allocated by spending ministries was 89.5 points, while that of the MOSF was just 63.1 points. Considering the nature of the performance budgeting system, differences between spending ministries and the MOSF seem to be a natural phenomenon. However, the most important thing is that spending ministries' average self-assessment scores were nearly 90 points out of 100. This surely shows that spending ministries had an optimism bias at the self-assessment stage and they did not assess their programmes objectively. Moreover, this led to the MOSF's drastic cutting of spending ministries' self-assessment results, as mentioned in 2.3.4. Secondly, of the four detailed sections of SABP, the performance planning and results sections showed a high degree of difference between the spending ministries' scores and those of the MOSF, the average difference over six years being 5.3 points and 16.5 points respectively, whilst there were almost no score gaps in the rationale and design and the management sections. This means that the overall SABP score differences between spending ministries and the MOSF were mainly explained by differences in the performance planning and results sections. In addition, the big

differences in the performance planning and results sections can also be explained by the logical relationships between Questions 2-1, 2-2 and 4-1 in the SABP checklist, that is, if there are big differences in the performance planning section (Questions 2-1 and 2-2), then there will automatically be big differences in the results section (Question 4-1). Furthermore, it might imply that dysfunctional consequences of performance planning have greater impacts than those of performance measurement and performance reporting, as mentioned in 2.3.1. Thirdly, the high SABP score differences between spending ministries and the MOSF were not reduced in the six years after the SABP was implemented in 2005. This might mean that both the spending ministries' optimism bias and the MOSF's drastic cutting had been repeated every year for six years, and this could lead to a failure to achieve the purposes of SABP in the medium and long-term. So, differences in SABP scores between spending ministries and the MOSF have been high, significant and persistent, and it is surely worthwhile examining what has happened in the SABP system to cause this situation.

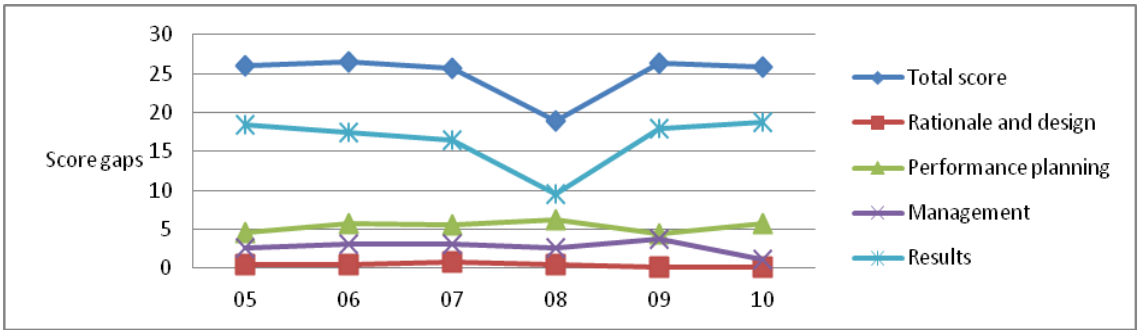
Table 3-9 Differences in SABP Scores Between Spending Ministries and the MOSF

(Unit: Points)

	Ministry	Total score	Plann -ing	(Rationale and design)	(Performance planning)	Mange -ment	Results
’05	Spending Ministry (A)	86.2	28.1	[14.2]	[13.9]	17.7	40.4
	MOSF (B)	60.1	23.1	[13.8]	[9.3]	15.1	21.9
	A-B	**26.1**	5.0	[0.4]	**[4.6]**	2.6	**18.5**
’06	Spending Ministry (A)	86.5	28.9	[14.7]	[14.3]	17.8	39.8
	MOSF (B)	59.9	22.9	[14.3]	[8.6]	14.7	22.3
	A-B	**26.6**	6.0	[0.4]	**[5.7]**	3.1	**17.5**
’07	Spending Ministry (A)	91.9	29.7	[14.9]	[14.7]	18.6	43.6
	MOSF (B)	66.2	23.6	[14.2]	[9.2]	15.5	27.1
	A-B	**25.7**	6.1	[0.7]	**[5.6]**	3.1	**16.5**
’08	Spending Ministry (A)	85.6	29.5	[14.9]	[14.6]	16.9	39.2
	MOSF (B)	66.7	22.8	[14.4]	[8.4]	14.3	29.6
	A-B	**18.9**	6.7	[0.5]	**[6.2]**	2.6	**9.6**
’09	Spending Ministry (A)	92.2	29.1	[14.7]	[14.3]	17.3	45.8
	MOSF (B)	65.9	24.5	[14.6]	[9.9]	13.6	27.8
	A-B	**26.3**	4.6	[0.1]	**[4.4]**	3.7	**18.0**
’10	Spending Ministry (A)	89.4	29.1	[14.9]	[14.2]	17.7	42.6
	MOSF (B)	63.6	23.1	[14.7]	[8.4]	16.6	23.9
	A-B	**25.8**	6.0	[0.2]	**[5.8]**	1.1	**18.7**
’05 - ’10	Spending Ministry (A)	89.5	29.1	[14.8]	[14.3]	17.7	41.9
	MOSF (B)	63.1	23.3	[14.3]	[9.0]	15.0	25.4
	A-B	**26.4**	5.8	[0.5]	**[5.3]**	2.7	**16.5**

Source: SABP Results from 2005 to 2010 (MOSF, 2012a)

Figure 3-8 Graph of Score Differences by SABP Section from 2005 to 2010



Source: SAPB Results from 2005 to 2010 (MOSF, 2012a)

Differences in SABP Grades between Spending Ministries and the MOSF

It will be also meaningful to examine the frequency distribution of SABP grades allocated by both spending ministries and the MOSF, because the budgets of programmes may be affected by SABP grades, according to the SABP manual. According to Table 3-10, an average of 95.9% of the spending ministries' SABP grades were distributed between "Effective" and "Somewhat effective", whilst an average of 85.8% of the MOSF's grades were concentrated in "Modest" and "Somewhat effective". In particular, an average of 74.9% of spending ministries' SABP grades was "Effective", whilst an average of 64.8% of the MOSF's grades was "Modest". Furthermore, these concentrations of SABP grades allocated by both the spending ministries and the MOSF were similar for six years. This might imply that there were big differences not only between the SABP grades but also between the SABP scores allocated by spending ministries and the MOSF for six years.

Table 3- 10 Frequency Ratio of SABP Grades from 2005 to 2010

(Unit: %)

		Total	Poor	Modest	Somewhat effective	Effective
'05	Spending Ministry	100.0	0.4	6.1	30.1	62.7
	MOSF	100.0	15.7	60.7	18.4	5.2
'06	Spending Ministry	100.0	1.2	4.7	27.6	66.6
	MOSF	100.0	11.3	67.2	16.3	5.2
'07	Spending Ministry	100.0	0.0	2.9	13.8	83.2
	MOSF	100.0	5.3	58.5	24.4	11.8
'08	Spending Ministry	100.0	0.3	2.6	8.9	88.3
	MOSF	100.0	3.4	61.5	28.1	7.0
'09	Spending Ministry	100.0	0.0	2.6	16.5	81.0
	MOSF	100.0	4.0	60.7	33.3	2.0
'10	Spending Ministry	100.0	0.0	3.0	23.5	73.6
	MOSF	100.0	6.3	80.0	11.0	2.7

’05-’10	Spending Ministry	100.0	0.3	3.8	21.0	74.9
	MOSF	100.0	8.2	64.8	21.0	6.0

Source: SABP Results from 2005 to 2010 (MOSF, 2012a)

3.3.5 Comparison of SABP with the PART

Before discussing previous studies on the results of both SABP and the PART, this section now compares SABP with the PART, because the PART in the US provided benchmarks for SABP in Korea. The PART is a systematic tool for assessing programme performance, a diagnostic tool for improving programme performance, and an informative tool for providing performance information for budgeting (OMB, 2005). PART scores are based on responses in four sections – programme purpose and design (20%); strategic planning (10%); programme management (20%); and programme results (50%) – and the usual number of common questions is 25, consisting of 5-10 questions in each section. Questions require a “Yes” or “No” answer like the SABP system does, and a numerical score of between 0 and 100 is converted into one of five grades: “Effective” (85-100); “Moderately effective” (70-84); “Adequate (50-69)”, “Ineffective (0-49)” and “RND (results not demonstrated)”. In particular, RND can be used when a programme does not have acceptable performance measures or lacks baseline and performance data. From 2002 to 2008, 1,015 programmes, 98% of all federal programmes, were assessed by the PART and the results can be seen in Table 3-11.

Table 3- 11 PART Scores and Programme Effectiveness

Programme effectiveness	PART scores	Number of programmes	%
Total	-	1,015	100.0
Effective	85-100	193	19.0
Moderately Effective	70-84	326	32.1
Adequate	50-69	297	29.3
Ineffective	0-49	26	2.7
Results not demonstrated (RND)	-	173	17.0

Source: http://georgewbush-whitehouse.archives.gov/omb/expectmore/about.html

Differences between SABP and the PART are as follows. Firstly, SABP was introduced as one of four fiscal reforms operated by the administration, while the PART started with a common recognition by the administration and the Congress of the necessity for reducing public expenditure. The consensus between the administration and the Congress was one of the factors that contributed to the success of the PART (Dull, 2006). Secondly, performance assessment and budget allocation in SABP is controlled by the Performance Management Bureau and the Budget Office within the MOSF respectively. However, in the PART the Resource Management Office within the OMB is responsible for both assessment and budget allocation. Considering the organisations related to performance budgeting, the PART seems to focus more on integrating performance results into budgeting than SABP does. However, interestingly, the extent to which performance results were reflected in budgeting was greater in SABP than in the PART, and that is one reason why the PART was not used in the Obama administration. Thirdly, programme types were different between the two systems. The types of programme on which the PART was used were: Direct Federal programmes (DF), Competitive Grant programmes (CO), Block/Formula Grant programmes (BF), Regulatory-Based programmes (RG), Capital Assets and Services Acquisition programmes (CA), Credit programmes (CR) and Research and Development

programmes (R&D). In particular, R&D programmes have been assessed separately from SABP since the National Science and Technology Commission (NSTC) was created in 2011. In addition, the weights in the PART are slightly flexible, while SABP weights are fixed. Finally, SABP results are applied by means of budgeting, performance management within the GPA, and personnel performance assessment in some ministries, while PART results are applied to budgeting. A summary of differences between SABP and the PART can be seen in Table 3-12 below.

Table 3- 12 Differences Between the SABP and the PART

		SABP	PART
▪Configuration	▪ Checklist & Weight	▪ Three sections: planning 30% (rationale and design 15%, performance planning 15%); management 20% and results 50% ▪ Weights are not changeable by ministries and agencies ▪ Weights of questions are not equal. ▪ Questions: 12 common questions and 1-2 specific questions depending on programme types	▪ Four sections: programme purpose/design 20%; strategic planning 10%; programme management 20%; and programme results 50% ▪ Weights are slightly flexible up to the point where questions are answered ▪ The weights of questions within a section, as a default setting, are equal ▪ Questions: 25 common questions and 5-10 specific questions depending on programme types
	▪Programme type	▪ Eight programme types: IT, Social Overhead Capital programmes (SOC), CA, Other Direct programmes, Investment programmes, Loan, Subsidy to Private Sector programmes, Subsidy to Local Government programmes * R&D programmes have been assessed separately from the SABP since 2011	▪ Seven programme types: Direct Federal programmes (DF); Competitive Grant programmes (CO); Block/Formula grant programmes (BF); Regulatory-based programmes (RG) ; Capital Assets and Services Acquisition programmes (CA); Credit programmes (CR) and Research and Development programmes (R&D)
	▪ Grade	▪ Four grades: Effective, Somewhat Effective, Modest and Poor ▪ RND (results not demonstrated) is not allowed	▪ Five grades: Effective, Moderately Effective, Adequate, Ineffective, RND
▪ Related organisation		▪ Performance Management Bureau within the MOSF * Budget Office within the MOSF reflects SABP results in budget allocation	▪ The Resource Management Office (RMO) within the OMB is responsible for both assessment and budget allocation
▪ Use of performance information		▪ Provides information as one of the key factors for budgeting. In particular, a “Poor” grade brings a 10 percentage points decrease in budgeting ▪ Apply SABP results to performance evaluation in the GPA (compulsory) and personnel performance assessment (voluntary)	▪ Providing information as one of the key factors for budgeting

Source: Modified Cho (2010)

3.4 Previous Studies and Implications

Previous studies that have analysed the results of SABP and the PART have focused on verifying the relationship between performance budgeting and the factors which affect budget decision-making or performance results. However, there have been almost no studies focusing on the differences in performance assessment results between spending ministries (federal government) and the MOSF (the OMB). This section presents previous studies under three categories: studies on the relationship between performance results and budgeting; studies on factors which affect budget decision-making or performance results; studies on the differences between spending ministries and the MOSF.

Firstly, Gilmour and Lewis (2005, 2006a) found, through regression analysis, that PART scores had a positive and statistically significant impact on recommended levels in the US President's budgets for the fiscal years 2004 and 2005. These results confirmed their hypothesis that programmes that had higher PART scores also got higher budgets. However, the "political content" of programmes[12], such as those located in a "Democratic department", appeared to produce results that differed between the FY 2005 budget and the FY 2004 budget. Contrary to their expectations, the study of the FY 2005 budget showed that the budgets of programmes located in so-called

[12] Gilmour and Lewis expressed the political content of programmes as 7 variables: belonging to a Democratic department; belonging to a core Democratic department; belonging to a department proposed for closure by the Republicans; the percentage increase in the previous year's budget; and the political environment in which a programme was established (Democratic President, Democratic Congress, Unified government).

"Democratic departments"[13] were not reduced statistically significantly by the Bush administration. Gilmour and Lewis suggest that possible reasons for this are substantial collinearity among measures of political content, the possibility of less political content in the budget of FY 2005, inadequate political content measures, and the indirect impacts of these on budgets (Gilmour and Lewis, 2005). On the other hand, in the study of the FY 2004 budget, they found that the budgets of programmes located in Democratic departments increased by 1.8 percent, whilst those of other programmes increased by an average of 5.6 percent, and this was statistically significant. It means that the programmes of Democratic departments were less supported than those of Republican departments (Gilmour and Lewis, 2006a). In addition, they found that programme characteristics such as programme type, size and age could also impact on budget allocation (Gilmour and Lewis, 2005, 2006a). They also found, through regression analysis, that the PART scores of political-appointee-run programmes were systematically lower in most management areas than those of senior-executive-run programmes, indicating the necessity of considering the proper role of political appointees in federal programme management (Gilmour and Lewis, 2006b). According to the GAO's questionnaire survey (2001), 43% of respondents answered that performance results played an important role in budget decision-making. However, Melkers and Willoughby (2001) revealed that just 39% of budgeting administration officers replied positively about the impact of performance results on budgeting.

[13] Gilmour and Lewis classified 7 departments as democratic departments: the Department of Commerce; the Department of Education; the Department of Energy, the Department of Housing and Urban Development; the Department of Labour, the Department of Health and Human Services; and the Environmental Protection Agency.

With regard to SABP, many Korean researchers reveal that there is a statistically significant relationship between SABP results (score and grade) and the budget allocations in the three budgeting stages (request by spending ministries; proposal by the administration; and final decision by the National Assembly), involving many variables which may affect budgeting, such as programme type (direct or subsidy), ministry type (welfare or economy related), ministry size (ministry or agency), and budget size (Park, 2005, 2008; Bang, 2009; Cho, 2010). Park (2005) firstly analyses the relationship between SABP results and budget allocation through regression analysis. He finds that the SABP results for 2005 do not have a statistically significant relationship with the budget for FY 2005, but there is a statistically significant relationship between the SABP results and the budget for FY 2006, stating that this may be evidence that performance results are being applied to budgeting. He also extends his analysis into the relationship between the SABP results from 2005 to 2007 and budgeting, and reveals that poor grades had negative effects on all three budget stages, and modest grades generally had a negative correlation with budgeting (Park, 2008). Bang (2009) also found that SABP results had a strong relationship with budgeting in 2005 and 2008, whilst the relationship became weak in 2006 and 2007. Also, he argues that SABP results are well applied to budgeting at the stage of proposals by the administration and at the final stage in the National Assembly, while spending ministries do not use SABP results actively. Moreover, he reveals that the relationships are negative in large budget programmes, contrary to expectations, showing that large programmes were restructured in order to prepare the way for an increase in funds for welfare expenditure from 2005 to 2008. In the case of subsidy and welfare-related programmes, the relationships are positive, although their SABP scores are low,

because members of the National Assembly are greatly interested in subsidy programmes and the administration pays more attention to welfare-related programmes than economy-related ones (Bang, 2009). In addition, Cho (2010) uses regression analysis to suggest that SABP results from 2006 to 2008 had statistically significant relationships with budget allocation in the three budgeting stages, and he states that some external factors, such as programme type, ministry type and budget size, affected budget allocation as well.

Secondly, some studies have focused on examining what factors may affect budget allocation or performance results. Yoon (2001), through regression analysis of 27 years' data, finds that a budget is decided according to the previous year's budget, inflation rates and the fiscal revenues of the government. Chang and Yoon (2002) also analyse empirically the factors which affect budget allocation on the basis of 27 years data from 1972 to 1998. Their analysis shows that partisan factors such as the percentage of seats held by the ruling party have the largest effect on budgeting, whilst other factors like GDP and presidential elections are not statistically significant. Kang (2007) states that the impact of the performance system may become greater when the performance objectives of an organisation are definite and concrete, performance indices are measureable, and the relationship between performance and budgeting is high. Kong et al. (2007) find that the performance planning section has more influence on SABP results than other sections in the checklist. Also, they argue that direct programmes show higher SABP scores than indirect ones, and that large programmes' SABP scores are higher than those of small programmes. Park (2005) reveals that each section in the checklist and each programme type have a different impact on SABP results, and Bang

and Yoon (2006) indicate through correlation analysis that the difficulty of programmes and programme types might influence SABP results. Table 3-13 presents previous studies on the relationship between performance results and budgeting, and factors which may affect budget allocation or performance results.

Table 3-13 Previous Studies on the Impact of Performance on Budget Decision-making

Researchers	Independent	Control variables	Dependent variables	Analysis method	Major findings
[Analysis of performance-budgeting relationship]					
Gilmour & Lewis (2005)	PART Score, Section	Democratic department programme, Rate of increase for previous year's budget, Type, size and age of programme	Government budget rate of increase	Regression analysis (FY 2005)	·(+)Score, section → △Budget ·(x)Political content → △Budget
Gilmour & Lewis (2006a)	PART Score	Democratic department programme, Rate of increase for previous year's budget, Programme age	Government budget rate of increase	Regression analysis (FY 2004)	·(+) Score →△Budget ·(o) Political content→ △Budget
Park (2005)	SABP Score, Grade	Rate of increase for previous year's budget	Government budget rate of increase	Regression, correlation (2006)	·(+) Results→△Budget ·(o)△Prior year→△Bud
Park (2008)	SABP Grade	Rate of increase for previous year's budget, Programme type, Budget and Ministry size	Budget rate of increase (three stages)	Regression analysis ('06-'07)	·(+) Grade →△Budget ·(o) Generally Control vars. → △Budget
Bang (2009)	SABP Score, Grade, Section	Welfare or economic related, Programme type and size	Budget rate of increase (three stages)	Regression and Survey ('05-'08)	·(+) Results → △Budget ·Some control variables affect budget changes
Cho (2010)	SABP Score, Grade, Section	Welfare or economic related programme, Program type, size and Ministry size	Budget rate of increase (three stages)	Regression and Survey ('06-'08)	·(+) Results → △Budget ·Some control variables affect budget changes
GAO (2001)	Score	-	Budgeting	Survey	·43%, greatly impact
Melkers et al. (2001)	Score	-	Budgeting	Survey	·39%, positive impact
[Analysis of Factors which Affect Budgeting or SABP results]					
Yoon (2001)	Previous year's budget	Inflation rate, Fiscal revenue	Budget rate of increase	Regression ('72-'98)	·(o) 3 variables → △Budget
Chang &Yoon (2002)	GDP, Ruling party's seat, etc.	Environmental factors, Budgeting system, Officers' behaviour	Budget rate of increase	Regression ('72-'98)	·(o) percentage of seats held by ruling party → △Budget ·(x) GDP, Presidential election etc.

Kang (2007)	-	Programme and organization characteristics	SABP results	Comparative Case analysis, Interview	·(o) Program, org. → SABP results
Kong et al. (2007)	SABP section	Programme type, Budget size	SABP results	Correlation and Regression	· Performance planning section has greater impact ·(o) Type, Size → SABP results
Park (2005)	SABP section	Programme type	SABP results	Correlation andRegression	·(o) Section, Type → SABP results
Bang et al.(2006)	-	Difficulty, Programme type	SABP results	Correlation	·(o) Difficulty, Type → SABP results

Note: (+) refers to positive relationships between variables, (o) means that there is a statistically significant relationship with variables, (x) means that there is no statistically significant relationship with variables, ∆Budget represents budget percentage changes (%) for the current year

Thirdly, there are some studies indicating possible differences in SABP results between spending ministries and the MOSF, and their potential side effects. Cho (2010) states that spending ministries show an "optimism bias" when assessing their programme in advance, indicating that the self-assessment scores of spending ministries are nearly 90 points out of 100 points and that, on average, this is 30 points higher than the review scores of the MOSF. Also, he mentions that the "SABP scores awarded by spending ministries and agencies vary little between programmes, so that they cannot provide useful pieces of performance information for prioritising programmes in making budget decisions" (Cho, 2010, p.294). Bang (2009) indicates that increases in spending ministries' self-assessment scores from 86.2 points in 2005 to 91.9 in 2007 may show the result of "optimism bias" by spending ministries, not the result of real performance improvement. Moreover, optimism bias can lead to large differences in SABP results between spending ministries and the MOSF, and therefore they may cause problems when SABP results are applied to budget allocation (Bang, 2009). Park (2005) simply calculates the differences between the SABP results of the spending ministries and

those of the MOSF from the 2005 results, using the concept of "disagreement ratio[14]". He finds that the disagreement ratio in SOC programmes is nearly 19%, higher than in other types of programmes, and the disagreement ratios in the performance planning and results sections are about 60% and 80% respectively. However, Cho (2010) and Bang (2009) just state the differences in SABP results through simple descriptive statistical analysis, and they do not analyse further the factors which affect these differences. Although Park (2005) calculates the specific differences by using the concept of disagreement ratio, he too does not carry out further analysis regarding the factors which might affect the disagreement ratio, and his analysis is limited to the year 2005.

These previous studies provide some useful insights for analysing the differences between spending ministries and the MOSF in this study. Firstly, analyses focusing on the relationship between performance results and budgeting are based on the premise that performance budgeting itself is being well implemented and its results are reliable and reflect the genuine performance of a programme. However, as mentioned in other studies, there have been high, significant and persistent differences in SABP results between spending ministries and the MOSF, and these raise a question about this assumption and the possible necessity for analysing whether SABP results represent the "actual" and "genuine" performance of a programme. Consequently, it is worthwhile analysing the process of how the final SABP results are arrived at and examining whether the SABP system is working well or not.

[14] He explains that the disagreement ratio between spending ministries and the MOSF in each programme follows the following formula: (number of disagreeing responses / total number of responses in SABP checklist) x 100.

Secondly, the fact that there have been hardly any studies on the differences between the SABP results of spending ministries and those of the MOSF, even though this might be very important, means that feasible policy alternatives that would improve the SABP system have not yet been fully discussed. High, significant and persistent differences in SABP results between spending ministries and the MOSF are surely unexpected or unanticipated dysfunctional consequences, especially for the MOSF, the designer of the SABP system. Analysing the unexpected consequences of SABP will shed a light on new policy implications which cannot be obtained from the analysis of expected consequences, because unexpected consequences have their own logics in the way they lead to these consequences. Creating a new policy without careful understanding or consideration of these logics will make the problem more complex or the effort useless (Lee, 2009). Consequently, this study seeks to suggest feasible policy implications for revising the SABP system by examining the factors which might affect the differences between the SABP results of spending ministries and those of the MOSF.

Thirdly, specific factors which might affect the differences in SABP results can be drawn from previous studies on the relationship between performance results and budgeting, because the reason why spending ministries have an optimism bias in the self-assessment stage is that ultimately they want to obtain bigger budgets. Thus, factors which might affect budgeting in previous studies, such as the SABP score, budget percentage changes for the current year, programme type (direct or indirect), budget size of programmes, types of organisation (ministry or agency) and purpose of programmes

(economy-related or welfare-related), can also be used to analyse the differences in SABP results between spending ministries and the MOSF.

These implications will be specified by developing research questions, operationalising related variables which affect the differences in SABP results, and constructing a multiple regression model in Chapter 4.

3.5 Conclusion

This chapter has described many countries' actual performance budgeting systems, including those of the US, the UK, Canada, Australia and Korea, and drawn some useful implications for this study. Many OECD countries have utilised performance budgeting as an analytic tool for budget decision-making and focused on outcome-oriented information in order to understand the real and ultimate effects of a programme. In addition, they have developed performance budgeting with other financial reforms such as top-down budgeting, programme budgeting and medium-term budgeting, and have mainly adopted the programme review method rather than performance indicator monitoring or programme evaluation. SABP in Korea, the PART in the US, the Strategic Programme Review in Canada and the Strategic Review in Australia are good examples of the programme review method. Considering these trends in performance budgeting among many OECD countries, the chapter has explained the SABP in detail, focusing especially on the differences between the SABP results of spending ministries and those of the MOSF. The chapter finds that these differences have been high, significant and persistent for six years. However, most previous studies have just

focused on whether the performance results affect budget allocation or what factors influence budgeting and performance results, on the premise that the performance budgeting system is operating well, and not examining specific reasons for the differences. Although there have been some initial studies that have shown an interest in the differences in SABP results, there have been almost no studies analysing the factors which might affect these. Therefore, the next chapter will specify and fill in the gap between actual policy and academic research by suggesting meaningful research questions and operationalising various factors which affect the differences in SABP results.

CHAPTER 4

RESEARCH METHODOLOGY

4.1 Introduction

As mentioned earlier, most studies on SABP have analysed the relationship between SABP results and subsequent budget allocations, and have not focused on the process for arriving at SABP results. However, differences in SABP results between spending ministries and the MOSF seem to have been high, significant and persistent; so it is surely worthwhile examining these. This chapter attempts to describe appropriate research methodology for analysing this phenomenon. This research employs both quantitative and qualitative approaches in order not only to explain the factors which might cause differences between the SABP results of spending ministries and those of the MOSF, through the analysis of numerical data such as six years of SABP results, but also to explore the existence of various dysfunctional consequences of SABP and their impacts on the differences between spending ministries and the MOSF through the analysis of qualitative data such as in-depth interviews. Moreover, the findings of both approaches will be discussed to suggest feasible policy alternatives that will produce a better SABP system.

The chapter begins by presenting the research objectives and the research questions established as a result of the review of relevant literature, and then matches the underlying philosophical worldviews in the research questions with the research methodology adopted in the study, which employs both quantitative and qualitative approaches. Regarding the quantitative approach, the chapter defines and operationalises the related variables influencing the differences in SABP results, and constructs a model to verify the relationships between variables. With respect to the qualitative approach, eight budgetary programmes in three case study areas that have undergone SABP are examined to explore various dysfunctional consequences of SABP and their impacts. The chapter explains the criteria for the selection of case study areas and specific research methods, focusing especially on in-depth interviews with Korean government officials of both spending ministries and the MOSF as well as experts in SABP. It also includes a research process for gathering relevant data for understanding differing perspectives on the dysfunctional consequences of SABP among ministries.

4.2 The Research Questions

This section explains how the research questions were chosen and why these questions are important for the study. The study starts by raising questions about the effectiveness of the SABP system, focusing on the differences between spending ministries and the MOSF. If the differences between spending ministries and the MOSF have been high, significant and persistent, then we cannot be sure that the SABP results are reliable and represent the “actual” or “genuine” performance of a programme, and moreover that SABP has fulfilled its purposes well. From the review of relevant studies on the results

of SABP in Chapter 3, it was clear that few researchers had focused on the differences between the SABP results of the spending ministries and those of the MOSF. Most studies had paid attention to the relationship between SABP results and budget allocation (Park, 2005, 2008; Bang, 2009; Cho, 2010) and analysed the factors which affected budgeting or SABP results (Yoon, 2001; Chang and Yoon, 2002; Kang, 2007; Kong et al., 2007; Pak, 2005; Bang et al., 2006). Even where differences in SABP results were indicated, the analysis remained at a simple descriptive level, not explaining in detail the factors which might affect the differences (Bang, 2009; Cho, 2010; Park, 2005). So, explaining these factors will provide useful insights for suggesting feasible policy implications for the refinement of SABP.

In addition, as mentioned in Chapter 2, this study categorises various dysfunctional consequences of performance budgeting into “unintended” and “intended” ones by spending ministries, whilst all are “unintended” by the MOSF, based on the principal-agent theory. According to this typology, exploring whether dysfunctional consequences of SABP actually exist, and what the impact of these might be on the differences between spending ministries and the MOSF, will also shed light on developing feasible policy alternatives.

Against this background, four research objectives are set out:

- To explain the factors which affect the differences between spending ministries and the MOSF

- To examine the types of, the extent of, and the reasons for, dysfunctional consequences of SABP

- To explore the influence of these dysfunctional consequences of SABP on spending ministries' self-assessment scores and the MOSF's review scores

- To suggest the role of spending ministries and the MOSF in improving the SABP system

Accordingly, the main research questions are:

(1) What are the factors which affect the differences in SABP results between spending ministries and the MOSF?

(2) What are the types of, extent of, and reasons for dysfunctional consequences of SABP?

(3) How do these dysfunctional consequences impact on both spending ministries' self-assessment scores and the MOSF's review scores?

(4) How can the SABP system in Korea be improved and refined?

These questions are explored using actual SABP results and previous studies on dysfunctional consequences of performance budgeting in the public sector. The first research question is closely related to the quantitative approach, focusing on whether

factors within the SABP system, such as the MOSF's review scores, budget percentage changes to programmes, programme types, the budget size of programmes and the types of organisation involved, can impact on the differences in SABP results which are mentioned in Chapter 3, by employing numerical data drawn from SABP results for six years. Examining these factors will lead to a more complete understanding of dysfunctional consequences of SABP and, along with exploration of the qualitative research questions discussed below, will help to draw out useful policy implications for revising the SABP system.

The second research question takes a qualitative approach and explores whether various dysfunctional consequences, "unintended" and "intended" by spending ministries but all "unintended" by the MOSF, as mentioned in Chapter 2, actually exist in the SABP system, and if they do, what their extent is and what the reasons for them are. It does this by looking at eight programmes in three case study areas. The second research question is important because the answers will indicate whether, and to what extent, SABP has actually been exposed to the various dysfunctional consequences described in many existing studies of performance budgeting in the public sector (Kelman and Friedman, 2009; Bevan and Hood, 2006; de Bruijn, 2002; Thiel and Leeuw, 2002; Smith, 1995; LeGrand and Bartlett, 1993; Bouckaert and Balk, 1991).

As an extension of the second research question, the third research question is designed to explore the impact of these dysfunctional consequences, focusing especially on spending ministries' optimism bias in the self-assessment stage and the MOSF's drastic cutting of spending ministries' results. This question is important because, by

examining the relationship between dysfunctional consequences of SABP and both spending ministries' and the MOSF's SABP scores, we can develop a deeper understanding of the cause of differences between spending ministries and the MOSF.

The last research question is the most important and is intended to generate information that is likely to have useful policy implications for refining and enhancing the SABP system in order to increase the efficiency and effectiveness of budgetary programmes in Korea. This question elicits the fact that the most important thing in performance budgeting is not the existence of dysfunctional consequences but the possibility of feasible policies to solve them. Also, it suggests that policy makers should be cautious when it comes to designing and refining the performance budgeting system.

4.3 The Selection of a Research Design

4.3.1 Philosophical Worldview and Research Approach

Research methodology and research methods should be in line with the researcher's philosophical worldview or paradigm (Bryman, 2008; Creswell, 2009; Creswell and Plano Clark, 2011; Grix, 2010; Namkoong, 2010), and they are designed to provide proper data to address the research questions and objectives (Yin, 2009). Creswell takes the term "worldview" to mean "a general orientation about the world and the nature of research that a researcher holds" (Creswell, 2009, p.6), and other researchers have referred to these ideas as "paradigms" (Kuhn, 1970) or "epistemologies" or "ontologies" (Archer, 1988). Bryman (1988, p.4) defines a paradigm as "a cluster of beliefs and

dictates which for scientists in a particular discipline influence what should be studied; how research should be done; [and] how results should be interpreted"; and Archer (1988, p.273) states that epistemology is related to "the nature of knowledge" and ontology is associated with "the nature of reality". Creswell (2009) divides researchers' worldviews into four categories: postpositivist, constructivist[15], advocacy/participatory, and pragmatist. According to Creswell (2009, pp.7-10), postpositivists focus on assessing cause and outcomes based on a deterministic philosophy, for example, in experiments; constructivists look for "the participants' views of the situation being studied"; an advocacy/participatory worldview is interested in "an action agenda for reform that may change the lives of the participants"; and pragmatism "emphasises the research problem and uses all approaches available to understand the problem and it applies to mixed methods research in that inquirers draw liberally from both quantitative and qualitative assumptions". He also indicates that quantitative approaches are more closely related to postpositivists' philosophical assumptions, qualitative approaches tend to employ constructivist or advocacy/participatory assumptions, and mixed methods approaches apply pragmatic assumptions. Similarly, Bryman (2008) and Grix (2010) state that quantitative research is more concerned with positivism and objectivism (the foundationalist's assumption) whilst qualitative research is more related to interpretivism and constructivism (the anti-foundationalist's assumption).

[15] Positivism is an epistemological position that advocates the application of the methods of the natural sciences to the study of social reality, whilst interpretivism is a contrasting epistemology to positivism and requires social scientists to grasp the subjective meaning of social action. Objectivism is an ontological position that implies that social phenomena can and should be considered objective entities, external to social actors, while constructivism considers that social phenomena are continually being accomplished by social actors (Bryman, 2008, pp.13-23).

With respect to the possibility of a mixed approach, there are two main versions of this kind of approach (Bryman, 2008, p.606): an epistemological version and a technical version. From an epistemological point of view, mixed methods research is not possible because quantitative and qualitative approaches are considered to be grounded in incompatible epistemological principles. The technical version, however, focuses on the greater strength of mixed methods research when it comes to collecting and analysing data. It too recognises that quantitative and qualitative approaches are based on "distinctive epistemological and ontological assumptions"; but it emphasises the flexible connections between these. It states that "mixed methods research becomes both feasible and desirable" (ibid, p.606) and "has become unexceptional and unremarkable in recent years" (Bryman, 2006, p.97) although some researchers express unease about this position (Buchanan, 1992; Pawson and Tilly, 1997). The definition of mixed methods research has been changed from mixing two methods to mixing in all phases of the research process (e.g., mixing philosophical positions, inferences and the interpretations of results) (Creswell and Plano Clark, 2011). King et al. (2001) indicate that the differences between quantitative and qualitative approaches only relate to style and technical methods, and they are not methodologically and substantively important. Also, Namkoong (2010) states that if the differences between two approaches lie in research methods they can be used complementarily for explaining and exploring one phenomenon. Greene et al. (1989, p. 256) define mixed method designs as "those that include at least one quantitative method (designed to collect numbers) and one qualitative method (designed to collect words), where neither type of method is inherently linked to any particular inquiry paradigm". Also, Tashakkori and Teddlie (1998, p.ix) define mixed methods as "the combination of qualitative and quantitative

approaches in the methodology of a study". However, this study makes no further attempts to discuss methodological issues such as whether mixed methods are possible or to what extent combining quantitative and qualitative research is desirable. Instead, it seeks to develop an appropriate research design for addressing the research objectives and research questions on the premise that it is possible to apply both quantitative and qualitative approaches in order to explain and understand one phenomenon more completely and comprehensively under a pragmatic worldview.

4.3.2 Research Design of the Study

The objectives of the study are both to explain the factors which affect the differences in SABP results between spending ministries and the MOSF by employing a quantitative approach and to explore dysfunctional consequences of SABP and their impacts on these differences based on a qualitative approach, in order to suggest feasible policy alternatives for reducing the differences and preventing dysfunctional consequences. In other words, these aims include not only verifying the relationships between some factors and the differences between the SABP results produced by spending ministries and those of the MOSF but also understanding the subjective meanings of government officials' perceptions, beliefs and behaviours. These objectives are in line with a pragmatic philosophical worldview and this approach does not "mix" the different philosophical worldviews, but uses an umbrella paradigm such as "pragmatism" for the study (Creswell and Plano Clark, 2011). Under the assumptions of pragmatism, the focus of the research is on the consequences of study, on the use of multiple methods as well as different forms of data collection and analysis, and on "what" and "how" to

research (Creswell, 2009; Cherryholmes, 1992; Morgan, 2007). By employing both quantitative and qualitative approaches, the study can bring a more thorough account and mutual corroboration of the existence of, and the reasons for, the differences between spending ministries' results and those of the MOSF in the SABP system. Moreover, quantitative and qualitative approaches can answer different research questions and enhance the integrity of findings, and therefore combining two approaches will give a more useful insight to policy makers who want to improve the effectiveness and efficiency of SABP (Greene et al., 1989; Bryman, 2006).

In this sense, the first research question is related to the quantitative approach while the second and third research questions are associated with the qualitative approach, and the last research question is linked with both approaches. As a result, data collection and data analysis are to be implemented separately and concurrently, with equal priority, by employing quantitative and qualitative analytic procedures up to the point where conclusions are drawn during the overall interpretation. The main source for quantitative data is six years of numerical results for SABP, while the sources for qualitative data are document analysis, archival records analysis and, particularly, in-depth interviews. Against this background, the rest of this chapter will describe the quantitative and the qualitative approaches used in detail.

4.4 Quantitative Approach

4.4.1 Introduction

The purpose of this section is to address the first research question, and to obtain useful insights for suggesting feasible policy alternatives. That is, this section will examine the factors which affect differences between spending ministries and the MOSF by employing various kinds of statistical methods (e.g., correlation analysis, t-test, ANOVA, regression analysis). More specifically, the study examines potential factors, such as MOSF review score, budget percentage change, programme type, budget size of programme and purpose of programme, which might affect these differences. Thus, the dependent variable is the differences in SABP results between spending ministries and the MOSF, which will be operationalised later. Also, as shown by many previous studies on SABP results, variables which are internally related to the differences can be treated as independent variables, and variables which are externally related to the differences can be dealt with as control variables. In this sense, the study sets "MOSF's review score" and "specific programme's self-assessment score minus the spending ministry's average score in that year" as independent variables; and it divides control variables into budget process-related factors, programme-characteristics-related factors and policy/politics-related factors. Budget percentage change for the current year can be taken as the first group; programme type (direct or indirect), budget size (small or large), and type of organisation (ministry or agency) can be taken as the second group; and purpose of programme (economy- or social welfare-related) and the number of SABP implementations on programmes can be taken as the third group. The next section

firstly operationalises the dependent, independent and control variables and explains their meanings for the study; it then constructs a model for analysing the relationships between various factors and difference in SABP results; and finally it explains the methods of data collection and data analysis.

4.4.2 Definitions and Meanings of Variables

Dependent Variables

① Disagreement ratio (DR)

The differences in SABP results between spending ministries and the MOSF can be measured by compiling the "disagreement ratio of responses in the SABP checklist" for each programme. As mentioned in Section 3.4, Park (2005) applies this concept to assess differences in SABP results and operationalises it in the following formula: (number of disagreeing responses / total number of responses in SABP checklist) x 100 in each programme. Table 4-1 illustrates simply how the disagreement ratio can be calculated for each programme. If there are five questions in the SABP checklist for a programme and three of these five questions are answered differently by spending ministries and the MOSF, then the disagreement ratio of this programme is 60% (3/5*100). This variable therefore captures all the differences in responses between spending ministries and the MOSF. In particular, it does not differentiate between a Q2 or Q3 situation (in which the spending ministry answers "Yes" but the MOSF answers

"No") and a Q4 situation (where the spending ministry answers "No" but the MOSF answers "Yes").

Table 4-1 How to Calculate Disagreement Ratio (Example)

Q1		Q2		Q3		Q4		Q5	
Spending	MOSF	Spending	MOSF	Spending	MOSF	Spending	MOSF	Spending	MOSF
Yes	Yes	Yes	No	Yes	No	No	Yes	No	No

Consequently, the DR does not consider the opposite meanings represented by a Q2 or Q3 situation (or, as we call it, a "Yes but No" type disagreement) and a Q4 situation (a "No but Yes" type disagreement), although the two different types of disagreements can be influenced in different directions by the independent and control variables, which will be explained later. Thus, we need to consider another type of DR, as discussed below, reflecting the opposite meaning of two types of disagreement between spending ministries and the MOSF.

② Relative Optimism Ratio of Spending Ministries (ROR)

In order to consider the opposite meanings of a "Yes but No" type disagreement and a "No but Yes" type disagreement, we can give an opposite sign to each disagreement e.g., +1 and -1 respectively, and then we can extract other kinds of dependent variable. The study calls this variable a "relative optimism ratio (ROR) of spending ministries" and focuses more on this variable, because it can represent the opposite implications of two types of disagreement for diverse independent and control variables. Table 4-2 shows how the ROR of a spending ministry can be calculated for each programme and the ROR of this programme is 20% (+1+1-1/5*100).

Table 4-2 How to Calculate the Relative Optimism Ratio of a Spending Ministry (Example)

Q1		Q2 (+1)		Q3 (+1)		Q4 (-1)		Q5	
Spending	MOSF	Spending	MOSF	Spending	MOSF	Spending	MOSF	Spending	MOSF
Yes	Yes	Yes	No	Yes	No	No	Yes	No	No

By comparing the impacts of diverse variables on two kinds of dependent variables (DR and ROR), the study can give more comprehensive explanations of the relationships between variables. The total number of responses for 2,920 programmes to the SABP checklist is 42,648. Spending ministries and the MOSF answered differently on 11,515 of these responses, so the DR for six years is 27.0%. The average DR in the first SABP cycle, from 2005 to 2007, is 25.4%, while that for the second cycle, from 2008 to 2010, is 29.9%. On the other hand, among the 11,515 responses, 10,443 (90.7%) are “Yes but No” type responses and 1,072 (9.3%) are “No but Yes” type. Thus, the ROR for six years is 22.0%. In addition, the ROR in the first cycle is 20.9% while that in the second cycle is 24.2%. Interestingly, at first glance it seems that the differences between spending ministries and the MOSF have not been reduced even though SABP has been implemented several times. However, specific relationships between the number of SABP implementation and the DR or ROR will be explained later. The details of the DR and ROR for each year are presented in Table 4-3 below.

Table 4-3 Number of Responses to the SABP Checklist, the DR and ROR

(Unit: number of programme and responses, %)

	First SABP cycle			Second SABP cycle			Six years
	’05	’06	’07	’08	’09	’10	
Number of programmes	555	577	585	384	346	473	2,920
Total number of responses	8,353	9,840	9,396	4,992	3,806	6,281	42,648
Number of different responses	2,081	2,612	2,335	1,640	1,100	1,747	11,515

Number of “Yes but No” type	1,947 (93.6)	2,324 (89.0)	2,127 (91.1)	1,484 (90.5)	1,019 (92.6)	1,542 (88.3)	10,443 (90.7)
Number of “No but Yes” type	134 (6.4)	288 (11.0)	208 (8.9)	156 (9.5)	81 (7.4)	205 (11.7)	1,072 (9.3)
DR (disagreement ratio)	24.9	26.5	24.9	32.9	28.9	27.8	27.0
	25.4			29.9			27.0
ROR (relative optimism ratio)	21.7	20.7	20.4	26.6	24.6	21.3	22.0
	20.9			24.2			22.0

Note: Ratios for each year are reported in parentheses.

In addition, Table 4-4 shows the ROR of each section of SABP for both the first cycle (from 2005 to 2007) and the second cycle (from 2008 to 2010) of SABP. In both cycles the ROR is mainly from the performance planning and results sections. The ROR from the performance planning and results sections over the six years is 34.85% and 38.15% respectively.

Table 4-4 Relative Optimism Ratio of Each Section in SABP

(Unit: %)

SABP section	Planning (Rationale and design)	Planning (Performance planning)	Management	Results	Whole section
Six years	**1.93**	**34.85**	**14.82**	**38.15**	**22.44**
’05-’07	1.88	**33.06**	13.23	**35.21**	**20.85**
’08-’10	1.95	**36.67**	16.49	**41.03**	**24.04**

Note: The ROR between Tables 4-3 and 4-4 may be a little different due to the round-off point.

The ROR for the performance planning section during the first cycle is 33.06%, while during the second cycle it is 36.67%, and the ROR for the results section during the first cycle is 35.21%, while during the second cycle it is 41.03%. Thus, we see that the ROR for both the performance planning and the results sections seems not to have fallen, although SABP has been implemented several times, and therefore the analysis will provide more useful information if it is divided between the first cycle and the second

cycle of SABP. Moreover, from the fact that most of the ROR can be explained by the ROR for the performance planning and results sections, we can also see that it will be more meaningful if the ROR is examined by dividing it into the ROR for the performance planning section, that for the results section and that for the whole section. These divisions can be applied to the analysis of the DR as well. Consequently, dependent variables of the study are "the DR in performance planning, results and the whole section" and "the ROR in performance planning, results and the whole section" based on two time frames, the first and the second cycles of SABP.

Independent Variables and Their Meanings

As mentioned earlier, factors which might affect the DR and ROR can be divided into two categories: factors related internally to differences between spending ministries and the MOSF; and factors related externally. The former can be independent variables, while the latter can be treated as control variables. Taking into account variables used in previous studies, as discussed in Section 3.4, the study sets "SABP score of the MOSF" and "specific programme's self-assessment score minus the spending ministry's average score in that year (specific score minus average score)" as the independent variables.

The first independent variable is the "SABP score of the MOSF". The MOSF reviews spending ministries' self-assessment scores and gives a score from 0 to 100 points while weighting the rationale and design, performance planning, management and results sections at 15, 15, 20 and 50 points respectively. The relationship between this variable and the DR and ROR can determine whether there has indeed been optimism bias by a

spending ministry in the self-assessment stage and whether this has persisted. As for the DR, for example, if spending ministries have given about 90 points to nearly all their programmes in the self-assessment stage, the possibility of disagreement between spending ministries and the MOSF becomes higher as the MOSF's review score becomes lower. However, if spending ministries have self-assessed their programmes strictly and objectively, like the MOSF, then it is not clear whether the DR becomes higher when the MOSF review score becomes lower, because there is a possibility that spending ministries have given low self-assessment scores when the MOSF review score is low. Thus, it can be expected that if there has been a high and persistent optimism bias by spending ministries in the self-assessment stage, the MOSF review score may have a "negative relationship" with the DR. With respect to the relationship between this variable and the ROR, if there has been high and persistent optimism bias by spending ministries in the self-assessment stage, the probability of a "Yes but No" type disagreement becomes higher, and that of a "No but Yes" type disagreement becomes lower, when the MOSF review score becomes lower. Both these situations will lead to increasing the ROR of spending ministries. However, if there has been no high and persistent optimism bias by spending ministries in the self-assessment stage, we cannot be sure whether the ROR becomes higher when the MOSF review score becomes lower. Consequently, it can be expected that if there has been high and consistent optimism bias by spending ministries in the self-assessment stage, then the MOSF review score may have a "negative relationship" with the ROR.

The second independent variable is "specific score minus average score". The relationship between this variable and the DR and ROR can determine whether there

has indeed been drastic cutting by the MOSF of a spending ministry's generous self-assessment score or whether the MOSF has not drastically cut a spending ministry's already tight self-assessment score. That is, if this variable is positive and becomes larger, it means that a spending ministry assesses a specific programme more generously than its other programmes. In this situation, two possibilities can be hypothesised: the first case is that the MOSF drastically cuts the spending ministry's generous self-assessment score; and the second case is that the MOSF does not cut it drastically. In the former case, the DR becomes higher. Also, the ROR becomes higher, because the probability of a "Yes but No" type response increases, while that of a "No but Yes" type response decreases, and both will lead to a higher ROR. However, in the latter case it is not clear whether the DR and ROR become higher or not. Thus, it can be expected that this variable is likely to vary "positively" with both the DR and the ROR when the MOSF is drastically cutting a spending ministry's generous self-assessment results.

On the other hand, if this variable is negative and becomes smaller, it means that the spending ministry assesses a specific programme more tightly than its other programmes. In this situation, we can also hypothesise two possible cases, as above: the first case is that the MOSF does not drastically cut the spending ministry's self-assessment score; and the second case is that the MOSF does drastically cut it. In the former case, the DR becomes lower; and, the ROR also becomes lower, because the possibility of a "Yes but No" type response decreases, while that of a "No but Yes" type response increases, and both will lead the ROR to be lower. However, in the latter case it is not certain whether the DR and ROR become lower or not. As a result, it can be

expected that this variable may have a "positive relationship" with both the DR and the ROR if the MOSF does not drastically cut a tight self-assessment score by the spending ministry.

The one thing which needs to be noted is that spending ministries' self-assessment scores and the MOSF's review scores do not influence each other. For example, the MOSF does not decide its review scores after considering spending ministries' self-assessment scores. In other words, spending ministries and the MOSF decide their scores independently and behave independently. There might, of course, be a correlation between spending ministries' self-assessment scores and the MOSF's review scores, but this is not the result of influence between the two kinds of score. Consequently, "the MOSF review score" and "specific score minus average score" can be used as independent variables for the DR and ROR without any simultaneity bias[16] in the regression analysis.

Control Variables and Their Meanings

Other factors, such as budget changes for the current year, programme type, budget size of programme, type of organisation, purpose of programme and the number of SABP implementations on programmes, can also have an external effect on differences in SABP scores. If differences between the SABP scores of spending ministries and those the MOSF are more likely to be influenced by these factors than by internal factors, then we need to consider the impact of these factors when it comes to interpreting the

[16] Simultaneity arises when one or more of the independent variables are jointly determined with the dependent variable, typically through an equilibrium mechanism.

relationship between independent and dependent variables. Therefore, the analysis model needs to control these factors so that they do not affect interpretation of the relationship between independent and dependent variables in the results of regression analysis later. In particular, the study categorises these variables into three groups for the convenience of analysis: budget-process-related factor (budget changes for the current year), programme-characteristics-related factors (programme type, budget size of programme, type of organisation), and policy/politics-related factors (purpose of programme, the number of SABP implementations on programmes).

(1) Budget-process-related factor: Budget percentage changes for the current year

Both spending ministries and the MOSF consider each programme's budget changes in the self-assessment and review stages of SABP because SABP results affect the programme's budget allocation. Budget percentage changes for the current year can be expressed in the following equation: for example, in the budget year 2010, budget percentage changes are [(2010 programme budgets - 2009 programme budgets) / 2009 programme budgets] * 100. If the budget change for the current year is low, then the spending ministry has a tendency to obtain a larger budget in the following year, and this tendency might lead to increasing disagreement between the spending ministry and the MOSF as well as to "Yes but No" type responses. On the other hand, if the budget change for the current year is low, then the MOSF might regard this programme's feasibility as low, and this tendency might lead to not only increasing disagreement between the spending ministry and the MOSF but also increasing "Yes but No" type responses and decreasing "No but Yes" type responses. Thus, if the budget changes for

the current year are low, then both the DR and the ROR might increase. Consequently, this variable might have negative relationships with both the DR and the ROR.

(2) Programme-characteristics-related factors

① Programme type: Direct and indirect programmes

According to the SABP manual, all budgetary programmes are categorised into eight types: (1) IT system programmes; (2) Social Overhead Capital programmes; (3) Capital Acquisition programmes; (4) Subsidy to Private Sector programmes; (5) Subsidy to Local Government programmes; (6) Loan programmes; (7) Investment programmes; (8) Other Direct programmes (MOSF, 2010a). However, as mentioned earlier, this study employs seven types, excluding IT system programmes, in order to maintain the consistency and comparability of data, because IT system programmes were incorporated into SABP after 2008. Direct programmes managed directly by central government include three types of programmes (Social Overhead Capital, Capital Acquisition, and Other Direct programmes), whilst indirect programmes implemented by other organisations contain four types of programmes (Subsidy to Private Sector, Subsidy to Local government, Loan, and Investment programmes). Programme characteristics such as whether a programme is direct or indirect might also affect the DR and ROR. Spending ministries generally focus on obtaining larger budgets for direct programmes which they can manage directly, through which they can exert more influence on society and interested groups, and which are more related to their own power and interests. On the other hand, indirectly managed programmes, such as those

dealing with subsidies to the private sector and local government, generally have a complex implementation process[17] as well as a high level of difficulty level in their execution, and these programmes are less related to spending ministries' own direct interests. Considering these tendencies of spending ministries, it can be expected that directly managed programmes might have positive relationships with both the DR and the ROR.

② Budget size of programmes: Small and large programmes

Programmes can be grouped according to their budget size and this study uses the concept of quartiles. Kong et al. (2007), Bang (2009) and Cho (2010) also divided budgetary programmes into four quartiles based on the median. This study divides programmes into four groups on the basis of programme budgets for the budget year t and the median is 8.3bn Won[18]. Programmes in the first quartile are less than 2.4bn Won, those in the second are from 2.4 to 8.3bn Won, those in the third are from 8.3 to 32.3bn Won, and those in the fourth are over 32.3bn Won. Small programmes are placed in the first quartile whilst large programmes are in the fourth quartile. Generally, spending ministries pay more attention to programmes with large budgets than to those with small ones, because large budget size programmes are in many cases closely related to important policies, so high level government officials tend to be more concerned with them. Moreover, spending ministries' middle and low level officials

[17] General process of indirect programme: announcement → selection of suitable trustees → decision about matching ratio of funds → confirmation of specific programme plan → budget distribution to trustees → implementation; general process of direct programme: announcement → selection of trustees → budget distribution → implementation

[18] The Won is the unit of Korean money. 1,000 Korean Won is equivalent to nearly 0.60 UK pounds, as of February 2013.

worry about being regarded as incompetent if they fail to obtain large budgets for large programmes, because the MOSF tends to secure budgets for the presidential agenda, and for programmes the ministers and the National Assembly are interested in, by cutting from programmes with large budgets rather than from many small programmes. Also, in order to change budget structures more efficiently, the MOSF focuses on adjusting programmes with a large budget, not programmes with small ones. Given this background, it might be expected that programmes with small budgets might have negative relationships with both the DR and the ROR, whilst programmes with large budgets might have positive relationships with both the DR and the ROR.

③ Type of organisation: Ministry and agency programmes

The type of organisation by which programmes are implemented might also affect the DR and ROR. In general ministry programmes are more policy-oriented, complex and important than those of agencies. Agencies' programmes are simpler and more process-oriented. Thus, it is possible that the DR and ROR might be higher in ministry programmes than in those of agencies, and it might be expected that ministry programmes would have positive relationships with both the DR and the ROR.

(3) Policy/Politics-related factors

① Purpose of programme: Economy-related and welfare-related programmes

Economy-related programmes (Group 1) are operationalised by economy-related

ministries' programmes and social welfare-related programmes (Group 2) mainly by welfare-related ministries' programmes. As regards the classification of government expenditures, there is the UN's classification of the functions of government (COFOG), based on government expenditure purposes: general public services, defence, public order and safety, economic affairs, environmental protection, housing and community amenities, recreation, culture and religion, health, education, and social protection. Although economic or welfare expenditures are generally classified according to the UN's COFOG, this study also regards the related ministries' programmes as one of these two types, because classifying 2,920 programmes into one of two kinds of programme on the basis of COFOG is almost impossible (Bang, 2009; Cho, 2010). With respect to Group 1, there are programmes belonging to eight ministries for economic affairs (the Ministry of Strategy and Finance, the Ministry of Science and Technology, the Ministry of Agriculture, the Ministry of Commerce and Energy, the Ministry of Information and Communication, the Ministry of Construction and Transportation, the Ministry of Maritime Affairs, and the Small and Medium Business Administration) (Bang, 2009; Cho, 2010). Group 2 includes the programmes of four ministries for health and social protection (the Ministry of Health and Welfare, the Ministry of Labour, the Ministry for Gender Equality, and the Commission for Youth) (Bang, 2009; Cho, 2010). Gilmore and Lewis (2006) also group programmes on the basis of "what department is responsible for the programme" to analyse the impact of the political content of programmes on budget allocation. The budgets of both economy-related and welfare-related programmes have been influenced by the policies or politics of administrations. The Korean Government has concentrated on developing its economy for several decades, so a greater share of the budget has been spent on

economy-related programmes. However, after the Asian financial crisis in the late 1990s the Korean Government realised the necessity for building a stronger social safety net. Since then, welfare-related programmes' budgets have for many years increased at a faster rate than that of budgets as a whole, and this trend has been particularly strengthened since 2003, putting an emphasis on the construction of social welfare programmes. Consequently, during the period the study focuses on, welfare-related programmes' budgets have increased continuously, while economy-related programmes' budgets have decreased. Considering these structural reforms of public expenditures, welfare-related programmes might be less affected by SABP results than economy-related programmes. This means that the DR and ROR in welfare-related programmes might be low, whereas those in economy-related programmes might be high. As a result, it might be said that economy-related programmes have positive relationships with both the DR and the ROR, whilst welfare-related programmes have negative relationships with both the DR and the ROR.

② Number of SABP implementations on programmes

Another possible control variable related to the policy or politics of administrations is the number of SABP implementations on programmes. From 2005 to 2010, of 2,920 programmes, 2,451 were evaluated by SABP once, 425 programmes twice, and 44 programmes three times. The distribution of programmes evaluated by SABP two or more times is shown in Table 4-5. Among 469 programmes, the number of programmes evaluated two or more times in the first cycle of SABP is 51 whilst the number in the second cycle is 418. This means that most of the programmes were assessed by SABP

once in the first cycle (the proportion of programmes assessed by SABP two or more times is just 2.4%) while a much greater number of programmes were assessed two or more times in the second cycle (34.7%). Considering the learning effect on spending ministries, it might be expected that if the number of times SABP is implemented increases, then the DR and ROR might become lower. As a result, it might be said that the number of SABP implementations has a negative relationships with both the DR and the ROR.

Table 4-5 Distribution of Programmes Assessed by SABP Two or More Times

(Unit: number of programmes)

Number of times SABP implemented		05	06	07	08	09	10	Total
		The first cycle			The second cycle			
Twice	1st	165	136	113	9	2		425
	2nd			31	144	119	131	425
Three times	1st	18	18	8				44
	2nd			20	13	11		44
	3rd				3	6	35	44

4.4.3 Constructing a Model: Multiple Regression Analysis

An analysis framework to verify the above relationships can be expressed with the multiple regression model shown in Table 4-6 below. While the analysis period is the six years from 2005 to 2010, the model uses two time frames – the first cycle (2005-2007) and the second cycle (2008-2010) – in order to compare the findings of the two cycles. Dependent variables are represented by Y_{ijt} (i = 1 is DR, i = 2 is ROR; j = 1 is the performance planning section, j = 2 is the results section, j = 3 is the whole section; t = 1 is the first SBAP cycle, t = 2 is the second SABP cycle, t = 3 is the six-year period). “SABP score of the MOSF” and “specific programme’s self-assessment score minus the

spending ministry's average score in that year", the independent variables, are represented by X_{1t} and X_{2t} respectively. A control variable which has ratio scale measurement is expressed as X_{3t} (budget percentage changes for the current year) and control variables which have nominal scale measurement are indicated by dummy variables from D_1 to D_7. Dummy variables are coded into (1,0) by using the operationalised concepts. The programme type variable (X_4) is programmes directly managed by government (D_1); the programme budget size variable (X_5) is programmes with small budgets (D_2) and those with large budgets (D_3); the characteristics of the organisation (X_6) variable is programmes belonging to a ministry (D_4); the purpose of programme variable (X_7) is programmes related to the economy (D_5) and programmes related to social welfare (D_6); and the number of SABP implementations variable (X_8) is programmes assessed by SABP two or more times (D_7). α is the constant, meaning the intercept of the Y axis; and β is the unstandardised coefficients in the multiple linear regression equation. ε denotes an error term.

Table 4-6 Analytical Model for Verifying the Factors which Affect the DR and ROR

$$Y_{ijt} = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + (\beta_3 X_{3t} + \beta_4 D_1 + \beta_5 D_2 + \beta_6 D_3 + \beta_7 D_4 + \beta_8 D_5 + \beta_9 D_6 + \beta_{10} D_7) + \varepsilon$$

<Independent variables>

- X_{1t}: SABP score of the MOSF at t
- X_{2t}: Specific programme's self-assessment score minus the spending ministry's average score at t (specific score minus average score)

<Control variables>

- X_{3t}: Budget percentage changes in the current year
- Programme type (X_4):
  (D_1) programmes managed directly by government (direct programmes 1, others 0)
- Budget size of programme (X_5):
  (D_2) small programmes (small programmes 1, others 0)
  (D_3) large programmes (large programmes 1, others 0)
- Characteristics of organisation (X_6)
  (D_4) programmes belonging to a ministry (ministry programmes 1, others 0)

■ Purpose of programme (X_7)
(D_5) programmes related to the economy (economy-related programmes 1, others 0)
(D_6) programmes related to social welfare (welfare-related programmes 1, others 0)
■ The number of SABP implementations (X_8)
(D_7) programmes assessed by SABP two or more times (two or more times 1, others 0)

<Dependent variables>
■ Y_{ijt}: Differences in SABP score between spending ministries and the MOSF
(i = 1 is DR, i = 2 is ROR; j = 1 is performance planning section, j = 2 is results section, j = 3 is whole section; t = 1 is the first SBAP cycle, t = 2 is the second SABP cycle, t = 3 is six years)

<Others>
■ α and β are the constant and unstandardized coefficients respectively in multiple linear regression equation
■ ε denotes an error term

The grounds for variables and their scales can be seen in Table 4-7 below. Dependent variables use a ratio scale because they are "the disagreement ratio between spending ministries and the MOSF" and "the relative optimism ratio of spending ministries". Two independent variables are expressed with SABP scores, so they also use a ratio scale. Among eight control variables, budget percentage change for the current year is represented by a ratio scale while the other seven control variables use nominal scales as dummy variables.

Table 4-7 Grounds for Variables and their Scale of Measurement

Category		Name of variable	Scale	Coding	Ground
Independent variable		- (X_{1t}) SABP score of the MOSF - (X_{2t}) Specific programme's score minus average score of the spending ministry in that year	ratio ratio	score score gap	Gilmour & Lewis (2005, 2006a), Park (2005, 2008), Bang (2009), Cho (2010), Chang&Yoon (2002), Kong
Control variable	Budget process	- (X_{3t}) Budget percentage changes	ratio	ratio	
	Programme characteristics	- (D_1) Direct programmes - (D_2) Small programmes - (D_3) Large programmes - (D_4) Ministry programmes	nominal	1,0	

	Policy related	- (D_5) Economy related - (D_6) Welfare related - (D_7) SABP two or more times	nominal	1,0	et al. (2007), Yoon (2001)
Dependent variable		- (Y_{ijt}) DR and ROR	ratio	-	

4.4.4 Data Collection

The data employed in this study are the SABP results of 2,920 programmes for 49 central government departments and agencies from 2005 and 2010. Since a third of all the programmes of each ministry (including agencies) are assessed by SABP, SABP has a three-year cycle. Thus, the data from 2005 to 2010 include two SABP cycles. The study collected SABP results data through the websites of the MOSF, the KIPF, spending ministries and the National Assembly. In particular, the MOSF, the controller and designer of the SABP system, has published SABP data since 2005. Therefore, the MOSF and other organisations' websites all initially contributed to collecting the SABP results applied in the study. In addition, when it came to analysing specific differences between spending ministries and the MOSF, for example, counting the number of "Yes but No" type and "No but Yes" type responses in each programme, the official data published on the websites were not sufficient, so the study also used internal data from the MOSF with the help of MOSF officials. Data including SABP scores, SABP grades and budgets in each year were recorded using the Excel programme, in order to make it easy to analyse them through econometric analysis. Data related to programme type, programme budget size, characteristics of organisations, purpose of programme and the number of SABP implementation were collected by categorisation of the SABP results or descriptive statistical analysis such as four quartiles based on the median.

4.4.5 Data Analysis

Data analysis was carried out using the following procedures. Firstly, the dependent, independent and control variables relating to the differences between spending ministries and the MOSF were extracted from both the descriptive statistical analysis and the review of relevant literature. Descriptive analysis was employed to compare the trends of SABP results and the differences in SABP results between spending ministries and the MOSF from 2005 to 2010. In addition, factors which might affect the differences in SABP results were ascertained and divided into independent and control variables (e.g., budget-process-related, programme-characteristics-related and policy/politics-related variables) based on the review of previous studies. Secondly, an initial framework for analysing the relationship between diverse variables and the differences in SABP results was to be developed by various preliminary analyses. Relationships between independent and dependent variables could be determined by correlation analysis, and relationships between diverse control variables and dependent variables could be certified by t-test. The results of preliminary analyses were to be applied when it comes to constructing the multiple regression analysis. Thirdly, relationships between diverse variables and the differences in SABP results were to be verified in detail through multiple regression analysis employing the SPSS (Statistical Package for Social Science). That is, the predicted relations explained in section 4.4.2 were to be examined through the results of regression analysis. In order to verify the statistical significance of the regression model, F-values verifying that all regression coefficients were not zero and t-values verifying whether each coefficient was substantially different from zero were to be examined. In addition, multicollinearity

among variables, which makes regression analysis impossible, was also to be tested. Then, R square (the variability in the outcome linked to the independent variables) and adjusted R square (the extent to which R square can be generalised) were to be analysed.

4.5 Qualitative Approach

4.5.1 Introduction

The purpose of this section is to address the second and third research questions, and to suggest feasible policy alternatives for preventing or reducing various dysfunctional consequences mentioned in Section 2.3.2 based on the conceptual framework developed in Chapter 2. That is, this section seeks to explore whether dysfunctional consequences by spending ministries actually exist, and if so, what are the extents of them, focusing on differing perspectives between spending ministries and the MOSF. It then looks at the impact of these dysfunctional consequences on the behaviours of both spending ministries and the MOSF, and explores feasible policy implications for a better SABP system. In order to do this, firstly, the study employs multiple case study as a qualitative approach to investigate and understand more thoroughly a particular contemporary phenomenon within its real life context (Yin, 2009). Secondly, the study conducts documentary analysis, archival records analysis and particularly in-depth interviews in order to collect data from multiple sources of evidence. Thirdly, for qualitative data analysis, Nvivo 9 is utilised during the process of capturing, coding and reporting the findings of the case study. The rest of this section will describe a multiple case study as

a research strategy, the selection of case study areas, a research process for data collection, and data analysis.

4.5.2 Multiple Case Study as a Research Strategy

There are many research strategies that can be used to collect and analyse data (de Vaus, 2001): experimental, longitudinal, cross-sectional, comparative, action research and case study. Among these research strategies, the case study is a suitable approach for finding idiographic as well as nomothetic explanations (de Vaus, 2001) and dealing with a set of events which investigators will rarely be able to control (Yin, 2009). The present study selected the case study as a research strategy not only because it would enable me to understand more fully a contemporary phenomenon in its real-life setting but also because investigating the existence of dysfunctional consequences of SABP and the impact on the SABP results of both spending ministries and the MOSF involves very sensitive and complex issues which I could not control (Yin, 2009; Benbasat et al, 1987). The case study could also be expected to give a good opportunity to understand the holistic and meaningful natures of the complex behaviours of spending ministries and the MOSF in the Korean context (Yin, 2009; de Vaus, 2001; Bryman, 2008; Grix, 2010). Furthermore, the study needed a good number of case studies under different conditions in order to investigate how consistently its findings held up (Yin, 2009), because exploring dysfunctional consequences of SABP includes different types of programmes (e.g. social overhead capital, investment, loan, subsidy to private sector or local government), diverse spending ministries, and various government officials at different levels (e.g. low, middle and high). De Vaus (2001, p. 226) argues that “given sufficient resources and access to cases, multiple case study will normally be more

powerful and convincing and provide more insights than single case study." In line with Yin (2009) and de Vaus (2001) this research sought to test and develop more compellingly, rigorously, and robustly the conceptual framework (Herriott and Firestone, 1983), and to suggest well-designed policy alternatives to refine the present SABP system. Through the use of multiple cases of evidence, the study could strengthen construct validity and external validity (Cook and Campbell, 1979). To enhance reliability, the study developed a specific case study database with all relevant data such as interview transcripts, research memos, final SABP reports, and evaluation reports of the National Assembly and the BAI (Yin, 2009).

As for the time dimension of the study, like most case studies and explanatory case studies the study incorporates a time dimension, particularly a retrospective design, because it collects information relating to an extended period for use on one occasion by drawing on qualitative research methods including archival records and document analysis and interviews with government officials who took part in or observed past events (de Vaus, 2001). With limitations on time and other resources, except for people's ability to recall the past, adopting a retrospective dimension is appropriate, rather than tracking changes forward through time. The important thing, as de Vaus (2001: 228) argues, is "to build up a clear and reasonably detailed picture of the sequence in which events took place and of the context in which they occurred."

4.5.3 The Selection of Three Case Study Areas and Eight Programmes

The case study puts more emphasis on theoretical generalisation, which involves strategic selection of cases, than on statistical generalisation, which depends on the representativeness of cases (de Vaus, 2001). Walsham (1993, p.15) argues that the external validity of case study work can be achieved by "the plausibility and cogency of the logical reasoning in describing the results from the cases and in drawing conclusions from them rather than the representativeness of cases in a statistical sense." To enhance the external validity of the case study this study employed three case areas and eight specific budgetary programmes within them.

The study used a number of criteria and carried out extensive screening in order to choose case study areas. Firstly, case study areas were chosen because of the likelihood that they would display some indications of dysfunctional consequences of SABP, because the main interest of the study was to explore the dysfunctional consequences of SABP and to examine the impact of these on the SABP results of both spending ministries and the MOSF. Secondly, areas of high political significance were selected, which was closely related to the first criterion, because the higher the programmes' political importance, the greater the expected optimism bias by spending ministries and the tighter the review of the MOSF might be. Therefore, the possibility of discovering dysfunctional consequences of SABP might be higher. Thirdly, the case study areas were selected to reflect the different conditions under which SABP operated, such as different types of programmes, diverse spending ministries, different levels of employees and programmes with different sizes of budget, in order to obtain the

plausibility and cogency of logical reasoning under different conditions. Fourthly, the time period for the case study areas was the second SABP cycle, from 2008 to 2010, since many of the programmes from the first SABP cycle (2005 to 2007) ceased after that period, due to the large amount of restructuring of programmes that took place in the government reorganisation of early 2008 after the presidential election.

On the basis of these criteria for the strategic selection of the case study areas, I firstly examined what the focus fields of each year's budgets had been from 2006 to 2011[19], through document analysis, and the results are shown in Table 4-8. I then chose three case study areas from the common core fields for the six years, and finally I selected eight budgetary programmes within three case study areas.

Table 4- 8 Focus Fields of Budgets from 2006 to 2011

'06	'07	'08
- **Future growth expansion field,** such as SOC, R&D, education - **Job creation** - **Support for classes** which are suffering from polarisation - Ensuring national security	- **Future growth expansion field** - **Satisfaction of basic needs** of citizens such as social welfare, education, etc - **Social job expansion** - Ensuring national security	- **Future growth expansion field** - **Preparation for low fertility and population aging in Korea** - **Social job expansion** - Balanced national development (including expansion of local public finance) - Reducing the national debt
'09	'10	'11
- **Job Creation field** - **Future growth expansion field** - **Stabilisation of ordinary people's lives** - Small and efficient government	- Recovery of economic vitality and **future growth expansion** - **Stabilisation of ordinary people's lives and job creation** - Maintaining law and order - Upgrading the national image	- **Stabilisation of ordinary people's lives** - **Future growth expansion (including job creation)** - Boosting local economies and maintaining public order - Reducing the national debt

[19] The study looked at SABP results from 2005 to 2010, and as results were reflected in the budgets of programmes in the following year, I examined the budgets from 2006 to 2011.

Source: MOSF press release materials, National Assembly Budget Office Report (2006-2011), and Guidance for making a budget for each year

Looking at Table 4-8, we can see that the Korean Government's focus fields could be put into the following four categories (core areas in parentheses).

- First field: Stabilisation of ordinary people's lives (housing, pensions, healthcare)
- Second field: Job creation (youth employment)
- Third field: Future productive capacity expansion (SOC, R&D, education)
- Fourth field: Other (maintaining public order, ensuring national security, and boosting local economies)

The study chose "the stabilisation of ordinary people's lives", "job creation", and "future productive capacity expansion" as the focus fields of the Korean Government for each year. Core areas in the stabilisation of ordinary people's lives were public housing, pensions and healthcare. With respect to job creation, youth employment was seen as the core area, because although the youth employment problem has been consistently treated as one of the hottest issues in most OECD countries, including Korea, for many years, most of these countries' programmes have not yet had significant results. Regarding the expansion of future productive capacity, SOC, R&D, and education would be the core areas. From these seven important areas within three fields, three case study areas of interest for the study – public housing, youth employment and SOC – could be drawn, for the following reasons.

Regarding the pensions and healthcare areas in the first field, the budgets of pensions and healthcare programmes were often decided by rules laid down by the relevant laws, regardless of their SABP results. So, it seemed meaningless to link SABP results with such mandatory expenditures. Thus, in the first field the area of public housing was chosen. In the second field, the area of youth employment area was selected, because as mentioned earlier youth employment has been the core area of the job creation field in many countries for a long time. With respect to the third field, R&D programmes' performances have since 2011 been assessed by another evaluation system run by the NSTC, because of the special nature of research and development expenditures. In addition, most of the education expenditures stemmed from transfers from central to local government, and this kind of internal transfer within governments has been excluded from SABP. For example, in 2010 the total amount of the budget in the education field was about 38,300bn Won, but 32,300bn Won (84.3%) consisted of transfers from central to local government. Thus the R&D and education areas were not appropriate case study areas, and so the remaining important area, SOC, was selected as a case study area.

After choosing three case study areas, eight specific programmes which had high political importance were strategically selected on the following three criteria: whether a programme was on the presidential agenda; whether it had a high priority among spending ministries' programmes; whether the National Assembly, the BAI or the NGO had an interest in it. In the public housing area, three programmes related to public housing were chosen: construction of public rental housing, subsidies for housing rental costs, and improvement of old public housing stock. Expansion of public housing was

one of the key projects of the Presidential Transition Committee (December 2007 to February 2008). The Presidential Transition committee suggested the expansion of public housing as one of 43 core presidential agendas in February 2008 (Presidential Transition Committee, 2008) and this was included officially as one of the 100 presidential agendas after the new administration was inaugurated. Subsequently, the Ministry of Land, Transport and Maritime Affairs (MLTM) published the "1.5 million Bokeumjari public housing construction plan for 2009 to 2018" in September 2008 and presented this to the President in December 2008 as one of the core projects for 2009 (Yeonhapnews, 2008). In 2011, the Minister of the MLTM stated in an interview that the construction of 1.5million units of public housing would be continued until 2018, as planned (Primenews, 2011). In addition, the MLTM announced a comprehensive public housing policy every year (MLTM, 2011, 2010). Moreover, the National Assembly Research Service included the programme for the construction of public housing as one of the government activities that would be subject to inspection (NARS, 2011), and the BAI launched a comprehensive audit and inspection of the public housing construction programme, raising concerns over the supply and demand situation in public housing (BAI, 2011). The construction of public rental housing, subsidies for housing rental costs, and the improvement of old public housing stock have been key programmes in the public housing field. As regards budget size, the 2008 budgets for construction of public rental housing and subsidies for housing rental costs programmes, at 3,400bn Won and 3,100bn Won respectively, were the largest and second largest of all the MLTM's programmes. It can therefore be stated that the priority of these programmes within the ministry is very high.

With respect to youth employment, it seems unnecessary to point out that budgetary programmes related to youth employment have high political importance, but the problem was how to choose suitable programmes for the research from 36 youth employment programmes in 12 ministries and agencies. The National Assembly Budget Office (NABO) (NABO, 2010a) put the 36 youth employment programmes into five categories using the OECD's classification of employment [20] drawn up to analyse countries comparatively: on-the-job training (eight programmes); experience, international training and internship (14 programmes); employment incentives (four programmes); direct job creation (eight programmes); and public employment services and administration (two programmes). As a result of an Analytic Hierarchy Process (AHP) by 20 labour specialists, the NABO gave priority to five types. The specialist group ranked the importance of the five types in the following order: on-the-job training → experience, international training and internship → employment incentives → public employment services and administration → direct job creation. In addition, the second of these types was closely related to the presidential agenda of producing 100,000 global youth leaders. Considering the choices of both AHP and the presidential agenda, of the five types, experience, international training and internship (the second type) had higher political importance than the other types and was more appropriate for the study. After careful programme screening of this type I found that six of the 14 programmes were assessed by SABP from 2008 to 2010. The six programmes in the second group were: youth job experience ('08), junior college students' overseas internship ('10), global youth leadership training ('10), global trade specialist training ('10), global plant

[20] The OECD (2009) put labour market policy into nine categories: public employment services and administration, training, job rotation and job sharing, employment incentives, supported employment and rehabilitation, direct job creation, start-up incentives, out-of-work income maintenance, early retirement.

specialist training (’10), and global exhibition specialist training (’10). However, global plant specialist training and global exhibition specialist training programmes were included in the global trade specialist training programme, because the purposes of these programmes were very similar and the Ministry of Knowledge and Economy (MKE) implemented all three programmes at the same time. Finally, considering the availability of evidence and interviewees, junior college students’ overseas internship (’10), global youth leadership training (’10) and global trade specialist training (’10) programmes were chosen for the case study in the youth employment area.

In the SOC area, national highway and railway construction were the main programmes, and they had high political importance. The SOC area is divided into four large categories: national highway, railway, airport, and port construction. The budgets for national highway and railway construction take up the majority of the SOC area’s funding (13,400bn Won out of 25,000bn Won, or 53.2%). The SOC programmes were all included in the presidential agenda, but the national highway and railway construction programmes were also selected as two of 30 major pioneering projects for boosting the metropolitan economy which were published in September 2008. Consequently, the national highway and railway construction programmes had higher priority among the MLTM programmes and were more appropriate for the case study. As a result, the study chose eight programmes within three case study areas based on the various criteria mentioned above: in the area of public housing, construction of public rental housing (P1), subsidy for housing rental costs (P2), and improvement of old public housing stock (P3); in the area of youth employment, global youth leadership training (P4), global trade specialist training (P5), and junior college students’ overseas

internships (P6); and in the SOC area, road construction (P7), and railway construction (P8).

Descriptions of the Eight Programmes

This section gives a brief explanation of the eight budgetary programmes in three case study areas. The three programmes in the public housing area all have characteristics of high public interests as well as political importance, and central-government-initiated construction has long been accepted as a desirable social safety net for solving the housing problems of the poor. The public housing construction programme (P1) was started in 1998 and a reasonable estimation of supply and demand has been developed since 2003 through the MLTM's comprehensive housing plan. The subsidy for housing rental costs programme (P2) has lent money to poor tenants at 70% lower interest rates than the market level since 1990. The programme for improvement of old public housing stock (P3) was started in 2009 in order to enhance the housing environment for the poor and boost the local economy. The performance indicator for P1 is "public rental housing construction rate"; for P2 it is "budget implementation rate"; and for P3 it is "rate of budget grants to local government". These are all indirect programmes (loan, loan, and subsidy to local government respectively) and although the MLTM self-assessed them as effective programmes, the MOSF reviewed them as modest programmes. Interestingly, on Question 2-1 for P1 and P2 the MOSF first decided "No", but finally changed its decision to "Yes"[21]. The specific reason will be explored and

[21] From 2005 to 2008, the Budget Office of the MOSF firstly reviewed spending ministries' self-assessment results, and then the Performance Management Bureau of the MOSF gave them a final assessment. However, from 2009 the Budget Office of the MOSF has not reviewed these

discussed in Chapters 6 and 7. The main features of the three programmes are summarised in the Table 4-9 below.

Table 4- 9 Main Features of the Three Programmes in Public Housing

Key features		P1	P2	P3
Performance indicator		Construction rate	Budget implementation rate	Rate of budget grants
Programme period		1998 -	1990 -	2009-2013
Programme type		Indirect (loan)	Indirect (loan)	Indirect (subsidy)
Budget size (bn Won)		3,422	3,066	50
SABP implementation		2008	2008	2010
Spending ministry's self-assessment (MLTM)	Total score	95.0 (effective)	91.7 (effective)	100.0 (effective)
	Q2-1	Yes	Yes	Yes
	Q2-2	Yes	Yes	Yes
	Q4-1	Yes	Yes	Yes
	Q4-2	Yes	Yes	Yes
MOSF's review	Total score	64.2 (modest)	60.8 (modest)	62.7 (modest)
	Q2-1	No → Yes	No → Yes	Yes
	Q2-2	No	No	No
	Q4-1	To a small extent	To a small extent	To a small extent
	Q4-2	No	No	N/A (new programme)

Note: Q2-1 is related to the appropriateness of the performance indicator; Q2-2 is related to the relevance of the performance target; Q4-1 is related to the achievement of the planned performance target; and Q4-2 is related to the implementation of comprehensive and objective programme evaluation including both quantitative and qualitative analysis

The three programmes in the youth employment area have been the focus of a great deal of attention by the administration, the National Assembly and various new media, because although youth employment has been a very important problem, it has not yet been solved. Thus, as mentioned above, the administration set up a project to recruit and train 100,000 global youth leaders as one of the presidential agendas, and the three programmes were all included in this project. The global youth leadership training programme (P4) was started in 1994 by the Ministry of Employment and Labour

results but just given some opinions when the Performance Management Bureau has finally decided the results.

(MOEL) in order to increase foreign employment for young people by giving them the opportunity to train abroad for three to ten months. The global trade specialist training programme (P5) has been implemented since 2007 by the MKE in order to give an opportunity for learning about foreign plants and global trading skills to four-year university students. In addition, the Ministry of Educations, Science and Technology (MEST) has implemented a junior college students' overseas internships programme (P6) since 2005 to enhance the employment rate of two-year university students by providing an opportunity for them to have foreign internship experience for about six months. In particular, the global youth leadership training programme (P4) has been highlighted, so the programme's performance indicators and performance targets were decided by several cabinet meetings with related governments. As with the three programmes in the public housing area, the relevant spending ministries all self-assessed these programmes as effective, but the MOSF reviewed P4 and P6 as modest programmes, while it found P5 was a poor programme. Interestingly, in Q2-1 for P5, the spending ministry answered "No" although it might know well that Q2-1 was logically linked with Q2-2 and Q4-1, and that if Q2-1 was "No", it might be difficult to avoid a poor grade finally. Also, in Q2-1 for P6, the MOSF gave a "Yes" answer on condition that the spending ministry would develop more appropriate performance indicators later. This will be explored in Chapters 6 and 7. These differences in SABP results between spending ministries and the MOSF imply that there might be dysfunctional consequences in the SABP system and these might impact on the behaviours of both spending ministries and the MOSF. The main features of the three programmes are summarised in Table 4-10 below.

Table 4-10 Main Features of Three Programmes in Youth Employment

Key features		P4	P5	P6
Performance indicator		The number of students who completed programme, Employment rate	The number of collaborating companies, Accretion rate of academic credit, Matching fund rate	Improvement rate of language skills, Degree of student satisfaction
Programme period		1994 -	2007 -	2005 -
Programme type		Indirect (investment)	Indirect (subsidy)	Indirect (subsidy)
Budget size (bn Won)		25	9	5
SABP implementation		2010	2010	2010
Spending ministry's self-assessment	Total score	87.5 (effective)	87.5 (effective)	100.0 (effective)
	Q2-1	Yes	No	Yes
	Q2-2	Yes	Yes	Yes
	Q4-1	To a large extent	Yes	Yes
	Q4-2	Yes	Yes	Yes
MOSF's review	Total score	64.3 (modest)	49.0 (poor)	56.5 (modest)
	Q2-1	Yes	No	Yes (conditionally)
	Q2-2	Yes	No	No
	Q4-1	To a large extent	To a small extent	To a small extent
	Q4-2	No	No	No

The road and railway construction programmes in the SOC area have always been highlighted by interested groups because these programmes have large budgets and are directly related to the local economy. The road construction programme (P7) was started in 1999 by the MLTM in order to link major cities, ports and tourist attractions as well as to reduce congestion and distribution costs. The MLTM added outcome-oriented performance indicators (traffic capacity) as well as two output indicators (budget implementation rate and road extension rate). The MOSF recognised the MLTM's efforts, so Q2-1 was marked "Yes" by both spending ministry and the MOSF. The railway construction programme (P8) was initiated in 1993, having a similar purpose to the road construction programme. Interestingly, on Q2-1 for P8 the MOSF gave a "Yes" on condition that the MLTM would add an outcome-oriented performance indicator later. In addition, on Q4-2 for both P7 and P8 the MOSF gave a "Yes", which

was different from their verdict on the other six programmes, because the two programmes implemented comprehensive and objective programme evaluations, including both quantitative and qualitative assessment, to follow the relevant law, which stipulated compulsory programme evaluation for programmes whose budget was over 50bn Won. These will be explored in Chapter 6. As with the other programmes, the spending ministry self-assessed the two programmes as effective while the MOSF reviewed them as modest. The main features of the two programmes are summarised in the Table 4-11 below.

Table 4-11 Main Features of Two Programmes in SOC

Key features		P7	P8
Performance indicator		Budget implementation rate, Road extension rate, Traffic capacity	Double-tract rate
Programme period		1999 -	1993 -
Programme type		Direct (SOC)	Direct (SOC)
Budget size (bn Won)		743	1,687
SABP implementation		2010	2010
Spending ministry's self-assessment	Total score	100.0 (effective)	100.0 (effective)
	Q2-1	Yes	Yes
	Q2-2	Yes	Yes
	Q4-1	Yes	Yes
	Q4-2	Yes	Yes
MOSF's review	Total score	60.5 (modest)	65.5 (modest)
	Q2-1	Yes	Yes (conditionally)
	Q2-2	No	No
	Q4-1	To a small extent	To a small extent
	Q4-2	Yes	Yes

4.5.4 Research Methods

After choosing the case study as a research strategy, the next consideration was how to collect data for it. Data can generally be collected from a variety of sources such as

documentation, archival records, interviews, and direct and participant observations (Yin, 2009; Becker and Bryman, 2004). Firstly, documentary information and archival records are helpful to most case studies, because they can give stable, unobtrusive, exact, precise and broad coverage for specific information. However, they can be produced for specific purposes and may have an unknown bias related to their author, so, this study used them as corroborating information and for further investigation rather than as definitive findings (Yin, 2009). Examples of documentation included are (Yin, 2009, p.103):

- Letters, memoranda, e-mails, and other personal documents such as notes
- Agendas, announcements and minutes of meetings and other written reports
- Administrative documents – proposals, progress reports, and other internal records
- Formal studies or evaluations of the same "case" that the researcher is studying
- News clippings and other articles appearing in the mass media or in community newspapers

Also, archival records include the following (Yin, 2009, p.105):

- Public use files such as census and other statistical data
- Service records, such as those showing the number of clients served over a given period of time
- Organisational records, such as budget or personnel records
- Maps and charts of the geographical characteristics of a place
- Survey data, such as data previously collected about a site's employees, residents, or participants

Secondly, interviews are essential as one of the most important data sources for a case study, because a case study generally seeks to explore actors' real perceptions or intentions. Even when documents include actors' perceptions or intentions, they sometimes hide real perceptions or intentions and present them in official rather than real ways. In this sense, interview data can complement documentation and archival

records by focusing directly on case study topics and providing more insightful data for perceived causal inferences and explanations (Yin, 2009). Interviews are generally classified into three types: structured, unstructured and semi-structured (Arksey, 2004; Yin, 2009)[22]. Unstructured and semi-structured interviews have "more willingness to allow interviewees to use their own words and develop their own thoughts" than structured interviews, which use a more structured set of questions (Denscombe, 2003, p.167). Unstructured interviews are more interactive and more responsive to the language and concepts used by respondents, so they are usually used in very broadly defined agenda, initial pre-piloting stages, and life-story and narrative focus (Gillham, 2005). In semi-structured interviews, a researcher is enabled to compare data from actors (e.g., between cases, individuals or groups) through more structured questions, as well as to get actors' feelings, perceptions, interpretations and strategies on the situation by asking open questions with prompts (Namkoong, 2010). Semi-structured interviews are used when the focus of the interview is a defined and significant development phase, so the questions need to be focused and well-ordered (Gillham, 2005).

Thirdly, direct and participant observations can be a useful source for a case study, especially when a researcher is investigating real-time events. A researcher can observe actors' behaviours directly or participate in the events that are the focus of the case study. These methods can cover the context of "cases" effectively, despite being time-consuming, risking possible bias from the observer's manipulation of facts, and the fact that they can be expensive (Yin, 2009).

[22] Elite interviews and group or focus group interviews are regarded as other types of interview (Gillham, 2005).Elite interviews are used in topic areas researchers do not know well and in parts of significant networks. Group interviews can be used when broad and exploratory investigation is necessary and focus group interviews are narrower and more specific than general group interviews.

This study collected information about dysfunctional consequences of SABP from documentation, archival records and in-depth semi-structured interviews. Other methods, particularly observation, seemed inappropriate, because the study sought to investigate events which had already happened, not events in real time. In addition, the study applied to data collection three principles suggested by Yin (2009, pp.114-124). Firstly, the study used multiple sources of evidence, such as documentation, archival records and interview data, in order to make the outcomes more convincing, persuasive and accurate, because "multiple sources of evidence essentially provide multiple measures of the same phenomenon". Secondly, the study created a case database with all relevant data (e.g., reports by the National Assembly, reports by the BAI, research memos and interview transcripts) in order to increase the reliability of the case study (Yin, 2009; Miles and Huberman, 1994). Thirdly, the study made an attempt to develop a chain of evidence (e.g., case study questions ↔ case study database ↔ case study report) to enhance the reliability of the information in the case study. An external observer can follow the evidence and trace the steps from initial research questions to ultimate conclusions or from conclusions to research questions.

4.5.5 Data Collection

After a conceptual framework for exploring dysfunctional consequences of SABP was developed through an extensive literature review, eight budgetary programmes within three case study areas were selected in July 2011; and then, detailed documentary analysis and pilot interviews were employed in order to increase the validity of, and

minimise the number of poorly structured questions in the interview guides for the field work. These preparations provided a basis for the field work.

(1) Documentary Data

Most of the documentary data was gathered through the websites of central government, the National Assembly, the professional public institutes, and the mass media. In particular, the MOSF's reports on the SABP results of the eight budgetary programmes were very useful, because they included different perspectives from spending ministries and the MOSF, and implied that dysfunctional consequences actually existed, and might be extensive. In addition, the NABO's annual analysis of the government's performance reports and the BAI's audit and inspection reports directly indicated some types of dysfunctional consequences of SABP in the eight programmes, providing useful information from a neutral position. In order to avoid "biased selectivity and reporting bias" (Yin, 2009, p.102), I analysed many documents, questioning their form and functions, and mainly used them for corroborating information and further investigation, not to provide definitive findings. Major documentation and archival records which were analysed in the study are shown in Table 4-12 below.

Table 4- 12 Major Documentary Data Sources

Category	Major documents collected and analysed	
Data bases	Digital budget information system (for budget data from 2005 to 2010)	
Documents and archives	Official governmental documents	Official documents of the MOSF, KIPF, NABO, BAI, MLTM, KDI, OECD, GAO, OMB, etc.
	Internal documents	SABP results, MOSF internal reports, etc.

	Formal studies	Korean Policy Studies Review, Korean Review of Public Administration, Korean Journal of Public Administration, Korean Journal of Policy Analysis and Evaluation, etc.
	Press releases	Yeonhapnews, Joong-ang daily, Primenews, etc.
	Statutes and rules	GPA, NFA, GPRA, GPRAMA

(2) Interview Data

Semi-structured interviews with key participants were a main source for the case study, because these were appropriate methods to extract useful information from complex and dynamic social situations (Arksey, 2004). The first 38 interviews were carried out face-to-face between 25th October and 15th November 2011 in Korea; and the second 31 supplementary interviews were conducted by internet phone and e-mail from 20th October to 20th November 2012 in order to develop feasible policy alternatives for a better SABP system by combining and verifying the policy alternatives suggested during the first interviews. Thirty eight interviewees were divided into three categories: 19 spending ministry officials, 13 MOSF officials, and six experts. Thirty two government officials who carried out SABP in the eight programmes were chosen, and six experts on SABP were selected to broaden the research base. Among the 32 government officials, four were senior managers (high ranking), 22 were managers (middle ranking), and six were assistant managers (low ranking). As for selecting interviewees, the study employed "purposive sampling", because this method can be useful to identify subgroups and compare results (Miles and Huberman, 1994). The distribution of the first 38 interviewees is shown in Table 4-13 below and a detailed list of interviewees is provided in Appendix 1.

Table 4-13 Distribution of the first 38 Interviewees

Case area	Spending ministry			MOSF			Sub-total	Expert	Total
	Senior manager	Manager	Assistant manager	Senior manager	Manager	Assistant manager			
Public housing	-	6	-	2	1	1	10	6	38
Youth employment	1	2	4	-	5	-	12		
SOC	1	5	-	-	3	1	10		
Total	2	13	4	2	9	2	32		

Note: Among 19 spending ministry officials, six were budget division and 13 were programme division officials. Among 13 MOSF officials, five were Performance management bureau officials, eight were Budget Office officials.

Pilot Interviews

Before conducting the first 38 interviews, five pilot interviews with Korean government officials who were studying at the University of Birmingham and had experience of SABP were carried out in August 2011. Among the five officials, three were from spending ministries, two were MOSF officials, one was an assistant manager and four were managers. Before commencing the pilot interviews, I explained the purpose of the study and how the interview data would be used, and then I elicited the consent of the interviewees. Also, I explained that the interviews would be recorded but re-assured the interviewees that all data would remain confidential and anonymous. During the pilot interviews, I tested the validity and clarity of the questions, identified question redundancy, and calculated the approximate time needed to complete the interview (Bell, 2006). My reflections on the pilot interviews were as follows.

Firstly, most interviewees agreed that the interview questions were generally clear and easily understood. However, one interviewee mentioned that the unique features of

Korean government culture, such as the general one-year rotation of personnel system and ministers' short tenure of office, could lead to difficulty in focusing on the long-term or whole objectives of programmes. Also, some interviewees answered that the strategy of a spending ministry's budget division could be different from that of the spending ministry's programme division. For example, the budget division was more concerned with good relationships with the MOSF than the programme division was, and this might lead it to recommend a highly ambitious performance target level. As a result, I added questions related to the unique features of Korean government culture and to differences in perspectives on SABP between spending ministries' budget divisions and their programme divisions. Secondly, some pilot interviewees said that the interview time was too long, as the interview lasted from two hours to two and half hours. Also, they pointed out some redundant questions, the contents of which were similar to those of other questions. In order to reflect these opinions, I combined or eliminated some questions through careful review of the interview questions. Thus the interview questions were reduced from 22 to 16, and the time required was reduced by about an hour. Thirdly, I realised that different topic guides might be needed in order to review the different perspectives of the spending ministries and the MOSF, as well as those of the experts. Finally, on the researcher's background and role, the possibility of subjectivity was raised because I, as a government official of the MOSF, had experience of reviewing spending ministries' self-assessment results and implementing SABP. Therefore, I recognised the necessity of having interview questions and interview data reviewed by supervisors and PhD colleagues to reduce possible bias. Also, I tried to avoid bias by deliberately considering the opposite view to my opinions. After the pilot interviews, topic guides for spending ministries and the MOSF, as well as for experts,

were completed through several revisions and supervisions. Appendices 2 and 3 describe these topic guides.

First Interviews and Supplementary Interviews

After the pilot interviews in August 2011, I prepared to make contact with 38 interviewees. Preparation for interviews was carried out through four steps: (1) identifying interviewees and confirming the list through a supervision meeting; (2) making calls to obtain consent to an interview and fixing an interview date, time and place; (3) sending an e-mail including interview topic guides and SABP reports; and (4) confirming the interview date, time and place again via telephone. Firstly, I searched for 42 potential interviewees by name, position and telephone number through the MOSF's SABP reports for the eight budgetary programmes, because SABP reports include this information. In a supervision meeting, my supervisors and I discussed the appropriateness of the interviewee list, for example the number of officials from spending ministries and the MOSF, their positions and departments, and my supervisor agreed with the relevance of potential interviewees. Secondly, I started to make calls via internet phone to potential interviewees to obtain consent to interviews. During this period, I explained the background of the study, the detailed purpose of the interviews, and how interview data would eventually be used. My experience as a MOSF official was very helpful when it came to communicating with interviewees and building mutual trust between interviewer and interviewee. In particular, most of the potential interviewees from spending ministries were happy to agree to the interview, although they needed to say they had somewhat different perspectives on SABP from those of the

MOSF. Thirty two government officials and six experts finally agreed to be interviewed. Thirdly, the interview topic guides and the MOSF's SABP reports on related programmes were sent by e-mail to all interviewees in early October 2011. Interview topic guides included a consent form explaining the purpose of the interview, the way in which interviewees' privacy would be protected, the right to withdraw at any time, and information about the researcher, such as position and contact number. The MOSF's SABP reports on the eight programmes were also sent to interviewees in order to refresh their memory of SABP results in the past. Finally, before departing for Seoul, I called all the interviewees again to confirm the interview date, time and place although I had to reschedule some appointments after arriving at Seoul due to the sudden opening of the National Assembly. During this process, however, I had an opportunity to explain the topic guides in detail again, and if interviewees had some doubts about the exact meaning of questions, I was able to give them more explanations.

The first field work was carried out in Seoul from 25th October to 15th November 2011. The interviews were conducted face-to-face using semi-structured topic guides at different venues, such as the respondents' offices or other meeting places. Each interview typically lasted for about one hour. Interviews were conducted in Korean because all the respondents were Korean. They were audio-recorded and at the same time field notes were taken to facilitate analysis later. All the respondents showed very cooperative attitudes and gave their opinions on, and insights into, the SABP system very freely, even though they knew that their opinions might be different from those of the MOSF. Also, they were eager to suggest useful policy alternatives so that a better performance budgeting system could be produced in Korea. During the interviews, I

informed interviewees again that the study's aim was not to indicate the faults of spending ministries but to explore differing perspectives between spending ministries and the MOSF. I also explained how the results would be used, and that the opinions of interviewees would be treated with complete confidentiality.

After returning to the UK, I reviewed and summarised all the interview records as well as making 38 full interview transcripts in English for further analysis. The detailed process of data analysis will be explained in the next section. When analysing the interview data, I realised the necessity of supplementary interviews in order to suggest feasible policy alternatives for revising and redesigning the SABP system, because a variety of policy alternatives were raised by interviewees during the first interviews. Thirty one supplementary interviews were conducted in about a month, from 20th October to 20th November 2012, using internet phone and e-mail, with similar preparations to those for the first interviews in August to September 2012. However, some interviewees could not give interviews due to their busy schedules, so 14 spending ministry officials, 12 MOSF officials and five experts participated in the supplementary interviews. In addition, some members of the MOSF were replaced by others who had more responsibility for revising the SABP system, because the aim of the supplementary interviews was to determine feasible policy alternatives. The supplementary interviews generally lasted for about thirty minutes and many interviewees sent me written replies via e-mail. The supplementary interviews were successful not only because I could save time and cost on the study but also because I could communicate again as much as I needed with many interviewees, due to the advanced telecommunication technology.

4.5.6 Data Analysis

Data collected from multiple sources were analysed according to the conceptual framework outlined in Chapter 2, including making sense of the concepts of various dysfunctional consequences, categorising them, investigating their impact and developing the role of spending ministries and the MOSF to improve SABP. The data analysis process was carried out based on two strategies: "coding of data" and "synthesising cases". Coding is defined by Miles and Huberman (1994, p.56) as "to review a set of field notes transcribed or synthesised, and to dissect them meaningfully, while keeping the relations between the parts intact". The study employed computer-aided coding of documentary and interview data for assigning meaning to the collected information, for classifying the data into subgroups for each research question, and for developing relationships between them. Synthesising cases means "aggregating findings across a series of individual studies" (Yin, 2009, p.156). So, the study treated multiple cases as individual separate studies and probed whether different groups of cases showed similarities, reflecting dysfunctional consequences of SABP. Nvivo 9, which is qualitative data analysis software, was used for data coding and analysis of the first interviews. (In the coding and analysis of supplementary interviews, an Excel programme was used because the number of questions was smaller than for the first interviews.) This software was developed in order to analyse data which are difficult to categorise and organise, including functions related to classifications, queries, reports and models (Park, 2011). The process of Nvivo 9 was followed by producing project, document, node and attribute coding, and developing relationships and models, queries,

and reports (Park, 2011). In particular, the Nvivo programme enabled flexible approaches to coding and numerical counts of references in interview data (Bazeley, 2007). The study also used some strategies for increasing the validity of findings. Firstly, for data collection and data analysis, the study used multiple sources of data to minimise potential biases related to any single source. Secondly, as mentioned earlier, it developed a database consisting of documentary and interview transcript data, and tried to maintain a chain of evidences. Thirdly, interview transcripts and initial analyses of cases were reviewed with my supervisors and electronically stored on my computer with a password to enable other researchers to access them. Finally, the results of further analysis of the case study were discussed with many prestigious researchers at an academic conference in June 2012 and the conference paper was also sent to key participants in the case study. Comments from both the conference and key participants made me revisit and re-examine the various data sources to increase the validity of the findings.

4.5.7 Ethical Considerations

De Vaus (2001) proposes four ethical issues: confidentiality, privacy, avoidance of harm to participants, and informed consent. Bryman (2008) also offers four ethical principles: whether there is harm to participants; whether there is a lack of informed consent; whether there is an invasion of privacy; whether deception is involved. Moreover, Creswell (2009) argues that ethical issues relate to all phases of the research process from setting the research problem to writing and disseminating the research. This study includes the exploration of opinions and beliefs of government officials on

sensitive and complex issues related to differences in perspectives on SABP between spending ministries and the MOSF. It was therefore very important to obtain the honest opinions of respondents, and so the ethical issues mentioned above needed to be seriously engaged with. In particular, how to secure informed consent and how to avoid potential harm to participants were the main ethical issues of the study. In order to obtain informed consent, I contacted all participants in advance and explained the background of research, the purposes of the interviews, the future use of the interview results, participant feedback, participant withdrawal, confidentiality, data storage, and data access and disposal, as mentioned earlier. As a result, all participants agreed with the contents of the consent form. Also, during the interviews I explained these things again to the participants and they all agreed with me again. Furthermore, I sent my conference paper to all participants in order to ask whether the contents of the paper invaded their privacy or harmed them and whether the detailed purposes and findings were appropriate. Many participants, including key participants, gave me positive feedback and useful comments. After receiving these comments, I revisited and re-examined the interview data and the findings to increase the validity of the case study.

4.5.8 Reflexivity

Some important methodological issues, such as access to interviewees and my background and role, arose during the study. As regards access to interviewees, at first, I anticipated some difficulty in accessing spending ministry officials because expressing different perspectives and opinions on SABP to those of the MOSF might be psychologically stressful to them. However, fortunately, in practice, most of the

spending ministry officials who received my invitation to interview were happy to accept and participate in the study. Moreover, they answered very honestly and sometimes very boldly. For example, they responded that they set the performance target levels very conservatively, taking the next year's target levels into consideration, and this was a natural phenomenon for them, or that SABP was not a "self-assessment" system because the MOSF changed the spending ministries' self-assessment results. They may have thought of the interview as an opportunity to express their opinions more informally to an MOSF official and that the important thing was to find solutions, not the problem itself. Consequently, I had a chance to hear various frank voices on the subject of the SABP system.

On the other hand, approaches to MOSF officials and experts were very easy and convenient for me because most of them had worked with me before I started the study. In this sense, the second concern associated with my background and role arose because, as one of the MOSF's officials, I actually reviewed spending ministries' self-assessment results and implemented SABP. Familiarity with the interviewees and research areas could lead to significant subjectivity. To help with this problem, my supervisors, academic researchers at the conference and colleagues in the PhD programme reviewed and commented on the interview transcripts and findings, and their comments were helpful in identifying bias. I also made an effort to avoid potential bias by intentionally considering phenomena from the perspective of spending ministry officials or neutral experts. However, my experience of the MOSF officials gave me some advantages when it came to conducting the research. For example, my professional background gave me more insights into the inner procedures surrounding SABP implementation and

suggested more (i.e. a greater number of) feasible policy alternatives than other researchers might have come up with. In addition, I could obtain high quality documentary sources, such as the internal reports and data of the MOSF, and I had the opportunity to interview highly prestigious people such as a vice-minister or directors-general in central government. These advantages could make the research more accurate, influential and significant.

4.6 Conclusion

This chapter has discussed the research methodology for this thesis from the philosophical worldview to the research design and the methods adopted. It has identified the research questions for both quantitative and qualitative approaches from the relevant literature review. Under its pragmatic philosophical worldview, the study did not seek to mix different worldviews but to bring a more thorough understanding and mutual corroboration by employing both quantitative and qualitative approaches (Creswell and Plano Clark, 2011). As for the quantitative approach, the chapter defined and explained the variables and their meanings, constructing multiple regression models for verifying the factors which might affect the differences between the SABP scores of spending ministries and those of the MOSF. With respect to the qualitative approach, the chapter considered multiple case study as a research strategy, explaining why and how the case study areas and specific programmes were selected, and why the research methods adopted were appropriate for the study. The following chapters will present the results of both quantitative and qualitative analysis based on the empirical data.

CHAPTER 5

QUANTITATIVE ANALYSIS: FACTORS WHICH AFFECT THE DISAGREEMENT RATIO AND RELATIVE OPTIMISM RATIO

5.1 Introduction

This chapter examines the differences between spending ministries and the MOSF in the processes of SABP by analysing the relationships between the variables operationalised in Chapter 4 using various statistical methods. The chapter begins with a preliminary statistical analysis, including a correlation analysis, scatter diagram and t-test, to ascertain the relationships between the independent and control variables and the dependent variables. Based on the results of this preliminary analysis, the rest of the chapter explains the relationships between variables by employing multiple regression analysis, particularly focusing on verifying whether the results of analysis are consistent with the predictions mentioned in Chapter 4. The multiple regression analysis results will be presented for three different time periods: the first SABP cycle (2005-2007), the second SABP cycle (2008-2010) and the whole period (2005-2010). In addition, the dependent variables (the DR and ROR) are examined in the following SABP sections in each time period: performance planning, results and the whole section. Furthermore, in order to explore the strategic behaviours of spending ministries and the MOSF to a

given ROR level, the chapter analyses the dynamic pattern of the RORs of spending ministries where SABP has been implemented two or three times. The main results of the chapter will be interpreted and discussed in Chapter 7.

5.2 Preliminary Statistical Analysis of Variables

Preliminary analysis results are useful to examine the appropriateness of a regression equation in advance. Also, the results can be applied to interpreting the regression analysis results. This section firstly describes the differences between SABP scores and grades allocated by spending ministries and those of the MOSF, according to programme type, programme budget size, organisation type and programme purpose; and then an analysis of the relationship between the independent and control variables and the dependent variables will be presented.

5.2.1 Differences in SABP Scores and Grades

Firstly, differences in SABP scores and grades between spending ministries and the MOSF for the six years are shown as pooled data, according to programme type, in Table 5-1. As mentioned earlier, the seven programme types can be divided into two types (programmes directly managed by government and those indirectly managed), so the differences between each programme type are calculated by sub-total. In total, the average SABP score of spending ministries was 89.52, whilst that of the MOSF was 63.13, and the average SABP grade of spending ministries was 3.70, which is just below “Effective”, while that of the MOSF was just 2.25, which is nearly “Modest”.

Thus, the differences in SABP scores and grades were 26.39 and 1.46 respectively. The SOC programmes, among the three directly managed programme types, and the subsidy to local government programme, among the four indirectly managed programmes, had the highest differences in SABP scores and grades. The differences in the SABP scores and grades of directly managed programmes were 24.93 and 1.36 respectively, while those of the indirectly managed programmes were 27.20 and 1.51 respectively. Thus, we can see that spending ministries have an optimism bias in the self-assessment stage in all types of programme, and that the differences between directly managed and indirectly managed programmes (24.93 vs. 27.20 and 1.36 vs. 1.51) might imply that programme type had an influence on the differences between spending ministries and the MOSF.

Table 5-1 Differences in SABP Scores and Grades According to Programme Type (Six-Year Pooled Data)

(Unit: points)

Programme type		Programme number	Spending ministries' self-assessment		MOSF's review		Difference	
			Score(A)	Grade(C)	Score(B)	Grade(D)	A-B	C-D
Total		2,920	89.517 (9.998)	3.704 (0.550)	63.132 (11.994)	2.248 (0.686)	26.385	1.456
Directly managed programme	Sub-total	1,049	90.027 (10.133)	3.738 (0.550)	65.100 (11.995)	2.376 (0.686)	**24.927**	**1.362**
	SOC	153	90.412 (10.129)	3.719 (0.560)	64.399 (12.122)	2.333 (0.692)	26.013	1.386
	Capital Acquisition	38	87.100 (10.018)	3.579 (0.543)	65.443 (12.300)	2.421 (0.702)	21.657	1.158
	Other direct programme	858	90.088 (10.133)	3.748 (0.550)	65.210 (11.995)	2.381 (0.686)	24.878	1.367
Indirectly managed programme	Sub-total	1,871	89.231 (9.998)	3.686 (0.550)	62.029 (11.994)	2.176 (0.686)	**27.202**	**1.510**
	Investment	259	90.664 (10.003)	3.737 (0.551)	63.769 (12.014)	2.290 (0.687)	26.895	1.447
	Loan	245	88.192 (10.173)	3.649 (0.554)	61.172 (12.075)	2.127 (0.691)	27.020	1.522
	Subsidy to private sector	817	88.874 (10.136)	3.668 (0.551)	62.323 (12.007)	2.198 (0.687)	26.551	1.470
	Subsidy to local government	550	89.550 (10.136)	3.704 (0.550)	61.155 (11.989)	2.113 (0.686)	28.395	1.591

Note: 1. Standard deviations are reported in parentheses. 2. SABP grades are converted as follows: poor grade is 1, modest is 2, somewhat effective is 3 and effective is 4.

Secondly, differences in SABP scores and grades according to programme budget size, which are divided into four categories on the basis of quartile (8.3bn Won), are presented in Table 5-2. In first quartile programmes (small programmes), the differences in SABP scores and grades between spending ministries and the MOSF were about 26.99 and 1.47 respectively. In addition, in fourth quartile programmes (large programmes), there were also big differences between the SABP scores and grades of spending ministries and those of the MSOF, and these were around 26.39 and 1.45 respectively. Differences between small and large programmes in SABP scores (26.99 vs. 26.39) and SABP grades (1.47 vs. 1.45) might imply that programme budget size affects the difference between spending ministries and the MOSF.

Table 5-2 Differences in SABP Scores and Grades According to Budget Size (Six-Year Pooled Data)

(Unit: points)

Programme budget size	Programme number	Spending ministries' self-assessment		MOSF's review		Difference	
		Score(A)	Grade(C)	Score(B)	Grade(D)	A-B	C-D
Total	2,858	89.512 (9.998)	3.706 (0.550)	63.150 (11.994)	2.251 (0.686)	26.362	1.455
First quartile (budget<2.4bn Won)	659	88.911 (10.002)	3.665 (0.551)	61.922 (12.010)	2.197 (0.688)	**26.989**	**1.468**
Second quartile (2.4≤budget<8.3bn Won)	740	89.255 (10.009)	3.691 (0.551)	62.909 (12.003)	2.223 (0.687)	26.346	1.468
Third quartile (8.3≤budget< 32.3bn Won)	729	88.968 (9.999)	3.691 (0.550)	63.186 (11.996)	2.254 (0.686)	25.782	1.437
Fourth quartile (budget≥32.3bn Won)	730	90.857 (9.995)	3.774 (0.550)	64.465 (12.007)	2.326 (0.687)	**26.392**	**1.448**

Note: 1. Standard deviations are reported in parentheses. 2. 62 programmes have a missing value.

Thirdly, differences in SABP scores and grades between spending ministries and the MOSF, according to organisation characteristics, are described in Table 5-3. The SABP scores and grades of ministry programmes were about 27.24 and 1.51 respectively,

while those of agency programmes were around 23.54 and 1.29 respectively. These differences might imply that organisation characteristics affect the differences between spending ministries and the MOSF.

Table 5-3 Differences in SABP Scores and Grades Between Programmes Belonging to Ministries and Agencies (Six-Year Pooled Data)

(Unit: points)

Characteristics of organisation	Programme number	Spending ministries' self-assessment		MOSF's review		Difference	
		Score(A)	Grade(C)	Score(B)	Grade(D)	A-B	C-D
Total	2,920	89.517 (9.998)	3.704 (0.550)	63.132 (11.994)	2.248 (0.686)	26.385	1.456
Programmes belonging to ministries	2,248	89.343 (9.998)	3.699 (0.550)	62.108 (11.994)	2.193 (0.686)	**27.235**	**1.506**
Programmes belonging to agencies	672	90.099 (10.019)	3.722 (0.552)	66.559 (12.046)	2.432 (0.689)	**23.540**	**1.290**

Note: Standard deviations are reported in parentheses.

Finally, Table 5-4 presents the differences in SABP scores and grades according to programme purpose. Differences between the SABP scores and grades of spending ministries and those of the MOSF for economy-related programmes were about 25.59 and 1.38 respectively, while those for welfare-related programmes were around 28.37 and 1.63 respectively. These results might imply that programme purpose affects the differences between spending ministries and the MOSF.

Table 5-4 Differences in SABP Scores and Grades Between Economy-Related Programmes and Welfare-Related Programmes (Six-Year Pooled Data)

(Unit: points)

Purpose of programme	Number of programme	Spending ministries' self-assessment		MOSF's review		Difference	
		Score(A)	Grade(C)	Score(B)	Grade(D)	A-B	C-D
Economy-related programmes	1,018	89.586 (10.029)	3.688 (0.553)	63.994 (12.075)	2.308 (0.690)	**25.592**	**1.380**
Welfare-related programmes	520	88.273 (10.055)	3.662 (0.550)	59.900 (11.987)	2.033 (0.688)	**28.373**	**1.629**

Note: Standard deviations are reported in parentheses.

Differences in SABP scores and grades according to the number of times SABP had been implemented were not described here, due to the difficulty of sorting data. Instead, the impact of this variable on the dependent variables is mentioned in the following section's relationship analysis.

5.2.2 Relationship Analysis of Variables

The descriptive analysis in 5.2.1 might imply that variables such as programme type, programme budget size, organisation characteristics and programme purpose impact on the differences between the SABP results of spending ministries and those of the MOSF. However, this analysis cannot provide the necessary information about whether, and to what extent, these relationships are statistically significant. This section gives more useful information, indicating a statistical significance between variables by employing correlation analysis and a t-test before analysing the regression equation.

Firstly, the relationship between the independent and the dependent variables can be presented by the correlation analysis and scatter diagram. The MOSF review score had a statistically significant relationship with the DR and ROR, at the significance level of 0.01, in all time periods, and the effective direction was completely negative. In addition, a specific programme's self-assessment score minus the spending ministry's average score in that year (specific score minus average score) had a positive relationship with the DR and ROR in a statistically significant way, at the level of 0.01, in all time periods. Table 5-5 shows the Pearson correlation coefficients between the two independent variables and the DR for the whole section in the three time periods,

and Table 5-6 presents them between the two independent variables and the ROR for the whole section in the three time periods. Correlation coefficients between the two independent variables and the DR and ROR generally decreased from the first SABP cycle to the second SABP cycle as follows: the correlation coefficient between the MOSF review score and the DR decreased from -0.548 to -0.414; the coefficient between the MOSF review score and the ROR also fell, from -0.397 to -0.329; and the coefficient between the specific score minus the average score and the ROR declined from 0.274 to 0.176. By contrast, the coefficient between specific score minus average score and the DR increased from 0.076 in the first SABP cycle to 0.141 in the second cycle. These trends of coefficients can also affect the trends of R square in the regression analysis and the implications of this will be discussed in Chapter 7. In the six-year period, the pooled data for the two independent variables had closer relationships with the ROR than the DR. The MOSF review score had a -33.2% relationship with the ROR and the specific score minus average score had a 21.6% relationship with the ROR, while the relationships between these variables and the DR were -22.0% and 0.92% respectively.

Table 5-5 Pearson Correlation Coefficients Between Two Independent Variables and the DR for the Whole Section

	’05-’07	’08-’10	’05-’10
MOSF review score	-0.548***	-0.414***	-0.220***
Specific score minus average score	0.076***	0.141***	0.092***

Note: *** denotes statistically significant at the 0.01 level (one-tailed test). Also the case for Table 5-6.

Table 5-6 Pearson Correlation Coefficients Between Two Independent Variables and the ROR for the Whole Section

	’05-’07	’08-’10	’05-’10
MOSF review score	-0.397***	-0.329***	-0.332***
Specific score minus average score	0.274***	0.176***	0.216***

Furthermore, from Figures 5-1 and 5-2, I was able to confirm that the two independent variables had linear relationships with the DR and the ROR for the whole section in the three time periods.

Figure 5- 1 Scatter Diagram Between Two Independent Variables and the DR in the Whole Section

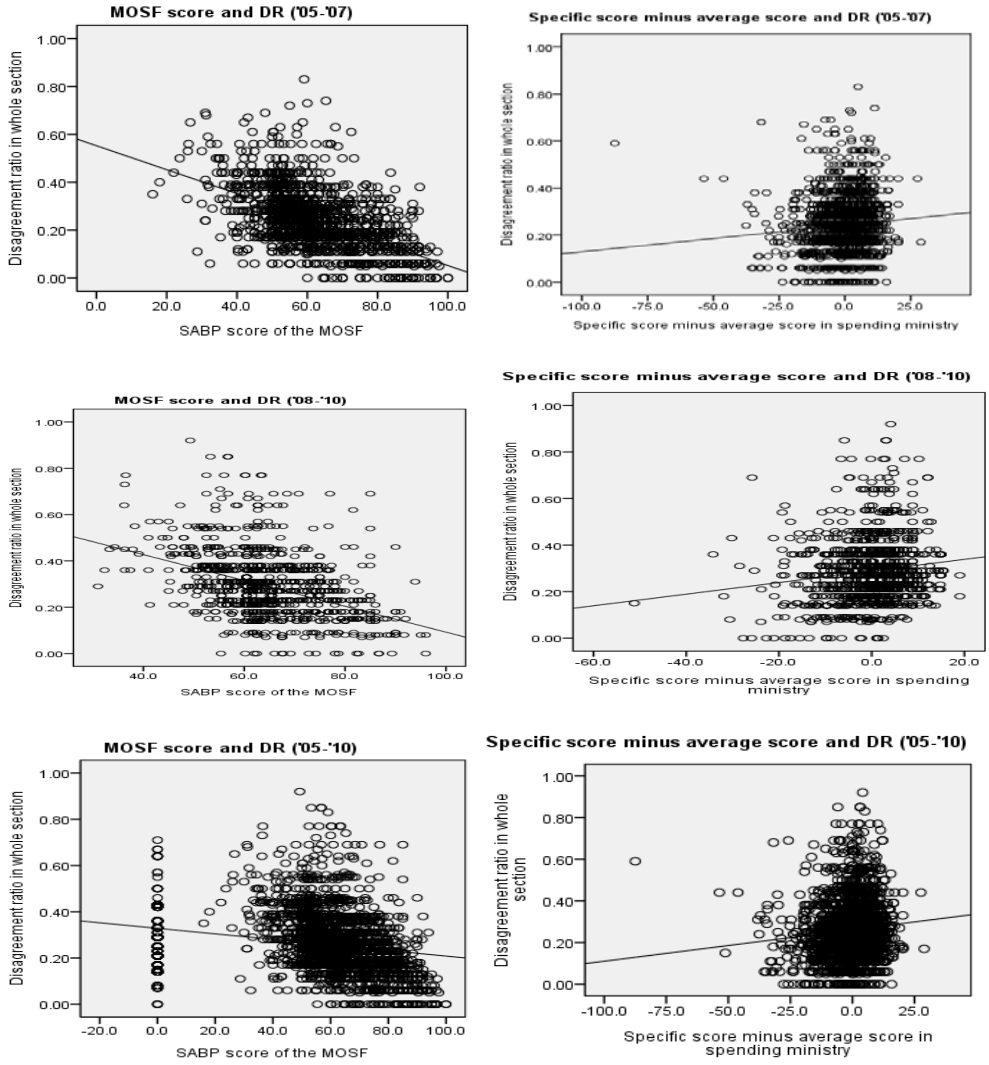
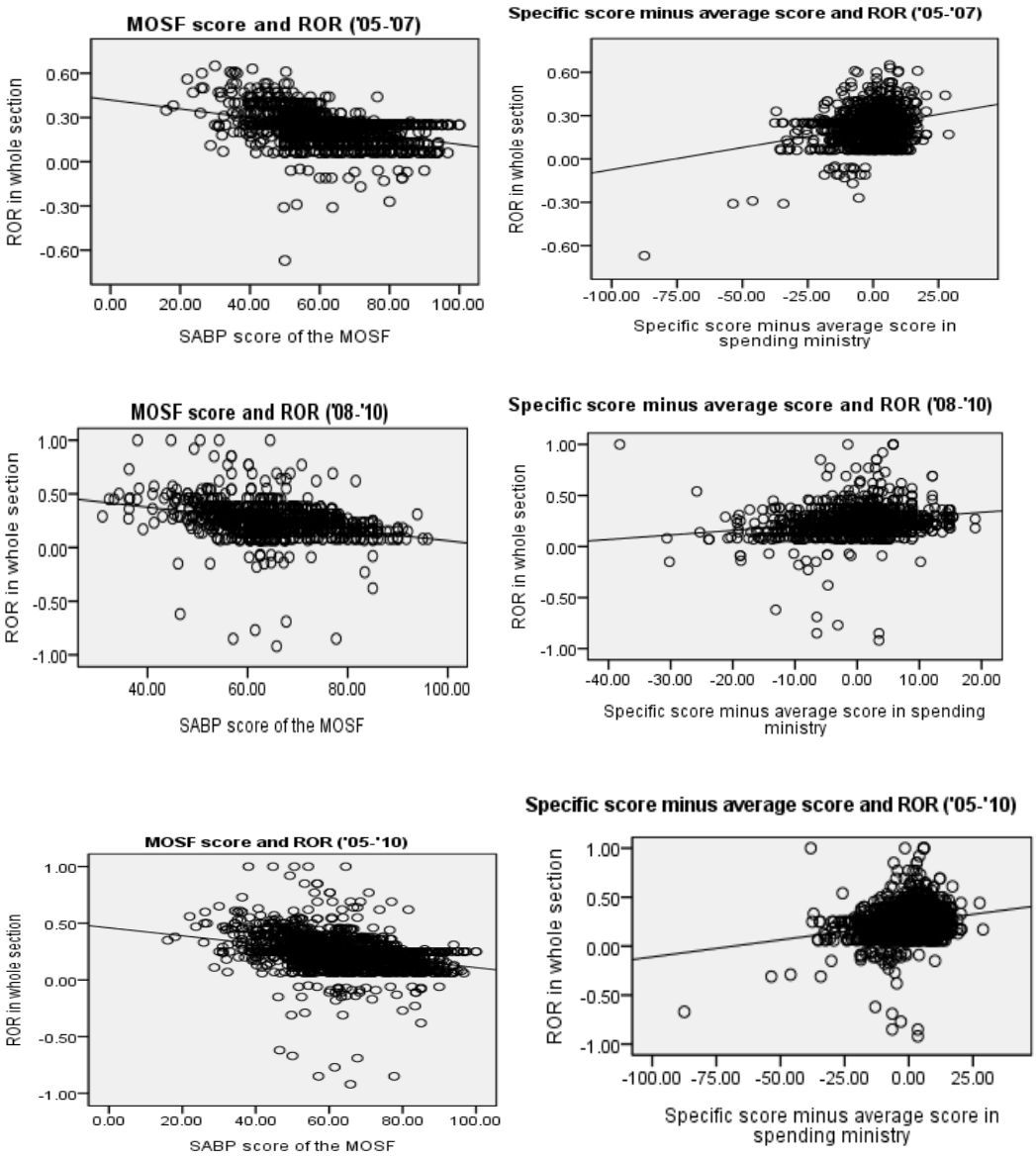


Figure 5- 2 Scatter Diagram Between Two Independent Variables and the ROR in the Whole Section



From the correlation analysis and the scatter diagrams it was possible to draw the conclusion that the two independent variables had statistically significant linear relationships with the DR and ROR, and these results confirmed that the two variables could be used as the independent variables in the regression equation mentioned in Chapter 4.

Secondly, the relationship between the control variables and the dependent variables could be presented by correlation analysis and a t-test. Budget percentage changes for the current year had a statistically significant relationship with the DR in the second SABP cycle and in the pooled data for the six-year period, at the significance level of 0.1. However, this variable did not have any statistically significant relationship with the ROR.

Table 5-7 Pearson Correlation Coefficients Between Budget Percentage Changes for the Current Year and the DR in Three Sections

	’05-’07	’08-’10	’05-’10
Performance planning section	0.031	-0.043*	0.028*
Results section	-0.031	0.045*	-0.010
Whole section	-0.004	0.048*	-0.014

Note: *denotes statistically significant at the 0.1 level (one-tailed test)

Table 5-8 Pearson Correlation Coefficients Between Budget Percentage Changes for the Current Year and the ROR in Three Sections

	’05-’07	’08-’10	’05-’10
Performance planning section	0.019	-0.031	0.012
Results section	-0.028	0.003	-0.021
Whole section	0.013	-0.025	0.003

The mean differences of the DR and the ROR for the whole section for the six-year period, according to programme type, were statistically significant at the level of 0.01. Table 5-9 shows the results of a t-test.

Table 5-9 Mean Differences of the DR and ROR for the Whole Section (’05-’10), According to Programme Type (t-test)

Programme type	DR	ROR
Direct programme	0.2504	0.2316
Indirect programme	0.2677	0.2435
Mean difference	-0.1725	-0.01188
t-value	-3.337***	-2.648***

*** p<0.01

The mean differences of the DR and ROR for the whole section for the six-year period, according to programme budget size, were not statistically significant, but the mean differences of the DR and ROR for the results section for the six-year period and those of the ROR for the performance planning section for the six-year period were statistically significant in the cases of small budget size and large budget size programmes respectively. Also, interestingly, both small budget and large budget programmes all had positive relationships with the DR and ROR. Tables 5-10 and 5-11 present these results.

Table 5-10 Mean Differences of the DR and ROR in the Results Section (’05-’10), According to Budget Size

Budget size	DR	ROR
Small budget programme	0.4748	0.4555
Other programme	0.4536	0.4260
Mean difference	0.02119	0.02953
t-value	1.802*	2.721***

* p < 0.1, *** p<0.01

Table 5-11 Mean Differences of the DR and ROR in Performance Planning Section (’05-’10), According to Budget Size

Budget size	DR	ROR
Large budget programme	0.3786	0.5201
Other programme	0.3708	0.4836
Mean difference	0.00782	0.03649
t-value	0.555	2.918***

*** p<0.01

With respect to organisation type, the mean differences of the DR and ROR for the whole section for the six-year period were statistically significant at the level of 0.01. Table 5-12 presents the results.

Table 5-12 Mean Differences of the DR and ROR in the Whole Section (’05-’10), According to Organisation Type

Organisation type	DR	ROR
Programmes belonging to ministries	0.2687	0.2449
Programmes belonging to agencies	0.2358	0.2189
Mean difference	0.03282	0.02601
t-value	5.249***	3.927***

*** p<0.01

As for economy-related programmes, the mean difference of the DR for the whole section for the six-year period was statistically significant at the level of 0.05 and the effective direction was negative, but that of the ROR was not statistically significant (see Table 5-13). With regard to welfare-related programmes, the mean difference of the DR was statistically significant at the level of 0.01 and the effective direction was positive. However, that of the ROR did not show any statistical significance (see Table 5-14).

Table 5-13 Mean Differences of the DR and ROR in the Whole Section (’05-’10), According to the Purpose of Programmes

Purpose of programme	DR	ROR
Economy-related programmes	0.2529	0.2433
Other programmes	0.2660	0.2370
Mean difference	-0.01311	0.00627
t-value	-2.484**	1.320

** p<0.05

Table 5-14 Mean Differences of the DR and ROR in the Whole Section (’05-’10), According to the Purpose of Programmes

Purpose of programme	DR	ROR
Welfare-related programmes	0.2926	0.2398
Other programmes	0.2550	0.2390
Mean difference	0.03762	0.00074
t-value	5.740***	0.116

*** p<0.01

With respect to the number of times SABP was implemented, the mean difference of the DR and the ROR for the whole section for the six-year period was statistically significant at the level of 0.05 and the effective direction was negative as shown in Table 5-15.

Table 5-15 Mean Differences of the DR and ROR in the Whole Section (’05-’10), According to the Number of Times SABP Implemented

The number of SABP implementations	DR	ROR
SABP twice or more	0.2559	0.2350
Other programmes	0.2689	0.2446
Mean difference	-0.01297	-0.00957
t-value	-2.488**	-2.013**

** p<0.05

Overall, considering carefully the results of these preliminary analyses, it may be said that the two independent variables, programme type and organisation type, and the number of times SABP had been implemented, showed statistically significant relationships with the DR and ROR. However, these results were mainly related to the DR and whole section for the six-year period, so an analysis of those in the performance planning and results sections might have rather different results. The next sections will

analyse in detail the factors which might affect the DR and ROR for the three sections in the three different time periods.

5.3 Factors Which Affect the Disagreement Ratio

Factors which affect the DR for the three time periods were analysed[23] based on three kinds of DR: the DR for the performance planning section, the DR for the results section and the DR for the whole section. The impact of the independent variables and control variables, as well as the goodness-of-fit test of the regression line (R Square, Adjusted R square and F ratio), were analysed sequentially.

5.3.1 Analysis of the Independent Variables

In Table 5-16, the MOSF's review score affected negatively the DR for three sections in the three different time periods, at the significance level of 0.01. For the six-year pooled data, the coefficients of the MOSF's review score are -0.015 for the DR for the performance planning section, -0.008 for the DR for the results section, and -0.002 for the DR for the whole SABP section. This means that if the MOSF's review score increases by 10 points, then the DR for the performance planning section, results section and whole section will decrease by 0.15%, 0.08% and 0.02% respectively with a 99%

[23] A multiple regression model was analysed by the SPSS (Statistical Package for the Social Sciences) and regression diagnostics were implemented. Hypotheses of an error term were verified through residual analyses such as residual plot (homoscedasticity), Durbin-Watson test (non-autocorrelation; near to 2), and normal probability plot (normal distribution). There was no multicollinearity between the variables because the VIF (variance inflation factor) was always less than 10 and the tolerance value was always less than 0.1. In addition, through both F test and t-test the significance of the linear regression line was verified. Enter method as a regression analysis method was used instead of the stepwise method because this study's purpose is the verification of the relationship between variables, not designing a prediction model.

probability. In both the first SABP cycle and the second cycle, the results were similar to those in the six-year period. Consequently, the MOSF's review score had a negative relationship with the DR for the three sections and in all the time periods, in a statistically significant way, at the level of 0.01.

Also, the other independent variable, specific score minus average score, had a positive influence on the DR for the three sections for all the time periods, at the significance level of 0.01. For example, in the six-year pooled data, the coefficients of this variable were 0.008 for the DR for the performance planning section, 0.007 for the DR for the results section and 0.002 for the DR for the whole section, and this means that if the difference increases by 10 points, then the DR for performance planning, results and the whole section would also increase by 0.08%, 0.07% and 0.02% respectively with a 99% probability. As for the first and second SABP cycles, the results were similar to those in the six-year pooled data. Consequently, this variable had a positive relationship with the DR in three sections for the three time periods, at the significance level of 0.01.

5.3.2 Analysis of the Control Variables

The control variables had different effects on the DR according to the different time period. Firstly, budget changes for the current year had a statistically significant effect on the DRs for the results section and the whole section in the second SABP cycle, at the significance level of 0.05, and on the DR for the results section in the six-year pooled data, at the significance level of 0.1. However, the effective direction in the second cycle was positive, while that in the six-year period was negative. Overall, this

variable might have a statistically significant relationship with the DR, but it was not clear whether the effective direction was positive or negative.

Secondly, programmes directly managed by government had a statistically significant effect on the DRs for three sections for the first SABP cycle, at the level of 0.01, but the effective direction of performance planning was negative, while the effective directions of the other two sections were positive. In the second SABP cycle, this variable had a negative relationship with the DR for the performance planning section, at the significance level of 0.1. In the six-year pooled data, this variable had a statistically significant impact on the DRs for the performance planning and results sections, at the level of 0.01. However, the effective direction of the former was negative, while that of the latter was positive. Overall, although this variable had a negative relationship with the DR for the performance planning section in the three time periods, considering the impact of this variable on the DRs for the results section and the whole section, it can be said that this variable had a positive relationship with the DR.

Thirdly, small programmes had a statistically meaningful effect on the DR for the performance planning section for the first SABP cycle and on the DR for the whole section in the six-year pooled data. However, the effective direction of the former was negative, while that of the latter was positive. Considering the impact on the DR for the whole section for the six-year period, it can be said that this variable might have a positive relationship with the DR.

Fourthly, large programmes had a statistically significant impact on the DR for the performance planning section in the first cycle and on the DRs for performance planning and the whole section in the six-year pooled data. The effective directions were all positive. Consequently, large programmes had a positive relationship with the DRs for the performance planning and results sections.

Fifthly, the ministries' programmes had a statistically significant relationship with the DRs for the performance planning and results sections in the first SABP cycle, and the effective direction was negative. In the second SABP cycle, this variable had a significant effect on the DRs for the performance planning and results sections, but the effective directions were opposite to each other. In the six-year pooled data, this variable had a significant influence on the DRs for performance planning and the whole section, but the effective directions were also opposite to each other. Overall, this variable had a statistically significant relationship with the DR, but the effective direction was not clear.

Sixthly, economy-related programmes only had a statistically meaningful impact on the DR for the whole section in the six-year pooled data, and the effective direction was negative. However, it did not have any statistically significant relationship with the DRs for the first and second SABP cycles. Consequently, economy-related programmes had a negative relationship with the DR for the whole section in the six-year pooled data.

Seventhly, welfare-related programmes had a statistically significant influence on the DR in the first SABP cycle, and the effective direction in the performance planning

section and the whole section was positive, while that in the results section was negative. In the second SABP cycle, this variable had a statistically significant negative relationship with the DRs for the results section and the whole section. In the six-year pooled data, this variable had a positive relationship with the DR for performance planning, while it had a negative relationship with the DR for the results section. Consequently, welfare-related programmes had a statistically significant impact on the DR in the three time periods, but considering the effective direction of the whole section in the first and the second cycle was the opposite, the effective direction of the variable was not clear.

Finally, the number of times SABP had been implemented had a statistically significant influence on the DR for the results section in the first SABP cycle, and the effective direction was positive. However, it had a negative relationship with the DR for the whole section in the second SABP cycle. In the six-year pooled data, this variable had a negative relationship with the DR for the performance planning section, while it had a positive relationship with the DR for the whole section. So, the number of times the SABP had been implement had a statistically significant relationship with the DR in the three time periods, but considering the effective direction of whole section in the second cycle and in the six-year period was the opposite, the effective direction of the variable was also not clear.

5.3.3 Analysis of the R square, Adjusted R square, and F ratio

The R square is the extent to which predictors in the regression model can explain the variation of the dependent variable. The highest R square is 0.394 for the performance planning section in the second SABP cycle; and the lowest one is 0.082 for the whole section in the six-year pooled data. This means that 10 variables can account for 39.4% of the variation of the DR for the performance planning section in the second SABP cycle, whilst they can only explain 8.2% of the variation of the DR for the whole section in the six-year pooled data. However, most of the R squares ranged from 0.23 to 0.35, so the values of R square could be said to be relatively high. Among the three sections, the R square for the performance planning section was the highest in the three time periods, and that for the whole section was generally the second highest. Also, the R square generally decreased from the first cycle to the second cycle. In general, the R square tends to increase as the number of predictors increases; but employing too many predictors makes it impossible to follow the principle of parsimony, one of the important rules for regression analysis, and is likely to result in multicollinearity between variables. In this case, it is necessary to adjust the influence of the number of predictors on the R square. An adjusted R square can complement the disadvantage of an R square and it indicates the extent to which the regression model can be generalised. Consequently, the smaller the difference between the R square and Adjusted R square, the better. The differences between the R squares and Adjusted R squares for all sections in the three time periods were very small, so the results of this regression model can be generalised to too great an extent. F values provide information about the predictive capability of the regression model. The F values which tested the significance of the

equation were all statistically significant, at the significance level of 0.01, for all the sections in all three time periods. Thus, the regression model employed in the study provided significantly better predictive powers for the effect of the related variables (the independent and control variables) on the dependent variables. Also, the t values, which determine whether the independent and control variables make a significant contribution to the model, could be regarded as statistically meaningful to too great a degree. The constant, which refers to the y intercept, and the gradient were both statistically significant, at the significance level of 0.01. As a result, considering the R square, Adjusted R square and F and t values of the regression model, the explanatory power and validity of this model were so high that we could use the results when explaining the factors which affected the DR.

Table 5-16 Factors Which Affect the DR in the Three Time Periods

Independent and control variables	2005-2007 (First SABP cycle)			2008-2010 (Second SABP cycle)			2005-2010		
	Performance planning section	Results section	Whole section	Performance planning section	Results section	Whole section	Performance planning section	Results section	Whole section
MOSF's review score	**-0.013*** (-29.187)**	**-0.009*** (-20.070)**	**-0.005*** (-28.957)**	**-0.021*** (-25.855)**	**-0.008*** (-11.398)**	**-0.006*** (-16.513)**	**-0.015*** (-36.815)**	**-0.008*** (-22.285)**	**-0.002*** (-13.607)**
Specific programme's self-assessment score minus average score of the ministry	**0.007*** (10.483)**	**0.008*** (12.611)**	**0.003*** (11.049)**	**0.009*** (8.522)**	**0.006*** (6.447)**	**0.004*** (8.249)**	**0.008*** (13.222)**	**0.007*** (13.742)**	**0.002*** (6.893)**
Budget changes for the current year	0.000 (0.948)	-0.001 (-1.606)	-9.091E-5 (-0.608)	-0.004 (-0.845)	**0.009** (2.409)**	**0.005** (2.316)**	0.000 (0.533)	**-0.001* (-1.732)**	0.000 (-0.895)
Programmes managed directly by government (1,0)	**-0.034*** (-2.914)**	**0.049*** (4.303)**	**0.015*** (3.076)**	**-0.032* (-1.869)**	0.009 (0.630)	0.002 (0.793)	**-0.038*** (-3.829)**	**0.031*** (3.437)**	0.002 (0.371)
Small programmes (1,0)	**-0.030** (-2.126)**	0.004 (0.332)	-0.003 (-0.463)	0.015 (0.739)	0.014 (0.780)	-0.003 (-0.376)	-0.013 (-1.133)	0.006 (0.554)	**0.019*** (3.070)**
Large programmes (1,0)	**0.032** (2.270)**	0.008 (0.577)	0.004 (0.720)	0.013 (0.697)	-0.021 (-1.308)	-0.009 (-1.096)	**0.026** (2.255)**	-0.001 (-0.143)	**0.023*** (3.874)**
Programmes belonging to ministry (1,0)	**-0.058*** (-3.752)**	**-0.028* (-1.886)**	0.007 (1.068)	**-0.036* (-1.835)**	**0.035** (2.078)**	0.013 (1.459)	**-0.047*** (-3.776)**	-0.004 (-0.315)	**0.017*** (3.609)**
Programmes related to economy (1,0)	0.004 (0.334)	-0.021 (-1.636)	-7.085E-5 (-0.013)	-0.005 (-0.285)	0.013 (0.804)	0.004 (0.444)	-0.002 (-0.199)	-0.009 (-0.891)	**-0.015** (-2.502)**
Programmes related to welfare (1,0)	**0.076*** (4.568)**	**-0.045*** (-2.806)**	**0.036*** (5.141)**	-0.035 (-1.583)	**-0.048** (-2.529)**	**-0.031*** (-3.053)**	**0.036** (2.590)**	**-0.046*** (-3.727)**	0.006 (0.895)
Programmes implemented SABP by two or more times (1,0)	-0.011 (-0.947)	**0.026** (2.330)**	-0.003 (-0.644)	0.011 (0.691)	-0.002 (-0.171)	**-0.014* (-1.953)**	**-0.021** (-2.186)**	-0.001 (-0.082)	**0.015*** (3.129)**
Constant	**1.221*** (36.658)**	**0.985*** (30.590)**	**0.564*** (40.037)**	**1.801*** (30.549)**	**0.980*** (19.468)**	**0.693*** (25.683)**	**1.373*** (45.714)**	**0.985*** (36.203)**	**0.321*** (52.978)**
F	**96.346***	**50.006***	**95.829***	**73.598***	**17.760***	**34.260***	**150.152***	**59.672***	**24.828***
Adjusted R^2	0.363	0.227	0.362	0.389	0.128	0.228	0.346	0.173	0.078
R^2	0.367	0.231	0.366	0.394	0.136	0.234	0.349	0.175	0.082

Notes: 1. Dependent variable is disagreement ratio in each programme. 2. Unstandardized coefficient value of each variable and t values in the parenthesis are reported. 3. * is statistically significant at the 0.10 level, ** at the 0.05 level, and *** at the 0.01 level.

5.4 Factors which Affect the Relative Optimism Ratio

Analyses of the factors which affected the RORs of spending ministries were carried out in the same way as in the previous section.

5.4.1 Analysis of the Independent Variables

From Table 5-17, it can be seen that the MOSF's review score had a statistically significant impact on the RORs for three sections in the three different time periods, at the significance level of 0.01. The effective direction was completely negative. Unstandardized coefficients of the variables in the three sections in the six-year pooled data were -0.012, -0.008 and -0.004 respectively. As mentioned earlier, this means that if the MOSF's review scores increase by 10 points, then the RORs for performance planning, results and the whole section will decrease by 0.12%, 0.08% and 0.04% respectively. The results for the other two time periods were similar to those for the six-year time period. As a result, the MOSF's review scores had a negative relationship with the RORs for the three sections in the three time periods, at the significance level of 0.01.

The other independent variable had a positive impact on the RORs for the three sections in the three different time periods, at the significance level of 0.01, because all unstandardized coefficients of this variable were positive. For example, in the six-year pooled data, unstandardized coefficients in the three sections were 0.008, 0.011 and 0.005 respectively. This means that if the specific score minus average score of that spending ministry increases by 10 points, then the RORs for the three sections will increase by 0.08%, 0.11% and 0.05%

respectively. Therefore, it can be said that this variable made a statistically significant contribution to the RORs in the three time periods in a positive way.

5.4.2 Analysis of the Control Variables

The control variables had different impacts on the RORs in the three time periods. Firstly, budget changes for the current year only had a significant effect on the ROR for the results section in the six-year pooled data, at the significance level of 0.1, and the effective direction was negative. However, in the other two time periods, this variable did not have any statistically significant influence on the RORs for the three sections. Thus, it can be said that this variable had a negative relationship with the ROR for the results section in the six-year pooled data.

Secondly, programmes directly managed by the government had a statistically significant effect on the ROR for the results section in the first SABP cycle and the six-year pooled data, at the significance level of 0.1, and the effective direction was all positive. However, it had no statistically significant impact on the other ROR. So, this variable had a positive relationship with the ROR for the results section in the first SABP cycle and the six-year pooled data.

Thirdly, small programmes had a statistically significant influence on the ROR for the results section in the second SABP cycle and the six-year pooled data, at the significance level of 0.01, and the effective direction was positive. Consequently, small programmes had a positive relationship with the ROR for the results section in the second SABP cycle and the six-year pooled data.

Fourthly, large programmes had a statistically significant impact on the ROR for the performance planning section in the three time periods, at the significance level of 0.01 and 0.05, and the effective direction was all positive. Thus, large programmes had a positive relationship with the ROR for the performance planning section in all the time periods.

Fifthly, programmes belonging to the ministry affected the RORs for the performance planning and results sections at the significance level of 0.1 and 0.05 respectively, and the effective direction was all negative. However, this variable also had statistically significant relationships with the RORs for the results section in the second cycle and the whole section in the six-year period, at the level of 0.1, and the effective direction was all positive. Considering the impact of the variable on the RORs in the second cycle and the six-year pooled data, and particularly on the ROR for the whole section in six-year pooled data, it can be said that this variable had a positive relationship with the RORs.

Sixthly, programmes related to the economy had statistically significant relationships with the RORs for the performance planning and results sections in the first cycle, at the level of 0.01, but the effective directions were opposite to each other. However, this variable had a statistically significant impact on the RORs for the results section of the second cycle and the performance planning section in the six-year pooled data, and the effective direction was all positive. Considering the most significant relationships, this variable had a positive relationship with the RORs.

Seventhly, welfare-related programmes affected the RORs for the performance planning and results sections, at the significance level of 0.01, but the effective directions were opposite to

each other. In the second cycle, this variable had a statistically significant impact on the RORs for the results section and whole sections, at the level of 0.05, and the effective direction was all negative. In addition, in the six-year pooled data this variable had a significant influence on the RORs for three sections, with the effective direction in performance planning being positive while those in the other two sections were all negative. Considering the various impacts of this variable on the RORs, especially the negative direction of the ROR for the whole section in the second cycle and the six-year period, it can be said that it had a negative relationship with the RORs.

Finally, the number of times SABP was implemented had a statistically significant impact on the RORs for whole section in the first cycle and the six-year pooled data, at the significance level of 0.05 and 0.01 respectively, and the effective direction was all negative. Consequently, considering the RORs for the whole section in the first cycle and the six-year pooled data were negative, this variable had a negative relationship with the RORs in a statistically significant way.

5.4.3 Analysis of the R square, Adjusted R square, and F ratio

The highest and the lowest R squares were both in the second SABP cycle. The former was 0.319, for the performance planning section; and the latter was 0.149, for the results section. As mentioned earlier, this means that 10 predictors can explain 31.9% of the variation in the ROR for the performance planning section and 14.9% of the variation in the ROR for the results section in the second cycle. Also, the R squares generally ranged from 0.20 to 0.30, so it seemed that it was relatively high level. Interestingly, the R squares for the three sections in the six-year pooled data were all over 0.20, indicating that the R square for the whole section

was 0.212. The R square for the performance planning section was relatively higher than that for the other sections in the second cycle and six-year period; but in the first cycle, the R square for the whole section reached the highest level, while that for the performance planning section reached the lowest level. The adjusted R square was similar to the R square for all the sections in the three different time periods, and this meant that the regression model could be generalised well. The F values which tested the significance of the regression model were all statistically significant, at the level of 0.01. So, the regression model employed in the study provided significantly higher explanatory powers for all predictors on the RORs. In addition, the t values determine whether the independent and control variables can also be said to make a significant contribution to the model. The constant, which refers to the y intercept, and the gradient were both statistically significant, at the significance level of 0.01. As a result, considering the R square, Adjusted R square and F and t values, the predictive power and validity of the regression model were so high that we could use the results of this model in terms of explaining the factors which affected the RORs of spending ministries.

Table 5-17 Factors Which Affect the ROR of Spending Ministries in the Three Time Periods

Independent and control variables	2005-2007 (First SABP cycle)			2008-2010 (Second SABP cycle)			2005-2010		
	Performance planning	Results section	Whole section	Performance planning	Results section	Whole section	Performance planning	Results section	Whole section
MOSF's review score	**-0.008*** (-12.711)**	**-0.009*** (-19.406)**	**-0.004*** (-23.243)**	**-0.021*** (-19.077)**	**-0.007*** (-9.455)**	**-0.006*** (-13.286)**	**-0.012*** (-21.654)**	**-0.008*** (-20.251)**	**-0.004*** (-22.924)**
Specific programme's self-assessment score minus average score of the ministry	**0.007*** (9.397)**	**0.011*** (17.726)**	**0.004*** (18.791)**	**0.009*** (7.141)**	**0.009*** (9.911)**	**0.006*** (8.786)**	**0.008*** (11.481)**	**0.011*** (19.544)**	**0.005*** (17.741)**
Budget changes for the current year	0.000 (0.452)	0.000 (-1.566)	-3.139E-6 (-0.025)	-0.001 (-0.288)	0.003 (0.914)	-0.001 (-0.535)	0.000 (0.300)	**-0.001* (-1.724)**	-7.366E-5 (-0.424)
Programmes managed directly by government (1,0)	0.020 (1.560)	**0.021* (1.833)**	0.003 (0.647)	-0.004 (-0.186)	0.019 (1.260)	0.006 (0.624)	0.003 (0.276)	**0.016* (1.815)**	0.001 (0.279)
Small programmes (1,0)	-0.011 (-0.735)	0.017 (1.290)	-0.001 (-0.257)	0.015 (0.654)	**0.050*** (2.842)**	0.015 (1.318)	0.000 (-0.021)	**0.030*** (2.778)**	0.006 (1.063)
Large programmes (1,0)	**0.043*** (2.888)**	0.017 (1.258)	0.002 (0.343)	**0.046** (2.060)**	0.001 (0.086)	0.008 (0.780)	**0.046*** (3.478)**	0.013 (1.292)	0.006 (1.132)
Programmes belonging to ministry (1,0)	**-0.029* (-1.803)**	**-0.031** (-2.134)**	0.008 (1.403)	-0.020 (-0.866)	**0.033* (1.958)**	0.012 (1.074)	-0.019 (-1.324)	-0.005 (-0.415)	**0.010* (1.742)**
Programmes related to economy (1,0)	**0.054*** (3.767)**	**-0.033*** (-2.614)**	-0.001 (-0.127)	0.034 (1.547)	**0.028* (1.734)**	0.018 (1.637)	**0.041*** (3.193)**	-0.007 (-0.746)	0.005 (1.031)
Programmes related to welfare (1,0)	**0.066*** (3.991)**	**-0.050*** (-3.108)**	-0.007 (-1.131)	-0.005 (-0.189)	**-0.040** (-2.067)**	**-0.033** (-2.587)**	**0.040*** (2.693)**	**-0.046*** (-3.693)**	**-0.016** (-2.541)**
Programmes implemented SABP two or more times (1,0)	0.001 (0.076)	0.013 (1.175)	**-0.009** (-2.161)**	-0.006 (-0.312)	0.018 (1.378)	-0.008 (-0.884)	-0.016 (-1.514)	0.004 (0.488)	**-0.015*** (-3.508)**
Constant	**0.947*** (22.242)**	**0.989*** (29.606)**	**0.470*** (38.206)**	**1.763*** (23.541)**	**0.825*** (15.909)**	**0.637*** (18.975)**	**1.223*** (30.990)**	**0.918*** (32.461)**	**0.513*** (36.485)**
F	**24.153*** **	**57.120*** **	**74.050*** **	**40.972*** **	**18.754*** **	**24.461*** **	**56.414*** **	**67.475*** **	**72.897*** **
Adjusted R^2	0.166	0.279	0.311	0.312	0.141	0.176	0.213	0.208	0.209
R^2	0.173	0.284	0.315	0.319	0.149	0.184	0.217	0.211	0.212

Notes: 1. Dependent variable is the ROR in performance planning, results, and whole section.

2. Unstandardized coefficient value of each variable and t values are shown in parentheses.

3. * is statistically significant at the 0.10 level, ** at the 0.05 level, and *** at the 0.01 level.

5.5 Comparison of Results Related to the DR and ROR

The previous two sections analysed the factors which affected the DR and ROR. This section compares the significant variables which affected the DR and ROR, and examines whether the results of regression analyses can verify the predictions of the impact of several variables on the dependent variables mentioned in Chapter 4. If the independent and control variables had statistically significant relationships with the DR and ROR for a specific section, then the extent of the impact might be partial. If there were statistically significant relationships with other sections, then the impact might be general. Moreover, if the variables had statistically significant relationships with the DR and ROR for other sections, especially for the whole section, then the extent of the impact could be regarded as strong. Two independent variables strongly supported the predictions: that is, the MOSF's review score had a negative relationship with both the DR and the ROR for all sections in the three different time periods, at the significance level of 0.01; and the specific programme's self-assessment score minus the average score of the spending ministry in that year (specific score minus average score) had a statistically significant impact on both the DR and ROR in a positive direction, at the significance level of 0.01. The impacts of control variables on the DR and ROR are summarised in Tables 5-18 and 5-19. Budget changes for the current year had a partially negative relationship with the ROR, but it was not clear whether this variable had a negative relationship with the DR, so the relationship between this variable and the ROR might partially support the prediction, whilst it was not clear whether the relationship between this variable and the DR supported it. Direct programmes managed by the government had a positive relationship with both the DR and ROR, so the

relationship between this variable and the DR and ROR might support the predictions. However, the extent of these impacts might be partial.

Table 5- 18 Relationships Between Control Variables and the DR

Control variables	2005-2007			2008-2010			2005-2010		
	Plan	Results	Whole	Plan	Results	Whole	Plan	Results	Whole
Budget changes					+**	+**		-*	
Direct programmes	-***	+***	+***	-*			-***	+***	
Small programmes	-**								+***
Large programmes	+**						+**		+***
Ministry programmes	-***	-*		-*	+**		-***		+***
Economy-related									-**
Welfare-related	+***	-***	+***		-**	-***	+**	-***	
Number of SABP		+**				-*	-**		+***

* p<0.1, ** p<0.05, *** p<0.01

Table 5- 19 Relationships Between Control Variables and the ROR

Control variables	2005-2007			2008-2010			2005-2010		
	Plan	Results	Whole	Plan	Results	Whole	Plan	Results	Whole
Budget changes								-*	
Direct programmes		+*						+*	
Small programmes					+***			+***	
Large programmes	+***			+**			+***		
Ministry programmes	-*	-**			+*				+*
Economy-related	+***	-***			+*		+***		
Welfare-related	+***	-***			-**	-**	+***	-***	-**
Number of SABP			-**						-***

* p<0.1, ** p<0.05, *** p<0.01

Small programmes had a positive relationship with both the DR and the ROR, so these relationships might not support the predictions, and the extent of the impacts might be

partial. Large programmes had a positive relationship with the DR and ROR, so this might support the predictions, and the extent of the impacts was completely partial. Programmes belonging to ministries had a positive relation with the ROR, so this might support the prediction, but the extent of the impacts might be partial. However, whether this variable had a positive relationship with the DR was not clear. Economy-related programmes had a negative relationship with the DR, while they had a positive relation with the ROR, so the relationship between this variable and the DR might not support the prediction, while the relationship with the ROR might support the prediction. Considering the extent of the impacts, the former was partial while the latter might be general. As for welfare-related programmes, it was not clear whether this variable had a positive or a negative relationship with the DR. However, this variable might have a negative relationship with the ROR. Consequently, whether the relationship between this variable and the DR supported the prediction was not clear, but the relationship between this variable and the ROR might support the prediction. The extent of the impact of the former seemed to be general, while that of the latter was strong. Similarly, we cannot be sure whether the number of times SABP had been implemented had a positive or a negative relationship with the DR. However, this variable had a negative relationship with the ROR, so the relationship between this variable and the ROR might support the prediction. The extent of the former and the latter might be general.

The predictions, including the effective direction between the variables and the verification results, are summarised in Table 5-20. Two independent variables strongly supported the predictions. As for the relationship between the eight control variables and the DR, just two variables (direct and ministry programmes) supported the

predictions, and two variables (small and economy-related programmes) did not support them, while it was not clear whether four variables (budget changes, ministry, welfare-related and number of times SABP implemented) supported the predictions or not. However, with respect to the relationship between the eight control variables and the ROR, seven variables supported the predictions and just one (small programmes) did not. The specific meaning of these results will be interpreted in Chapter 7.

Table 5- 20 Verification of Predictions

(O: Support, X: Not support)

	Predictions	Sign	DR	ROR
1	MOSF's review score may have a negative relationship with the DR and ROR if there has been a high and persistent optimism bias by spending ministries.	-	O (strongly)	O (strongly)
2	Specific score minus average score may have positive relationship with the DR and ROR if the MOSF has drastically cut spending ministries' generous self-assessment results.	+	O (strongly)	O (strongly)
3	Budget changes may have a negative relationship with the DR and ROR.	-	Not clear (partially)	O (partially)
4	Direct programmes may have a positive relationship with the DR and ROR.	+	O (partially)	O (partially)
5	Small programmes may have a negative relationship with the DR and ROR.	-	X (partially)	X (partially)
6	Large programmes may have a positive relationship with the DR and ROR.	+	O (partially)	O (partially)
7	Ministry's programmes may have a positive relationship with the DR and ROR.	+	Not clear (partially)	O (generally)
8	Economy-related programmes may have a positive relationship with the DR and ROR.	+	X (partially)	O (generally)
9	Welfare-related programmes may have a negative relationship with the DR and ROR.	-	Not clear (generally)	O (strongly)
10	The number of times SABP implemented may have a negative relationship with the DR and ROR.	-	Not clear (generally)	O (generally)

5.6 Dynamic Pattern of the RORs of Spending Ministries

This section analyses the dynamic pattern of the RORs of spending ministries from programmes which were assessed by SABP two or three times between 2005 and 2010. By analysing the dynamic pattern of the RORs, we can identify the strategic behaviours of both spending ministries and the MOSF at any given ROR level, and moreover this might lead to an opportunity to explore the algorithm of the optimum response by spending ministries and the MOSF at any given ROR level. This section uses ROR data rather than DR data, not only because most independent and control variables had statistically significant relationships with the ROR but also because they fulfilled predictions, as analysed in the previous section. For the convenience of the analysis, the study assumes two possible situations: the ROR of a specific programme is higher than that of the spending ministry in that year; and the ROR of a specific programme is lower than that of the spending ministry in that year. From these two possible situations we can develop two plausible predictions.

5.6.1 Two Predictions

Firstly, when the ROR of a specific programme is higher than that of the spending ministry in that year, it is highly plausible that the spending ministry self-assessed the specific programme more generously than other programmes, and the MOSF will drastically cut the spending ministry's self-assessment score, as analysed in 5.4. In that case, it might be plausible for the spending ministry to reduce this programme's self-assessment score in the next SABP round, because there has been a lot of criticism of

this phenomenon. For example, the National Assembly and the news media have criticised the spending ministries overestimated SABP scores. The NABO and several media have questioned the effectiveness of spending ministries' self-assessment scores (NABO, 2010c; Park, 2009; Asiaeconomy, 2009; Joongangilbo, 2009). As well as the spending ministries, the MOSF might also think of big differences between SABP scores as unintended dysfunctional consequences of SABP, so the MOSF might try to reduce the high ROR by raising its own review score in the next SABP. Actually, in order to reduce the big SABP score gap, from 2010 on the MOSF has given penalties to spending ministries where the SABP score gap between the ministry and the MOSF is 20 points or more (MOSF, 2010a, 2011b). From these behavioural expectations of both the spending ministries and the MOSF, the following prediction can be developed:

Prediction 1: If the ROR for a specific programme is higher than the average for the spending ministry, then the spending ministry will reduce its self-assessment score for this programme, while the MOSF will raise its review score, in the next SABP round, and consequently the ROR for the programme will fall in the next SABP round.

Secondly, when the ROR for a specific programme is lower than the average for the spending ministry in that year, it is highly probable that the spending ministry has self-assessed the specific programme more strictly than others of its programmes, and the MOSF will not drastically cut the spending ministry's self-assessment score, as analysed in 5.4. In that case, the spending ministry might try to raise this programme's self-assessment score in the next SABP round, with the intention of matching the programme's self-assessment score with those of other programmes. Conversely, the

MOSF might consider the review score of this programme as more generous than those of other programmes of the spending ministry, so the MOSF will reduce its review score of this programme in the next SABP round. Therefore, it might be possible to make the following prediction:

Prediction 2: If the ROR for a specific programme is lower than the average for the spending ministry, then the spending ministry will raise its self-assessment score for this programme, while the MOSF will reduce its review score, in the next SABP round, and consequently the ROR for the programme will increase in the next SABP round.

5.6.2 Data Analysis

As mentioned earlier, this section analyses the programmes assessed by SABP two or three times, and their distribution can be seen in Table 5-21. Among 2,920 programmes from 2005 to 2010, 2,451 were evaluated by SABP once, 425 twice, and 44 three times. Also, the study examines these programmes by dividing them into two types on the basis of the ROR for specific programme in the first SABP round: the ROR of specific programme was higher than that of the spending ministry; and the ROR of specific programme was lower than that of the spending ministry.

Table 5-21 Distribution of Programmes Assessed by SABP Two or Three Times

(Unit: number of programmes)

Number of SABP		’05	’06	’07	’08	’09	’10	Total
Twice	1st	165	136	113	9	2		425
	2nd			31	144	119	131	425
Three times	1st	18	18	8				44
	2nd			20	13	11		44
	3rd				3	6	35	44

Programmes Assessed by SABP Twice

(1) ROR of specific programme in first SABP round > ROR of spending ministry in first SABP round

Among 425 programmes, the number in this category was 209. For these programmes, the changes in the spending ministries self-assessment score, the MOSF's review score, the ROR, and the t-test results can be seen in Table 5-22.

Table 5- 22 Changes of Self-Assessment Score, Review Score and the ROR (t-test)

(Unit: points, t-value)

	Self-assessment score	MOSF's review score	ROR of specific programme
First SABP	90.17	56.23	0.303
Second SABP	91.28	64.87	0.247
Mean difference	-1.11	-8.64	0.057
t-value	-1.286	-8.777***	4.646***

*** $p < 0.01$

According to Table 5-22, when specific programme's ROR in the first SABP round was higher than the spending ministry's ROR in the first SABP round, the spending ministry tried to increase the specific programme's self-assessment score, which was different from Prediction 1. However, this was not statistically significant, so it is not clear whether the spending ministry increased the self-assessment score of the specific programme in a statistically meaningful way. The MOSF also increased the review score of the specific programme in the second SABP round, and this was statistically significant at the significance level of 0.01. The ROR for these programmes decreased from 0.303 in the first SABP round to 0.247 in the second SABP round, at the

significance level of 0.01. Consequently, Prediction 1 might be said to be generally but not strongly supported, and it might be said that the MOSF tried to reduce the high ROR by raising its review score in the next SABP round.

(2) ROR of specific programme in first SABP round < ROR of spending ministry in first SABP round

Among 425 programmes, the number in this category was 201. Table 5-23 indicates the trend of spending ministries' self-assessment score, the MOSF's review score, the ROR for a specific programme, and the t-test results.

Table 5- 23 Changes in Self-Assessment Score, Review Score and the ROR (t-test)

(Unit: points, t-value)

	Self-assessment score	MOSF's review score	ROR of specific programme
First SABP	87.73	69.52	0.128
Second SABP	91.89	66.82	0.215
Mean difference	-4.16	2.70	-0.086
t-value	-4.340***	-8.777***	-5.281***

*** $p < 0.01$

According to Table 5-23, when specific programme's ROR was lower than the spending ministry's ROR in the first SABP round, the spending ministry tried to increase the self-assessment score of this programme in the second SABP round, while the MOSF decreased the review score. As a result of these behaviours, the ROR for specific programmes increased from 0.128 in the first SABP round to 0.215 in the second SABP round. Moreover, these three results were all statistically significant at the significance level of 0.01. Consequently, Prediction 2 might be said to be strongly supported in a statistically meaningful manner, and this means that the spending ministry had an

incentive to raise a self-assessment score relatively lower than that of other programmes, whilst the MOSF reduced its relatively generous review score.

Programmes Assessed by SABP Three Times

(1) ROR of specific programme in first SABP round > ROR of spending ministry in first SABP round

Among 44 programmes, the number in this category was 22. For these programmes, the trend in the spending ministries' self-assessment score, the MOSF's review score, and the ROR for specific programmes when SABP was implemented three times can be seen in Table 5-24. In addition, Table 5-25 shows whether the behaviours of both the spending ministries and the MOSF were statistically supported by employing the results of Analysis Of Variance between Groups (ANOVA). The spending ministries' self-assessment score decreased from 91.79 in the first SABP round to 90.33 in the third SABP round, as mentioned in Prediction 1, although in the second SABP round the score increased to 93.61. However, the spending ministry's attempt to decrease the self-assessment score was not statistically significant. On the other hand, the MOSF's review score increased greatly from 52.90 in the first SABP round to 64.45 in the third, and it was statistically significant at the level of 0.01. As a result, the ROR for the specific programme fell from 0.334 in the first SABP round to 0.204 in the third, and this was also statistically significant at the level of 0.01. Overall, Prediction 1 was generally but not strongly supported, like the analysis result of programmes assessed by SABP twice, and it implies that the MOSF tried to reduce the high ROR for specific programmes by reducing the review score of the programme, whilst it is not clear

whether the spending ministries tried to reduce a self-assessment score that was more generous than those of other programmes in a statistically significant way.

Table 5-24 Changes in Self-Assessment Score, Review Score, and ROR of Specific Programme

(Unit: points, ratio)

	1st SABP	2nd SABP	3rd SABP
Spending ministry's self-assessment score	91.79	93.61	90.33
MOSF's review score	52.90	67.76	64.45
ROR of specific programme	0.334	0.201	0.204

Table 5-25 Mean Differences between the Self-Assessment Score, Review Score and ROR (ANOVA)

		Sum of squares	d.f	Mean square	F	Sig.
Self-assessment score	Between groups	91.860	2	45.930	0.547	0.582
	Within groups	4033.112	48	84.023		
	Total	4124.972	50			
Review score	Between groups	2070.781	2	1035.390	12.834***	0.000
	Within groups	3872.396	48	80.675		
	Total	5943.177	50			
ROR	Between groups	0.197	2	0.098	6.988***	0.002
	Within groups	0.675	48	0.014		
	Total	0.872	50			

*** $p < 0.01$

(2) ROR of specific programme in first SABP round < ROR of spending ministry in first SABP round

Among 44 programmes, the number in this category was 17. The trend in the spending ministries' self-assessment score, the MOSF's review score, and the ROR for specific programmes can be seen in Table 5-26. In addition, Table 5-27 shows the ANOVA results for these programmes. According to Tables 5-26 and 5-27, the spending ministries greatly increased the self-assessment score of these programmes from 80.55

in the first SABP round to 91.33 in the third, as mentioned in Prediction 2, and this was statistically significant at the level of 0.01. On the other hand, the MOSF increased the review score for these programmes, which was different from Prediction 2. However, this result was not statistically significant, so it is not clear whether the MOSF increased the score in a statistically meaningful way. As a result of these behaviours, the ROR for these programmes increased from 0.138 in the first SABP round to 0.242 in the third, and this was statistically significant at the level of 0.05. Consequently, Prediction 2 might be said to be generally but not strongly supported, and spending ministries might have an incentive to raise the self-assessment score, which was relatively lower than that for their other programmes.

Table 5-26 Changes in Self-Assessment Score, Review Score, and the ROR for Specific Programme

(Unit: points, ratio)

	1st SABP	2nd SABP	3rd SABP
Spending ministry's self-assessment score	80.55	93.96	91.33
MOSF's review score	60.91	65.27	66.33
ROR of specific programme	0.138	0.222	0.242

Table 5-27 Mean differences in Self-Assessment Score, Review Score and the ROR (ANOVA)

		Sum of squares	d.f	Mean square	F	Sig.
Self-assessment score	Between groups	2222.492	2	1111.246	10.808***	0.000
	Within groups	6477.572	63	102.819		
	Total	8700.064	65			
Review score	Between groups	363.013	2	181.506	2.014	0.142
	Within groups	5678.202	63	90.130		
	Total	6041.215	65			
ROR	Between groups	0.134	5	0.018	3.692**	0.030
	Within groups	1.143	63			
	Total	1.277	65			

** $p < 0.05$, *** $p < 0.01$

Summary

Firstly, when, in the first round, the ROR for specific programme was higher than the average for the spending ministry in that year, the ROR for this kind of programme was decreased in the later SABP round in a statistically significant way. During this process, the MOSF raised the review score to reduce the high ROR for specific programme, and this was statistically significant. However, the spending ministries' behaviour was not statistically significant, whether it raised or reduced the self-assessment score of this kind of programme. Secondly, when, in the first round, the ROR for specific programme was lower than the average for the spending ministry in that year, the ROR for this kind of programme was increased in the later SABP round in a statistically significant way. During this process, both the spending ministries and the MOSF behaved as mentioned in Prediction 2 in the programmes assessed by SABP twice. However, in the programmes assessed by SABP three times, the spending ministries raised the self-assessment score for this kind of programme in a statistically significant manner, although the MOSF's behaviour was not statistically significant. Consequently, it might be said that the MOSF tried strongly to reduce a high ROR by raising its review score in the next SABP round. On the other hand, the spending ministries had more interest in increasing a low ROR by increasing their self-assessment score than in decreasing a high ROR by decreasing their self-assessment score. The results of data analysis for the dynamic pattern of RORs are summarised in Table 5-28.

Table 5-28 Summary of the Results of the Dynamic Pattern of RORs

(O: support the prediction, X: not support the prediction)

		Self-assessment score	Review score	ROR	Prediction
SABP twice	ROR of specific programme > ROR of spending ministry	X (statistically not significant)	O (statistically significant)	O (statistically significant)	Accepted (generally)
	ROR of specific programme < ROR of spending ministry	O (statistically significant)	O (statistically significant)	O (statistically significant)	Accepted (strongly)
SABP three times	ROR of specific programme > ROR of spending ministry	O (statistically not significant)	O (statistically significant)	O (statistically significant)	Accepted (generally)
	ROR of specific programme < ROR of spending ministry	O (statistically significant)	X (statistically not significant)	O (statistically significant)	Accepted (generally)

Note: Shaded part was different from the behaviours mentioned in predictions

5.7 Conclusion

This chapter has analysed the factors which affect the DR and ROR of spending ministries. To do this, after ascertaining the differences between the SABP scores and grades of spending ministries and those of the MOSF, it firstly examined the relationship between the independent and control variables and the dependent variables, by employing correlation analysis, scatter diagrams and t-tests to see whether the variables extracted from the relevant literatures actually had some relationship with the dependent variables. The chapter found that all the independent and control variables had statistically significant relationships with the dependent variables through the preliminary statistical analysis, mainly of the six-year pooled data. Secondly, in order to verify the relationships between the variables in detail in the three different time periods, the chapter employed multiple regression analysis. After diagnosing the linear

regression model's goodness-of-fit through residual analyses and multicollinearity tests, such as VIF and tolerance value, it found that all the independent and control variables had statistically significant relationships with the DRs and RORs. Moreover, considering the relatively high R squares and Adjusted R squares (ranging from 0.20 to 0.35), the predictive power and validity of the regression model could be also regarded as high. As for verifying the predictions mentioned in Chapter 4, the two independent variables strongly supported the predictions. All the control variables except for small programmes strongly, generally or partially supported the predictions relating to the ROR. However, with regard to the predictions relating to the DR, only two control variables supported the predictions and two variables did not support them, while it was not clear whether four control variables supported the predictions or not, although they all had statistically significant relationships with the DR. The detailed reasons and meanings of these results will be discussed in Chapter 7. Finally, the chapter also analysed the dynamic pattern of the RORs, in order to explain the strategic behaviours of both spending ministries and the MOSF in relation to any given ROR level from the programmes assessed by SABP twice or three times from 2005 to 2010. The results found that the MOSF had strongly tried to reduce the high ROR by increasing the review scores, whilst the spending ministries had more interest in increasing the low ROR by raising self-assessment scores. The detailed meanings of these results will be also discussed in Chapter 7. The next chapter will analyse the qualitative data for exploring dysfunctional consequences of SABP and their impacts on the behaviours of both spending ministries and the MOSF on the basis of the principal-agent relationship.

CHAPTER 6

QUALITATIVE ANALYSIS FOR DYSFUNCTIONAL CONSEQUENCES OF SABP: A MULTIPLE CASE STUDY

6.1 Introduction

This chapter examines the difference between spending ministries and the MOSF by exploring the existence of, the extent of, and the reasons for dysfunctional consequences of SABP, and the impact of these on the behaviour of both spending ministries and the MOSF. Also, it suggests policy implications for improving the SABP system. It employs a multiple case study from eight budgetary programmes in three case study areas: in the public housing area, construction of public rental housing (P1), subsidy for housing rental costs (P2), and improvement of older public housing stock (P3); in the youth employment area, global youth leader training (P4), global trade specialist training (P5), and junior college students' overseas internships (P6); and in the SOC area, road construction (P7), and railway construction (P8). In order to fully understand the differences between the perspectives and behaviours of spending ministries and the MOSF, the chapter analyses qualitative data from documents, archival records and two rounds of in-depth interviews with Korean government officials and experts by using N-

vivo 9, the qualitative data analysis software. Most findings of the chapter are the result of the first interviews, and the supplementary interviews focus on the impact of dysfunctional consequences and possible policy alternatives. The existence of dysfunctional consequences is to be explored on the basis of the theoretical framework developed in Chapter 2: some dysfunctional consequences are "unintended" by both spending ministries and the MOSF, while others are "unintended" by the MOSF but are "intended" by spending ministries, as "agents" in the principal-agent relationship. When it comes to looking at the extent of dysfunctional consequences, the chapter pays special attention to the differing perceptions between spending ministries and the MOSF on each dysfunctional consequence. Reasons for dysfunctional consequences are to be analysed not only by internal reasons, such as SABP design and public programmes' characteristics, but also by external reasons, like the unique Korean government culture, the information asymmetry between spending ministries and the MOSF, and time and cost. The chapter also verifies that dysfunctional consequences have an influence on both spending ministries' optimism bias and the MOSF's strict control. Finally, policy alternatives that would produce a better SABP system by reducing or preventing these dysfunctional consequences are suggested according to three criteria: those that are urgent; those that should be implemented in the short-term; and those that can be implemented in the long-term.

6.2 Six Key Themes

From the first 32 interviews with government officials, 868 pieces of coded data were initially created from eight programmes in three case study areas and then, finally, the

data were classified into six key themes for exploring dysfunctional consequences of SABP. The number of codings (references) in each case area was similar, at around 300. Interestingly, unlike in the pilot interviews, there was no significant difference in perspectives on SABP between spending ministries' budget divisions and their programme divisions, so this was not dealt with as a separate key theme. By examining these six themes and their sub-themes, it was possible to explore the types of, the extent of, and reasons for dysfunctional consequences of SABP, as well as the influences of these dysfunctional consequences on the SABP results of both spending ministries and the MOSF. The six key themes were all closely linked to the conceptual framework mentioned in Chapter 2 and, among them the impact of dysfunctional consequences was the result of my interpretation of the first interviews, because the topic guide did not include explicit questions related to it. However, an explicit question was included in the supplementary interviews, and the results of these interviews will be explored later. Dysfunctional consequences of SABP had the largest number of codings, because the theme included the types of, extent of, and reasons for dysfunctional consequences. Policy alternatives had the second largest number of codings, because the category contained various kinds of alternatives. As for the background questions, most interviewees gave general opinions on the consequences of SABP. Although the number of codings related to the strategy of the MOSF and spending ministries and to opinions on spending ministries' efforts to prevent dysfunctional consequences was rather small, by repeatedly reviewing the results, I was able to draw out common patterns from the eight programmes in three case study areas. The six key themes can be seen in Table 6-1.

Table 6- 1 Coding and Frequency of Six Key Themes in the Three Case Study Areas

(Unit: number of coding, %)

Six key themes	Public housing		Youth employment		SOC		Total	
	Coding	Frequency	Coding	Frequency	Coding	Frequency	Coding	Frequency
General opinions on the consequences of SABP	68	24.7	51	16.3	39	13.9	158	18.2
Strategies of the MOSF and spending ministries	20	7.3	12	3.8	9	3.2	41	4.7
Dysfunctional consequences of SABP	116	42.2	143	45.8	153	54.4	412	47.5
Impact of dysfunctional consequences	9	3.3	20	6.4	14	5.0	43	5.0
Opinions on spending ministries' efforts to prevent dysfunctional consequences	13	4.7	12	3.9	14	5.0	39	4.5
Policy alternatives	49	17.8	74	23.8	52	18.5	175	20.1
Total	275	100.0	312	100.0	281	100.0	868	100.0

6.3 General Opinions on the Consequences of SABP

Most respondents gave an opinion as to whether the SABP system had generally achieved its aims or not, but the opinions were divided into positive and negative by a similar number of references from a similar number of interviewees (22 interviewees with 31 references had positive opinions, while 23 interviewees with 37 references were negative). Some interviewees commented both positively and negatively in the same interview. In each case study area, the number of respondents and references giving positive responses was similar to the number giving negative ones. The six experts showed a similar division of opinions: five experts making five references were positive; while four experts making seven references were negative. The reports of the NABO analysing the government's performance management also commented on both positive and negative aspects of SABP (NABO, 2007a, 2008, 2009a, 2010c, 2011). The links between the SABP results and budgeting on the one hand, and specialised and fair

evaluation for the public sector on the other, were indicated as the main reasons why the government achieved its aims. However, the penalty-oriented operation of SABP, the intrinsic limitations of the performance budgeting system in the public sector (e.g., difficulties in measurement, indirect results, and the time taken to achieve performance, etc.), the lack of interest by senior officials, and the lack of specialty in both spending ministries and the MOSF were indicated as the main reasons for the programmes not achieving optimum performance. Consequently, it can be said that the SABP system has increased the efficiency of budget allocation and the accountability of spending ministries through linking SABP results with budgeting, but at the same time the SABP system needs to be improved for an even better performance budgeting system. The number of respondents and references in the three case study areas can be seen in Table 6-2.

Table 6- 2 General Opinions on the Consequences of SABP in the Three Case Study Areas

(Unit: number of coding and person)

Main categories	Public housing			Youth employment			SOC			Total		
	Coding	MOSF	Spending	Coding	MOSF	Spending	Coding	MOSF	Spending	Coding	MOSF	Spending
Achieves its aims	11	3	3	11	4	5	9	2	5	31	9	13
Links SABP results and budgeting	7	3	2	7	4	3	3	2	1	17	9	6
Specialised and fair evaluation	2	1	1	4	3	1	3	2	1	9	6	3
Doesn't achieve aims	13	2	4	11	5	4	13	4	4	37	11	12
Excessive workload	4	1	3	2	1	1	-	-	-	6	2	4
Intrinsic limitation	1	-	1	3	1	2	3	1	2	7	2	5
Lack of discretion	3	1	2	1	-	1	-	-	-	4	1	3
Lack of senior officials' interest	4	2	1	1	-	1	-	-	-	5	2	2
Lack of specialty	4	1	2	-	-	-	1	1	-	5	2	2
Lack of trust	3	1	2	2	1	1	1	1	-	6	3	3
Penalty-oriented operation	8	2	5	5	2	3	4	1	2	17	5	10
Spending ministries' tactics	2	1	-	1	1	-	-	-	-	3	2	-
Others	6	3	3	3	1	2	2	-	2	11	4	7

6.4 Strategies of the Spending Ministries and the MOSF

Although questions about the strategies of the spending ministries and the MOSF could not be asked of all interviewees, because of the interviewees' busy schedules, I could discern the main characteristics of these strategies through the three case study areas. The formal strategy of the MOSF was closely related to SABP aims of increasing the efficiency of budget allocation and improving the performance of a programme (six references). This formal strategy had been taken for granted by MOSF officials and this had led to the strict application of the SABP checklist to spending ministries' self-assessment results and finally resulted in "drastic cutting" of the latter.

> "Strategically, the MOSF has put emphasis on budget efficiency by using the SABP results, even though sometimes spending ministries have appealed against the MOSF's drastic cutting of their results." (Manager, MOSF, interviewed on 9th November 2011)

However, interestingly, the Budget Office of the MOSF sometimes asked the Performance Management Bureau of the MOSF, informally, not to lower the SABP results as much as they could be lowered in the case of programmes on the presidential agenda (six references). The experts also pointed out the MOSF's informal strategy towards the presidential agenda.

> "In respect of the presidential agenda, both spending ministries and the MOSF are penalised for the award of poor grades, and this leads to disruption in the consistency of evaluation." (Expert, KIPF, interviewed on 4th November 2011)

On the other hand, the Performance Management Bureau of the MOSF showed a different attitude in its reactions towards the Budget Office's request, trying to keep to the principles of SABP: more budget resources going into efficient programmes; fewer into poor quality programmes. Moreover, the National Assembly and the BAI insisted that the Performance Management Bureau of the MOSF explain its reasons in cases where it did not obey the SABP principle, and this surveillance by external organisations seemed to strengthen the MOSF's principles.

> "Sometimes, if the programme was related to the presidential agenda, I would informally ask the officials of the Performance Bureau in my ministry to at least not give a poor grade…. However, the Performance Bureau tried to stick to their principles, despite the request from their own ministry." (Manager, MOSF, interviewed on 27th October 2011)

When both the formal and informal strategy of the MOSF were considered, it seemed that the MOSF generally focused on economic rationalisation in line with the principles of SABP, and consequently it tended to take a strict view of the spending ministries' self-assessment results.

With respect to the spending ministries' formal strategy, these ministries focused on developing an appropriate performance indicator and performance target levels (six references), and providing satisfactory explanations to the MOSF (nine references). Informally, they tried to keep in constant contact with MOSF officials (six references) and sometimes actively indulged in tactics such as changing the performance indicator itself and avoiding the development of a performance indicator that would be disadvantageous to them (four references). These tactics might be related to

"misrepresentation" by spending ministries, one of the dysfunctional consequences that will be explained later. Spending ministries' formal and informal strategy was completely based on opportunism and a maximisation of their performance, in order to obtain high SABP results and as big a budget as possible. These strategies could be seen as dysfunctional consequences of the SABP system and as leading, eventually, to an optimism bias when it came to self-assessment.

> "Formally, I tried to choose the performance indicator which would most easily obtain a high SABP score, such as an objectives- or outcome-oriented indicator." (Manager, MLTM, interviewed on 27th October 2011)

> "Informally, some officials tried to ask MOSF officials they were acquainted with to help them to obtain better than a 'poor' grade…" (Assistant Manager, MEST, interviewed on 2nd November 2011)

Table 6-3 presents the number of codings related to the formal and informal strategies of the MOSF and spending ministries in eight programmes in three case study areas.

Table 6-3 Formal and Informal Strategies of the MOSF and Spending Ministries

(Unit: number of coding)

Main categories			Public housing	Youth employment	SOC	Total
MOSF						
	Formal strategy	Focusing on budget efficiency	2	1	3	6
	Informal strategy	Not lowering the score of programmes on the presidential agenda	3	2	1	6
		Intentional drastic cutting	-	1	-	1
Spending ministries						
	Formal strategy	Report to the Minister	2	-	-	2
		Self-restructuring means	-	1	-	1
		Study performance indicator and target level	3	2	1	6
		Sufficient explanation	3	3	3	9
	Informal strategy	Keeping informal contact	5	1	-	6
		Using some tactics, such as changing performance indicator	2	1	1	4

6.5 Types of, Extents of and Reasons for Dysfunctional Consequences of SABP

This section explores the existence of, the extent of and the reasons for the diverse dysfunctional consequences of SABP mentioned in the conceptual framework. These were drawn from eight programmes in three case study areas by analysing related documents (e.g., research memos, the MOSF's final SABP reports, the National Assembly and the BAI's evaluation reports and new releases, etc.) and the results of the first round of interviews.

6.5.1 Tunnel Vision

According to the SABP manual (MOSF, 2010a), the MOSF puts emphasis on non-quantifiable as well as quantifiable performance[24], that is, in Question 2-1 the MOSF asks whether spending ministries have developed a quantifiable outcome performance indicator, and in Question 4-2 it asks them whether they have implemented comprehensive and objective programme evaluation to assess the non-quantifiable, or the qualitative, performance of a programme. In this regard, in the case studies in the public housing area, P1 and P2 had developed only an evaluation of the quantifiable performance of a programme[25], which was not related to an outcome, and had not implemented comprehensive and objective programme evaluation. Thus, the MOSF

[24] Question 2-1. Are there definitive relationships between the performance indicator and the purpose of the programme? 4-2. As a result of programme evaluation, is the programme operating effectively?

[25] The performance indicator for P1 is "public rental housing construction rate"; for P2 it is "budget implementation rate"; and for P3 it is "rate of budget grants to local government".

gave a "No" answer to Question 4-2 for the two programmes (MOSF, 2008a, 2010d). However, interestingly, in Question 2-1 the two programmes were given a "Yes" answer, although performance indicators seemed to focus only on quantifiable performance, not on outcome and, moreover, for P1 and P2, Question 2-1 was firstly reviewed with a "No" by the MOSF's Budget Office, but finally changed into a "Yes" by the MOSF's Performance Management Bureau. In this sense, the question of how P1 and P2 could finally get a "Yes" answer, and the question of whether the MOSF could ignore political pressure completely when it came to reviewing the spending ministries' self-assessment results, might need to be considered further in Chapter 7. In the case of P3, the MOSF gave a "Yes" response to Question 2-1, because the programme's main purpose was to subsidise the budgets of local government, and this was regarded as a quantifiable outcome-related indicator. However, P3 also received by a "No" answer to Question 4-2, because there was no profound programme evaluation. Backing up the MOSF's report, most respondents agreed that the spending ministries had a tendency to focus on the quantifiable performance of a programme (nine interviewees, with references). In the youth employment area, the MOSF also stated that although one of the two performance indicators for P4 (employment rate) was related to the outcome, the performance indicators for P5 and P6 only dealt with quantifiable performance[26], none of the three programmes presented full programme evaluation (MOSF, 2010c, 2010e, 2010f). Consequently, in Question 2-1, P4 was given a "Yes" answer and P5 received a "No" answer. In the case of P6, the MOSF gave a "Yes" answer on the

[26] The performance indicators for P4 are "the number of students who completed the programme" and "employment rate"; for P5 they are "the number of collaborating companies", "accretion rate of academic credit" and "matching fund rate"; and for P6 they are "improvement rate of language skills" and "degree of student satisfaction". The degree of satisfaction in P6 can be seen as a kind of qualitative performance indicator.

condition that the spending ministries would develop outcome-oriented performance indicators later. The NABO stated that the global trade specialist training programme, P5, used only quantifiable performance indicators, and there were no indicators of convincing, substantial performance (NABO, 2009a). Most interviewees also said that spending ministries tended to pay attention to quantifiable performance, but not to interesting non-quantifiable performance (nine interviewees, with nine references). In contrast to the previous two case study areas, two programmes, P7 and P8[27], in the SOC area had implemented comprehensive and objective programme evaluations for non-quantifiable performance (MOSF, 2010d). Thus, at first glance, tunnel vision seemed not to exist in the SOC area. However, it is compulsory for a programme for which the budget is over 50 billion Won to implement comprehensive evaluation, such as PFS, according to the MLTM's construction technology management regulations. This could imply that the MLTM implemented programme evaluation to obey the relevant law, not as a result of a voluntary decision. In addition, Question 2-1 for P8 received a "Yes" answer on the condition that the MLTM would develop outcome-related performance indicators later, as for P6. Moreover, the BAI report indicated that the road construction programme's performance indicators were not suitable evidence of the final outcome of the programme, and merely represented the programme's output (BAI, 2011a). Nine out of ten interviewees, making 13 references, in the SOC area also answered that spending ministries focused only on quantifiable performance.

> "I think spending ministries have a strong tendency to regard quantifiable performance as more important…" (Manager, MOSF, interviewed on 28th October 2011)

[27] Performance indicators for P7 are "budget implementation rate", "road extension rate" and "traffic capacity"; and for P8 it is "double-track rate".

> "In the case of the railway construction programme, quantifiable performance indicators such as construction rate or double-track rate were mainly used as performance indicators." (Manager, MMTM, interviewed on 11th November 2011)

Therefore, in the SOC area, the fact of implementing comprehensive and objective programme evaluation might not be critical evidence of an absence of tunnel vision. Rather, the interviews, the BAI report, and the compulsory regulations for the industry showed that the spending ministry had a tendency to focus on quantifiable performance. As for the overall interview results, a total of 27 interviewees (15 spending ministry officials, 12 MOSF officials) making 37 references agreed that tunnel vision existed, and there was no disagreement. All six experts also agreed that tunnel vision existed. From the above evidence, it might be said that both spending ministries and the MOSF strongly agreed that this phenomenon existed.

Most respondents argued that high SABP scores and consequent bigger budgets, increasing reliability and fairness, and ease of evaluation were the main reasons for this phenomenon. High SABP scores and bigger budgets represented a different expression of a spending ministry's maximisation strategy; and, moreover, this was a common reason for all dysfunctional consequences. Also, spending ministries considered that a quantifiable performance indicator had more advantages than a non-quantifiable indicator in terms of predictability, management, and responsibility. Apart from the spending ministries' maximisation strategy, the other reasons were all related to internal characteristics of the SABP design itself, rather than to factors in the external environment, such as information asymmetry or the unique Korean government culture.

6.5.2 Myopic Management

In order to prevent the risk of a short-term perspective, a long-term performance indicator is regarded as a core performance indicator for SABP. So, SABP has been conducted in three-year cycles, and the assessment of new programmes can be delayed until they begin to demonstrate their performance (MOSF, 2010a). Moreover, the MOSF requires spending ministries to develop an outcome performance indicator that relates to the long-term strategy and final goal of a programme (MOSF, 2011c). However, according to the MOSF's report, only P3, P4 and P7 submitted a fully developed outcome performance indicator, whilst the other five programmes (P1, P2, P5, P6, P8) submitted a simple output performance indicator (MOSF, 2010c, 2010d, 2010e, 2010f). Nonetheless, interestingly, only P5 received a "No", and seven programmes received a "Yes", in answer to Question 2-1. P1 and P2 first received a "No" answer, which was eventually changed into a "Yes" answer by the MOSF, and P6 and P8 finally received a "Yes" on the condition that the spending ministries would follow the MOSF's recommendation that they develop a full outcome performance indicator later, as mentioned in 6.5.1 (MOSF, 2010d, 2010e). The reason and implication will be discussed in Chapter 7. Besides the MOSF's reports, the NABO drew attention to the possibility of spending ministries' pursuit of short-term performance rather than long-term strategy (NABO, 2007b, p.119):

> "A performance indicator intrinsically evaluates present effort to achieve performance, so there must be a time-lag between present effort and an eventual outcome. Thus, emphasis on performance can provide a spending

> ministry with an incentive to pursue short-term achievement rather than a long-term and strategic goal."

Most interviewees responded that spending ministries paid more attention to short-term performance than to the long-term, and this phenomenon was the most cited dysfunctional consequence across all three case study areas (28 interviewees, with 38 references). Sixteen out of 19 spending ministry officials and 12 out of 13 MOSF officials agreed that this phenomenon existed, and there was no disagreement. One interviewee even said that spending ministries sometimes tried to purse extremely short-term performance indicators for a programme.

> "…comprehensive government policies have been released in two- or three-month cycles, and this phenomenon may be a good example of how much spending ministries are interested in extremely short-term performance." (Manager, MLTM, interviewed on 31st October 2011)

In addition, the six experts all commented on spending ministries' myopic tendencies. Overall, it might be said that both spending ministries and the MOSF strongly agreed on the existence of spending ministries' myopic management.

As for the reason for this phenomenon, high SABP scores and bigger budgets and the unique Korean government culture were the most cited across the three case study areas. Spending ministries were more interested in next year's budget than in the budget for two or more years later. Also, the unique Korean government culture, especially the general one-year rotation personnel system and ministers' short tenure of office, have greatly influenced spending ministries' myopic perspective.

> "The priority of policies in the Korean government is easily changed by high level government officials and the President. Moreover, the one-year rotation personnel system has reinforced the tendency to focus on short-term achievement." (Expert, KIPF, interviewed on 4th November 2011)

> "Spending ministries have mainly been interested in short-term performance and this is closely related to the unique Korean culture in politics and government. People select the President on the basis not of policy capability but of political tendency. So, the President often appoints non-experts to be ministers and changes ministers after about a year. This kind of minister also changes high ranking government officials…" (Expert, KDI, interviewed on 8th November 2011)

In addition, lack of time for building up long-term performance and the characteristics of public programmes, such as unpredicted, indirect, spillover effects and small assessment levels ("sub-programme" levels in the programme budgeting structure), were also cited as reasons for myopic management. The characteristics of public programmes and small assessment levels are related to internal reasons of SABP, while the unique Korean government culture and problems of lack of time are associated with the external environment of the SABP system.

6.5.3 Measure Fixation

The MOSF recognises the possibility of this phenomenon, so it requires spending ministries to develop representative performance indicators closely related to the desired outcomes for a programme, in order to prevent the problem caused by paying attention only to a measurable performance indicator (MOSF, 2010a). However, according to the MOSF's report, P5, P6 and P8 did not develop representative performance indicators,

and so P5 was finally given a "No", and P6 and P8 finally received a "Yes" in answer to Question 2-1, on the condition that the spending ministries would follow the MOSF's recommendation that they develop representative performance indicators later (MOSF, 2010c, 2010d, 2010e). The NABO was also aware of this problem, so it investigated whether the spending ministries ever submitted representative performance indicators for these programmes; and it continued to point out the lack of representative performance indicators every year (NABO, 2007a, 2008, 2009a, 2009b, 2010d, 2011). For P2, the NABO indicated the necessity of dividing the performance indicator into subsidies for purchasing costs and those for rental costs, in order to fully represent the programme (NABO, 2011). For P4, the NABO pointed out that "the failure rate at the halfway stage" should be added as a new performance indicator, in order to represent the programme's purpose, alongside "the employment rate" (NABO, 2011). The NABO also commented that P5's performance indicators were not closely related to the purpose of the programme, so they were not representative performance indicators (NABO, 2009b). Across the three case study areas, most respondents answered that the spending ministries focused on measurable performance (20 interviewees with 20 references). Ten out of 13 MOSF officials and ten out of 19 spending ministry officials agreed with this. Three out of six experts also mentioned the existence of measure fixation, whilst no experts commented on an absence of measure fixation.

> "Furthermore, in order to obtain bigger budgets, programmes have to be graded above 'modest', at least, so spending ministries have no choice but to depend on output-oriented, short-term, and measurable performance." (Expert, BAI, interviewed on 2nd November 2011)

However, interestingly, five interviewees from spending ministries did not fully agree

with this judgement. The five interviewees all argued that they paid attention not only to measurable performance but also to non-measurable performance, because they thought that non-measurable performance was closely related to their programme's ultimate goal.

> "I think I paid attention to non-measurable performance in my own way, such as trying to develop an informal additional performance indicator which could reflect the road construction programme's performance well. For example, I tried to add 'volume of traffic per hour' as a new performance indicator, because I thought this indicator could measure the quality of road service well, although I needed more time." (Manager, MLTM, interviewed on 31st October 2011)

Also, characteristics of public programmes such as intangibility, indirectness, externality, and complexity (Wilson, 1989; Thiel and Leeuw, 2002) impacted on spending ministries' tendency to focus on the measurable performance of a programme. Spending ministries insisted that they had no choice but to pursue measurable performance, due to the difficulty of managing and assessing non-measurable performance. Considering spending ministries' primary goal under SABP was to obtain high SABP scores and bigger budgets, rather than improving the genuine performance of a programme, spending ministries had less incentive to be interested in non-measurable performance. Overall, the MOSF might strongly agree with this verdict, while spending ministries might partly agree with it. The MOSF's suggestion of developing representative performance indicators in order to prevent the problems caused by measure fixation seemed not to be actively accepted by spending ministries.

High SABP scores and bigger budgets were the reasons most frequently cited for

measure fixation, and the characteristics of public programmes were mentioned as another reason. Also, lack of time to consider non-measurable performance was cited as one of the reasons for measure fixation. Apart from high SABP scores and bigger budgets, the characteristics of public programmes might be related to the internal logic of SABP, and the lack of time might be related to SAPB's external environment.

6.5.4 Suboptimisation

Across the three case study areas, a total of 15 interviewees (eight MOSF, seven spending ministry), making 17 references, gave positive opinions on the existence of this phenomenon. However, three of the spending ministry officials and one of the MOSF officials did not agree that spending ministries focused on partial objectives rather than the full objectives of the organisation. The interesting thing was that one of the MOSF interviewees denied the existence of the phenomenon. However, an even more interesting point was that the reason for the MOSF official's denial was rather different from that of the spending ministries: that is, this official emphasised the involuntary aspect of choice by spending ministries, whilst the spending ministries stressed their voluntary decision-making (see the underlined sections of the following).

> "…in the public housing area, the spending ministry seemed to pursue the whole objective or the first best solution of the organisation, because the performance indicator for this area had been passed in advance in a Cabinet meeting, and stipulated the provision of 1.5milion public housing homes in 10 years. Thus, I think the spending ministry had no discretion when it came to pursuing their objective." (Manager, MOSF, interviewed on 11th November 2011)

> "I think my division has pursued the whole objective of my ministry, and of the government as a whole, by focusing on building a road that was vitally necessary, prioritising the completion of this road, and increasing investment efficiency." (Manager, MLTM, interviewed on 31st November 2011)

> "I tried to pursue the overall objectives of the organisation, and if necessary I changed the performance indicators to do it." (Manager, MLTM, interviewed on 11th November 2011)

Four of the six experts also argued that spending ministries had a tendency to pursue partial objectives rather than whole ones. They indicated that spending ministries focused on partial objectives in order to avoid the risks of measurability and responsibility, and sometimes they tried to play tactical games by questioning the evaluator's expertise in their programme.

> "…in private road construction, which contains the obligatory 'loss of compensation' rule in the case of falling short of the contracted traffic volume or passengers, the spending ministry used the process rate of construction as a performance indicator instead of traffic volume or the number of passengers." (Expert, KIPF, interviewed on 4th November 2011)

> "I think in case of difficulty in achieving the first best solution or the overall objectives of the organisation, spending ministries tend to pursue the second best solution, or the narrower objectives, at any time." (Expert, University, interviewed on 9th November 2011)

Judging from the fact that many important decisions on public housing or the SOC area had been made in advance by cabinet meetings, regardless of what the spending ministries wanted, I was given the impression that the MOSF strongly agreed that suboptimisation by spending ministries existed. Spending ministries too might partly agree that this phenomenon existed.

High SABP scores and bigger budgets, and the unique Korean government culture, especially the one-year rotation personnel system, the sectionalism both within ministries and among ministries, the ministers' short tenure of office, and the excessive emphasis on the presidential agenda, were indicated as the main reasons for spending ministries' tendency towards suboptimisation. Also, as the experts mentioned, efforts to avoid the risk of responsibility and accountability might be other reasons for this phenomenon. Apart from high SABP scores and bigger budgets, all these reasons seemed to be closely related to the external environment in which SABP operated.

6.5.5 Ratchet Effect

Shin and Cha (2010) argue that although a little subjectivity is inevitable when it comes to setting a performance target level, the grounds for setting the level should be explicitly stated, to enable performance information users to assess the validity of the target level. They indicate that reasonable grounds for the target level of a programme should include comparison with the results of similar programmes besides the past trend of the programme's achievement. In this regard, the SABP manual (MOSF, 2010a) requests spending ministries to submit reasonable grounds for choosing a performance target level as well as providing the target level itself, in order that the appropriateness of the target level may be evaluated. The NABO (2009b) has stressed the importance of ambitious performance target levels, because low target levels could lead to a fall in the standard of responsibility and reliability in government activities. However, across the three case study areas, all the programmes except for P4 showed strong signs of the

ratchet effect. In the public housing area, all three programmes were given a "No" by the MOSF's review, because P1's target level was lower than that for the previous year, and moreover there was no reasonable explanation; P2's level was a mere reflection of past trends; and P3's 167% achievement rate was considered to be evidence of an unambitious target level (MOSF, 2010d). Nine interviewees (three MOSF officials, six spending ministry officials) making 11 references also agreed that this phenomenon existed. In the youth employment area, P4's target level was accepted as ambitious by the MOSF because this programme's goal had been determined from 2008 on by agreements between related organisations such as several ministries, economic institutions and university associations (MOSF, 2010e). This could be taken as evidence that continuous negotiations on performance target levels could solve the problems and risks of uncertainty and information asymmetry among the parties involved. On the other hand, P5's target level was given a "No" by the MOSF (2010c); and the NABO (2009b) also considered this programme's target level to be unambitious because "the accretion rate of academic credit" could be achieved automatically if participants finished the programme successfully and "the number of collaborating companies" was the same as in the previous year's target. In P6, the spending ministries did not submit a reasoned explanation for the target level, and just suggested 5% more than in the previous year (MOSF, 2010f). In addition, 11 interviewees (four MOSF, seven spending ministry) with 11 references mentioned this phenomenon. In the SOC case study area, P7 and P8 were given a "No" by the MOSF review because P7's performance target level was changed after the implementation of the programme, and no reasoned and specific explanation was submitted for the level set for P8 (MOSF, 2010d). Eight interviewees (four MOSF, four spending ministry) making ten references also agreed

that a ratchet effect existed. In terms of the results of interviews across the three case study areas, the highest number of interviewees acknowledged the existence of this phenomenon (28 interviewees with 32 references), the same as for myopic management. Interestingly, 17 out of 19 spending ministry interviewees were honest enough to accept the existence of the ratchet effect. Four out of the six experts also commented that spending ministries had no incentive to suggest highly ambitious target levels and tended to submit somewhat conservative targets. The above evidence suggests that this phenomenon was strongly accepted by both the MOSF and the spending ministries.

> "In the case of a programme graded 'poor', it was very difficult to obtain appropriate budgets for the next year; so attainability was the first criterion for setting the performance target level…" (Manager, MLTM, interviewed on 27th October 2011)

> "I think spending ministries absolutely did not suggest highly ambitious performance targets for their programmes, but just submitted very conservative performance target levels..." (Manager, MOSF, interviewed on 11th November 2011)

> "A performance target level is, of course, conservative, not ambitious, because the spending ministry has to consider the next year's target level." (Assistant Manager, MOEL, interviewed on 31st October 2011)

High SABP scores and bigger budgets were the reason most cited for the ratchet effect. The next most cited reason was the penalty awarded for poorly graded programmes. Another interesting reason was the information asymmetry between spending ministries and the MOSF. Spending ministries tended to use the MOSF's lack of specific knowledge of their programmes.

> "In addition, the MOSF seems to evaluate the validity of spending ministries' performance target levels on the basis of rational grounds for a target level; but I think the MOSF is not able to obtain accurate information about what rational grounds are, because the MOSF cannot know a programme as well as the spending ministry." (Manager, MKE, interviewed on 28th October 2011)

The last reason mentioned was programme supervisors' concern over highly ambitious target levels. Apart from high SABP scores and bigger budgets, all the reasons mentioned by interviewees were closely related to the external environment of SABP.

6.5.6 Misrepresentation

The reports of the MOSF, the NABO and the BAI indicated the possibility of distortions or manipulations by spending ministries seeking to avoid disadvantageous performance data. According to the MOSF's review of P7, the MLTM submitted a performance target level to the MOSF after the programme had already been implemented, so the MOSF did not accept the suitability of the performance target level (MOSF, 2010d). The NABO (2010c) suggested various types of misrepresentation, such as inadequate performance reports and distortion of programmes' performance. The NABO (2011) also commented that P4 needed an additional performance indicator giving the percentage of students who completed this programme as well as the number of students completing it, in order to measure performance more accurately. Moreover, the NABO (2011) stated that the performance indicator for the subsidy for housing rental programme (P2) should be divided to show the budget implementation rate for this programme and that for the programme for subsidy for housing purchase, to which it is closely related. Regarding P1 and P2, the BAI (2011a, 2011b, 2009) also indicated that

spending ministries measured their performance by using the methods they favoured, not the ones intended by the BAI, and that these ministries reported irrelevant facts or facts relating to past achievements, and often failed to report on the performance currently under consideration. Across the three case study areas, a total of 17 interviewees (11 MOSF, six spending ministry) making 17 references responded positively to the idea that spending ministries tended to overestimate their performance as much as possible and to hide disadvantageous performance data, whilst a total of ten interviewees making ten references disagreed with this. These ten interviewees were all spending ministry officials. The main reason for not accepting this judgement was the continuous stream of requests for performance data from the MOSF, the National Assembly, and the press. In other words, it was very difficult for spending ministries to hide a disadvantageous performance or overestimate results because the organisations mentioned persistently demanded exact performance data. However, six interviewees from spending ministries agreed that they had a tendency towards misrepresentation, although the extent of this tendency was much weaker than the MOSF was suggesting. The MOSF officials argued strongly for the existence of the tendency, using words such as "basically", "never", "definitely", while the spending ministry officials used more vague, defensive words, like "sometimes", "especially when", "if possible" (see the underlined sections of the following quotations).

> "…if I found some disadvantageous data in the process of evaluation, I would not submit this data. Instead, I sometimes gave other, subsidiary information to the MOSF, explaining that this data was more important…" (Manager, MLTM, interviewed on 27th October 2011)

> "The results section has the biggest proportion of SABP marks, so a spending ministry might sometimes have an incentive not to provide some

> disadvantageous performance data to avoid a poor grade. In addition, sometimes a spending ministry puts performance data together to hide the low performance of a programme." (Assistant Manager, MEST, interviewed on 2nd November 2011)
>
> "I think spending ministries have an incentive to overestimate their performance, especially when the performance indicator is more related to the quality and not quantifiable." (Manager, MLTM, interviewed on 11th November 2011)
>
> "Basically, spending ministries tried to hide disadvantageous data and they never submitted them. They tended to turn in only compulsory data…" (Senior Manager, MOSF, interviewed on 3rd November 2011)
>
> "Spending ministries never evaluated their performance accurately, and always tried to overestimate it…" (Manager, MOSF, interviewed on 9th November 2011)
>
> "Spending ministries had a strong tendency to overestimate their performance and deliberately did not include disadvantageous data." (Manager, MOSF, interviewed on 11th November 2011)

Five of the six experts also commented that spending ministries basically tried to conceal their disadvantageous performance data to avoid taking responsibility for it, and moreover they also tried to avoid setting a sensitive performance indicator that might force them to submit disadvantageous performance data. Taken together, the above pieces of evidence suggest that the MOSF might strongly agree with the existence of misrepresentation, while the spending ministries might partly agree with it.

High SABP scores and bigger budgets, information asymmetry, and the awarding of penalties were commented on as the reasons for this phenomenon. Spending ministries always wanted to obtain high SABP scores and bigger budgets, so they had an incentive to overestimate their performance and to hide disadvantageous performance data.

Information asymmetry and the awarding of penalties related to SABP could strengthen the incentive. Apart from high SABP scores and bigger budgets, the other two reasons were both related to the external environment of SABP.

6.5.7 Cherry-Picking

The NABO indicated that in order to increase the validity of P2 it had been necessary to adjust the income level at which people qualified for a subsidy, and although one of the purposes of P6 was to alleviate income polarisation among students by supporting low-income students, this programme had chosen students only on the basis of grades and language skills, regardless of their income or standard of living, in order to attain its goal more easily (NABO, 2011, 2009b). From the interviews, I learned that there was a great difference between the spending ministries and the MOSF in their perspective on whether spending ministries gave adequate consideration to priority groups in their programmes. Those recognising the cherry-picking phenomenon were all MOSF officials (seven interviewees with seven references), while those denying it were all spending ministry officials (ten interviewees with references). However, the argument of the MOSF interviewees was not as strong as that of the spending ministries. With respect to cherry-picking, the MOSF might partly agree that it existed, whilst the spending ministries strongly denied it (see the underlined sections of the following).

> "I partly agree that the spending ministries have sometimes tried to select priority groups on the basis of how far these helped them to attain their performance target…" (Manager, MOSF, interviewed on 27th October 2011)

> "There was a suspicion of the cherry picking phenomenon, but no supporting evidence." (Senior Manager, MOSF, interviewed on 3rd November 2011)

> "In the case of youth employment, this has been a very sensitive social issue, so it was not possible to choose priority groups more accurately." (Assistant Manager, MEST, interviewed on 2nd November 2011)

> "I never considered the attainability of performance targets when it came to selecting priority groups." (Manager, MLTM, interviewed on 11th November 2011)

The MOSF officials stated that spending ministries had a tendency to cherry-pick in order to obtain high SABP scores, and consequently bigger budgets for the following year. On the other hand, spending ministry officials answered that it was not possible to choose priority groups for a programme because their programmes' target groups were all ordinary citizens, not a specific group of people.

Table 6-4 presents detailed results for the first 32 interviews on eight programmes.

Table 6-4 Dysfunctional Consequences of SABP in Eight Programmes in Three Case Areas

(Unit: number of coding and person)

Dysfunctional consequences	Reason	Public Housing Area						Youth Employment Area						Social Overhead Capital Area					Total		
		Coding				MOSF	Spending	Coding				MOSF	Spending	Coding			MOSF	Spending	Coding	MOSF	Spending
		P1	P2	P3	Sub			P4	P5	P6	Sub			P7	P8	Sub					
Tunnel vision		6	7	2	15	4	5	2	4	3	9	4	5	5	8	13	4	5	37	12	15
	High SABP scores and bigger budgets	1	5	-	6	2	2	-	2	1	3	2	1	4	2	6	3	3	15	7	6
	Increasing reliability and fairness	1	1	2	4	2	2	-	3	3	6	3	3	2	2	4	3	1	14	8	6
	Easiness to evaluate	2	1	-	3	1	2	4	4	-	8	1	5	5	4	9	3	5	20	5	12
Myopic management		4	4	2	10	3	5	4	3	4	11	5	6	7	10	17	4	5	38	12	16
	High SABP scores and bigger budgets	1	1	1	3	2	1	3	1	2	6	1	4	4	3	7	3	4	16	6	9
	Characteristics of public programme	-	1	-	1	1	-	-	1	-	1	-	1	-	2	2	-	1	4	1	2
	Small assessment level	-	1	-	1	1	-	-	-	-	-	-	-	1	-	1	-	1	2	1	1
	Time and cost	2	1	1	4	1	3	-	1	1	2	-	2	2	1	3	2	1	9	3	6
	Unique Korean government culture	1	3	1	5	2	3	2	4	2	8	4	3	4	3	7	3	3	20	9	9
Measure fixation		3	2	-	5	2	3	3	3	2	8	4	4	4	3	7	4	3	20	10	10
	High SABP scores and bigger budgets	1	2	-	3	2	1	3	2	1	6	2	4	4	2	6	4	2	15	8	7
	Time and cost	1	-	-	1	-	1	-	1	-	1	1	-	2	1	3	1	2	5	2	2
	Characteristics of public programmes	-	-	-	-	-	-	1	1	1	3	1	2	-	-	-	-	-	3	1	2
Denial		-	1	1	2	-	2	1	-	1	2	-	2	1	-	1	-	1	5	-	5
Suboptimisation		1	2	-	3	1	2	3	2	3	8	4	4	5	1	6	3	1	17	8	7
	High SABP scores and bigger budgets	-	3	-	3	1	1	1	2	1	4	1	3	2	1	3	2	1	10	4	5
	Unique Korean government culture	1	2	-	3	1	2	2	1	2	5	3	2	2	-	2	1	1	10	5	5
Denial		-	2	-	2	1	1	-	-	-	-	-	-	2	-	2	-	2	4	1	3
	Cabinet meeting's decision	-	1	-	1	1	-												1	1	-
Ratchet Effect		5	4	2	11	3	6	4	4	3	11	4	7	5	5	10	4	4	32	11	17
	High SABP scores and bigger budgets	2	4	-	6	3	2	2	1	2	5	3	2	2	3	5	2	3	16	8	7
	Information asymmetry	-	2	-	2	2	-	1	3	-	4	1	3	1	1	2	2	-	8	5	3
	Penalty-oriented operation	-	-	2	2	-	2	1	2	1	4	1	3	3	2	5	2	2	11	3	7
	Supervisors' concern	-	1	-	1	-	1	-	1	1	2	-	2	2	-	2	1	1	5	1	4

Dysfunctional consequences	Reason	Public Housing Area						Youth Employment Area						Social Overhead Capital Area					Total		
		Coding				MOSF	Spending	Coding				MOSF	Spending	Coding			MOSF	Spending	Coding	MOSF	Spending
		P1	P2	P3	Sub			P4	P5	P6	Sub			P7	P8	Sub					
Misrepresentation		2	4	1	7	4	3	2	1	2	5	3	2	3	2	5	4	1	17	11	6
	High SABP scores and bigger budgets	-	2	-	2	1	1	1	-	-	1	1	-	-	1	1	1	-	4	3	1
	Information asymmetry	-	1	-	1	-	1	1	1	1	3	2	1	-	-	-	-	-	4	2	2
	Penalty-oriented operation	-	-	1	1	1	-	1	1	-	2	1	1	-	1	1	-	1	4	2	2
Denial		-	-	2	2	-	2	2	1	1	4	-	4	2	2	4	-	4	10	-	10
	Continuous data request													-	2	2	-	2	2	-	2
	No concern over budgets													1	-	1	-	1	1	-	1
Cherry-picking		1	1	-	2	2	-	1	-	1	2	2	-	3	-	3	3	-	7	7	-
Denial		-	1	2	3	-	3	1	1	1	3	-	3	1	4	5	-	4	11	-	10
	Difficulty in choosing priority groups							-	1	-	1	-	1	1	4	5	-	4	6	-	5
	Using other tactics							1	-	-	1	-	1						1	-	1

Note: P1 is construction of public rental housing, P2 is subsidy for housing rental costs, P3 is improvement of old housing stock, P4 is global youth leader training, P5 is global trade specialist training, P6 is junior college students' overseas internship, P7 is road construction, and P8 is railway construction programme.

6.6 Opinions on Spending Ministries' Efforts to Prevent Dysfunctional Consequences

Interestingly, from the first interviews, I learned that spending ministries thought of the non-quantifiable or qualitative, long-term, non-measurable and whole performance of a programme as important, even though they focused on the programme's quantifiable, short-term, measurable and narrow performance. They argued that the former characteristics were more closely related to the real and genuine performance of programmes; but due to the necessity for bigger budgets, the unique Korean government culture, and especially the characteristics of public programmes – such as complexity, invisibility, indirectness and non-measurability – they had no choice but to pursue the latter. However, spending ministries have continuously tried to attain the former characteristics at the same time, in order to prevent the problems that arise from focusing only on the latter ones.

> "Most of my ministry's programmes are related in some way to the welfare of citizens, so non-quantifiable performance is considered as more important, even though we are focused on achieving quantifiable performance indicators." (Assistant Manager, MOEL, interviewed on 31st October 2011)

> "Although pursuing a long-term strategy would be more desirable for improving people's welfare, in the actual policy scene, spending ministries have no choice but to focus on the short-term achievement of a programme.." (Manager, MLTM, interviewed on 27th October 2011)

> "However, I think non-measurable performance is also important alongside measurable performance.…Non-measurable performance reflects the long-term and qualitative characteristics of a programme…The more important and more genuine performance of a programme might exist in non-measurable objectives, and so it is necessary to manage this unofficially…" (Assistant Manager, MEST, interviewed on 2nd November 2011)

Backing up the answers of the spending ministry officials, the MOSF officials also indicated that the spending ministries had tried to pay attention to non-quantifiable, long-term, non-measurable or whole performance results of a programme, even though this kind of result seemed to be difficult to achieve.

> "…spending ministries in recent times have been trying to pay more attention to long-term strategy. For example, in the youth employment programme in the MOEL, a simple 'employment rate' used to be used for the programme's performance indicator; but at now an 'employment rate over six months' is asked for as a performance indicator, if possible…" (Manager, MOSF, interviewed on 2nd November 2011)

> "Interestingly, spending ministries had a tendency to focus on their Minister's agendas, although these were non-measurable objectives in SABP…" (Manager, MOSF, interviewed on 11th November 2011)

> "…non-quantifiable performance indicators were more important in many cases, especially in the welfare field. For example, is it possible to evaluate the performance of a public child-care centre just by surveying the customers' satisfaction? Can the National Health Insurance system only be evaluated by a surplus or deficit in the financial accounts?" (Senior Manager, MOSF, interviewed on 3rd November 2011)

Across the three case study areas, a total of 13, nine, 12 and four interviewees gave positive opinions on spending ministries' efforts to prevent only the quantifiable, short-term, measurable and narrow performance of a programme respectively. However, as regards the ratchet effect, misrepresentation and cherry-picking, neither the spending ministries nor the MOSF mentioned that the spending ministries had made specific efforts to prevent these from occurring. The number of codings and opinions offered on spending ministries' efforts to prevent dysfunctional consequences can be seen in Table 6-5.

Table 6-5 Opinions on Spending Ministries' Efforts to Prevent Dysfunctional Consequences

(Unit: Number of coding and person)

Effort to prevent	Public housing			Youth employment			SOC			Total		
	Coding	MOSF	Spending	Coding	MOSF	Spending	Coding	MOSF	Spending	Coding	MOSF	Spending
Tunnel vision	4	1	3	5	-	5	5	-	4	14	1	12
Myopic management	3	-	3	2	2	1	4	-	3	9	2	7
Measure fixation	4	1	3	4	-	4	4	3	1	12	4	8
Suboptimisation	2	1	1	1	1	-	1	-	1	4	2	2
Ratchet effect	-	-	-	-	-	-	-	-	-	-	-	-
Misrepresentation	-	-	-	-	-	-	-	-	-	-	-	-
Cherry-picking	-	-	-	-	-	-	-	-	-	-	-	-

These results can provide empirical evidence on the categorisation of "unintended" and "intended" dysfunctional consequences of SABP. This will be discussed in Chapter 7. Furthermore, various dysfunctional consequences are categorised in detail, according to their extent and the reasons for their existence, in the following Tables, 6-6 and 6-7, on the basis of the findings mentioned in 6.5 and 6.6.

Table 6-6 Categorisation According to Extent of Phenomenon

		Unintended dysfunctional consequences	Intended dysfunctional consequences
Both spending ministries and the MOSF strongly agree on existence of phenomenon		Tunnel vision Myopic management	Ratchet effect
MOSF strongly agrees on existence	Spending ministries partly agree on existence	Measure fixation Suboptimisation	Misrepresentation
MOSF partly agrees on existence	Spending ministries strongly deny existence	-	Cherry-picking

Table 6-7 Categorisation According to Reason for Phenomenon

	Unintended dysfunctional consequences	Intended dysfunctional Consequences
Internal reason such as SABP design, public programmes' characteristics	Tunnel vision	-
External reason such as unique Korean gov't culture, information asymmetry, time and cost	Suboptimisation	Ratchet effect Misrepresentation
Both reasons	Myopic management Measure fixation	-

Note: High SABP scores and bigger budgets was the common reason for all types of dysfunctional consequences, including cherry-picking, but it was not possible to categorise these into internal or external reasons.

6.7 Impacts of Dysfunctional Consequences on SABP Results

The pursuit of high SABP scores and bigger budgets, the most cited reason for all types of dysfunctional consequences in the first interviews, lead to spending ministries' optimism bias when it came to their self-assessment stage. On the other hand, the MOSF basically distrusted the spending ministries' self-assessment results because it thought that the latter tended to overestimate their self-assessment scores as much as possible, and moreover sometimes indulged in tactics such as deliberately lowering self-assessment results in the case of programmes already completed and programmes the MOSF or the National Assembly were particularly interested in, so as to give high SABP scores to other programmes that particularly interested them. The MOSF's distrust of the spending ministries' results lead to the strict application of the SABP manual at the MOSF's review stage.

With respect to tunnel vision, as mentioned earlier, there was no comprehensive and objective programme evaluation of the qualitative or non-quantifiable performance of

programmes from P1 to P6, and spending ministries submitted and pursued only quantifiable performance indicators. However, the spending ministries gave a “Yes” answer to the relevant question (4-2) for all programmes in the self-assessment stage, because they considered a satisfaction survey or simple evaluation by an outside organisation to be a comprehensive and objective evaluation, as mentioned in the SABP manual. On the other hand, the MOSF gave a definitive “No” to all programmes in relation to this question, because the spending ministries, from the perspective of the MOSF, were just focusing on quantifiable performance. Consequently, the tunnel vision of spending ministries might lead both to these ministries’ optimism bias and to the MOSF’s downgrading of the ministries’ results, and thus the difference between the SABP results of the spending ministries and those of the MOSF would be widened.

As for the ratchet effect, spending ministries gave a “Yes” to Question 2-2 for all programmes, although they tended to submit rather conservative performance target levels in order to achieve them easily and to obtain bigger budgets in the following year. Moreover, spending ministries argued that they achieved their target level successfully in seven programmes, although these didn’t include P4, and so in answering Question 4-1 they gave a “Yes” for seven programmes (except P4) and a “To a large extent” for P4. However, the MOSF carried out strict reviews of the grounds on which performance levels were assessed, as well as of the target levels themselves, because it did not trust the integrity of spending ministries’ target levels, and in the end, the seven programmes (except P4) received a “No” answer to Questions 2-2 and 4-1. This might imply that the MOSF did drastically cut spending ministries’ self-assessment results in the case of generous self-assessment. Overall, the ratchet effect could lead to spending ministries’

optimism bias and the MOSF's strict review of spending ministries' self-assessment results.

The interviews also supported these results of the documentary analysis. In the first interviews on the three case study areas, most respondents answered that dysfunctional consequences lead both to the MOSF's drastic cutting of spending ministries' self-assessment results and to the spending ministries' optimism bias at the self-assessment stage (16 interviewees with 19 references, and 19 interviewees with 24 references, respectively). In addition to the government officials, most of the experts also commented on this. The results of the first interviews can be seen in Table 6-8 below.

> "… the MOSF thinks that the spending ministries tend, for strategic reasons, to put forward their self-assessment score at rather a high level; so the MOSF has a tendency to apply the criteria very strictly to the ministries' self-assessment results." (Senior Manager, MOSF, interviewed on 1st November 2011)

> "…but spending ministries are very sensitive about their SABP results, and they only have an interest in the SABP results, not in improving their programmes' performance or budget efficiency." (Manager, MKE, interviewed on 28th October 2011)

Table 6-8 Impacts of Dysfunctional Consequences on the MOSF and Spending Ministries

(Unit: number of coding and person)

Main categories		Public housing			Youth employment			SOC			Total		
		Coding	MOSF	Spending	Coding	MOSF	Spending	Coding	MOSF	Spending	Coding	MOSF	Spending
MOSF	Drastic cutting of spending ministries' results	5	3	1	8	2	4	6	3	3	19	8	8
Spending	Optimism bias	4	2	2	12	3	5	8	3	4	24	8	11

In the supplementary interviews I again asked directly about the impact of dysfunctional consequences on SABP results, in order to check the results of the first interviews. In this round, twenty nine interviewees (11 MOSF, 13 spending ministry, five experts) out of 31 responded that spending ministries had a tendency to focus on performance in a quantifiable, short-term, measurable and narrower sense, trying to set performance target levels as low as possible, sometimes hiding and manipulating disadvantageous performance data, and highlighting the impact on priority groups in order to obtain a high SABP score and a budget increase. Moreover, they indicated that this tendency of the spending ministries towards an optimism bias when assessing their programmes could deepen the MOSF's distrust of the ministries' self-assessment results and could contribute to the MOSF's drastic cutting of these results. Two interviewees gave me no answers on this subject, but did not deny these phenomena.

In addition, many MOSF officials in the supplementary interviews, interestingly and practically, indicated that the starting point for these unexpected side effects is the differing perspectives on the appropriateness of performance indicators (Q2-1) and performance target levels (Q2-2) in the SABP checklist; because if a programme receives a "No" for Question 2-1, then it cannot obtain a score over 65, due to the logical linking of the questions; and thus the probability of a poor grade is increased. Thus, the spending ministries always made efforts to answer "Yes" to Questions 2-1 and 2-2, while the MOSF reviewed the accuracy of the answers to these questions particularly strictly.

6.8 Policy Alternatives

Interviewees suggested various kinds of policy implications, and I divided these into urgent, short-term, and medium- or long-term policies for preventing the dysfunctional consequences of SABP. This section analyses the results of both the first and the supplementary interviews. After analysing the first interview data, I found that some policy alternatives needed to be considered again, and with great care, because they were very important, or because there might be differing opinions between spending ministries and the MOSF when implementing them as policy alternatives. Of 14 policy alternatives suggested in the first interviews, six were re-examined through the supplementary interviews. In the supplementary interviews, in particular, interviewees were required to indicate the importance of these policy alternatives by putting them into one of the following five categories – Very important (5.0 points), Important (4.0), Modest (3.0), Not important (2.0) and Absolutely not important (1.0) – in order to compare the perspectives on policy alternatives between spending ministries and the MOSF and to put the policy alternatives into an order of priority. Policies discussed in the supplementary interviews were as follows: (1) refining the SABP checklist, especially by adding questions related to qualitative, long-term, non-measurable and whole performance; (2) changing the penalty-oriented operation of SABP; (3) introducing a relative evaluation system; (4) elevating the level assessed by SABP from "sub-programme" to "programme" level in the programme budgeting structure; (5) giving more discretion to spending ministries; and (6) abolishing the stage of self-assessment by spending ministries. The remainder of this section will explore the

various policy alternatives one by one, based on the results of both interviews. In addition, the feasibility of these policy alternatives will be discussed detail in Chapter 7.

6.8.1 Options for Urgent Policy Change

I decided that refinement of the SABP checklist, changing penalty-oriented operation, and attracting the interest of senior officials were urgent priorities, considering the number of references made to them by interviewees and the influence these items appeared to exert.

Refinement of the SABP Checklist

The MOSF has revised the SABP checklist every year e.g., changing the number of questions (reducing unnecessary questions and merging related questions), adjusting the weight of each question, and adopting a partial scoring system in many of the questions in order to reflect accurately spending ministries' efforts to improve the performance of a programme (MOSF, 2007, 2008b, 2009a, 2009b, 2010a, 2010g). Despite the MOSF's annual refinements, most respondents (25 interviewees, with 51 references) and experts mentioned the necessity for increasing the weight of questions or adding questions related to the evaluation of spending ministries' efforts to achieve non-quantifiable, long-term, non-measurable and whole performance for a programme in the first interviews. They also thought that the MOSF should require spending ministries to submit evidence of these efforts as an appendix, at least, to the official SABP report. Similarly, Kelman and Friedman (2009) suggest adding measures to prevent effort

substitution and adapting measures to reflect organisational learning about gaming will limit measure fixation. Thiel and Leeuw (2002) suggest that performance paradox can be prevented by developing new performance indicators which refer to different aspects of a programme.

Against this background, in order to examine the feasibility of this policy alternative, I asked some specific questions in the supplementary interviews: whether the interviewees agreed that these kinds of dysfunctional consequences existed; whether it was necessary to add questions related to the qualitative, long-term, non-measurable and whole performance of a programme; whether it was necessary to submit evidence to back up claims about performance; and how important interviewees perceived the policy alternative to be. First of all, all respondents from both spending ministries and the MOSF agreed that these phenomena did actually exist in the SABP system. By contrast, two experts focused more on the lack of expertise of spending ministry officials rather than potential problems with the SABP system. As for the necessity of adding questions and requiring evidence, 23 interviewees (10 MOSF, 11 spending ministry, two expert) indicated the necessity for adding questions, especially two questions related to the long-term and whole performance of a programme, as well as for spending ministries to submit relevant evidence. On the other hand, seven interviewees (two MOSF, two spending ministry, three expert) disagreed with these specific alternatives, because they thought that the current SABP checklist could avoid these side effects by appropriate use of Question 4-2 (comprehensive and objective programme evaluation), and they were worried that these policies would significantly increase the burden of work for spending ministries. However, overall, most

respondents agreed with the severity of these dysfunctional consequences and wanted firm action taken to minimise them. These results supported the first interview results well. Moreover, respondents in the supplementary interviews regarded the policy alternative as important, because the average importance of these changes was 4.0 out of 5.0 points (MOSF 4.4, spending ministries 3.8), though the MOSF considered them more important than did the spending ministries.

Changing Penalty-oriented Operation

With respect to the penalty orientation of the SABP system, the spending ministries argued that the MOSF had followed the penalty rule assiduously, but had not kept to the incentive rule. The strong relationship between SABP results and budget allocation, especially between poor grades and budget cutting, had contributed to the spending ministries having more interest in SABP at the initial stage. However, so far the spending ministries' optimism bias had not decreased as much as the MOSF had expected it would, and the main reason for this might be the penalty-oriented operation of the SABP system. Although it might be difficult for the MOSF to abandon its penalty-oriented policy basis, it might be time to reconsider the way in which SABP operated (12 interviewees with 17 references in the first interviews).

On the basis of the results of the first interviews, in the supplementary interviews I asked the following questions: whether the interviewees agreed with the existence of this phenomenon; whether it was important to give incentives to effective programmes; whether such a policy change might genuinely improve the performance of a

programme; and what interviewees' perception of the importance of this policy alternative was. In the supplementary interviews, 21 interviewees (nine MOSF, 12 spending ministry) agreed with the necessity for changing the penalty-oriented operation of SABP. Many respondents indicated that SABP results tended to be mechanically used as a tool for reducing a programme's budget, not for providing useful performance information for the programme's management, so spending ministries just tried to avoid a poor grade without making any effort to improve the programme's genuine performance. In addition, some interviewees mentioned that penalty-oriented operation of SABP encouraged spending ministries to just focus on trivial but high-scoring programmes instead of risky but important programmes. Moreover, the time lag, which meant SABP results for a past year's programme being applied to the programme's future budget, was also indicated as a reason why it was necessary to change the system's penalty-orientated operation. In general, most respondents disagreed with the mechanical application of SABP results to budget allocation, although at the same time they agreed with the necessity for feedback in one form or another. The interesting thing was that the MOSF officials considered this problem more serious than did the spending ministry officials. The average perceived importance score given by MOSF officials was 4.4 points, whereas that of spending ministries was just 3.6 points (making the overall average 3.9). This means that the designer and controller of the SABP system might consider the way it operated more cautiously than spending ministry officials. An even more interesting point was that two spending ministry officials and four out of five experts disagreed with changing the penalty-oriented operation. Spending ministry officials emphasised that the SABP results should be strongly linked with budget allocations in order to provide an

opportunity for spending ministries to reconsider their programmes seriously. Four experts also commented that it was too early to change penalty-oriented operation, because spending ministries still did not understand that SABP's main purpose was to increase the efficiency of budget allocation through linking performance results and budgeting. However, the overall opinion of interviewees was that penalty-oriented operation would not help to reduce various dysfunctional consequences of SABP, and it was time to consider changing it. As for increasing the incentives for effective programmes, most respondents replied that both incentives and penalties might be necessary to really improve the performance of programmes. Weighing incentives and penalties might enable spending ministries to reconsider their performance indicators and performance target levels more seriously. These changes would mean that the MOSF would not rely on SABP just for control purposes (Smith, 1995). Moreover, all the MOSF officials responded that this policy alternative could genuinely increase the performance of a programme to a significant extent (ten interviewees) or to a small extent (two interviewees), and these answers indicate that the MOSF officials perceived the importance of this policy. Eight spending ministry officials agreed that changing the penalty orientation of SABP could increase the level of performance of programmes, while four of them disagreed with this idea, because performance could be improved by many other methods. Four experts responded positively, while one expert disagreed, saying that this policy would not enhance the real performance of programmes, because spending ministries would try to find appropriate tactics whatever the policy implemented. Overall, the supplementary interview results to a large degree supported the first interviews results, and changing the penalty-oriented operation of SABP might

be essential for minimising and preventing various side effects and enhancing the effectiveness of the SABP system.

Increasing Senior Officials' Interest in SABP

More input into SABP by senior officials was another urgent policy alternative proposed by many interviewees (10 interviewees with 12 references in the first interviews). Seniors officials included the ministers and vice ministers of each ministry. Although the size of the Performance Management Bureau in the MOSF and the number of MOSF officials who were responsible for reviewing spending ministries' self-assessment results had decreased between 2005 and 2010, the National Assembly and the BAI had become more interested in SABP. The NABO had started to evaluate both spending ministries' performance planning reports and their actual performance reports every year. The BAI had strengthened its evaluation of both spending ministries' planning and their performance reports. In 2011 the BAI completed an audit of all spending ministries' performance reports after a pilot auditing project initiated in 2009. In order to refine and improve the SABP system in a short period, I came to the conclusion that senior officials' input was vital. In this area, a reform undertaken by the Obama administration gave some useful pointers. The Obama administration has started to monitor several HPPGs, at quarterly intervals, through the Vice Minister's meeting, and then to link performance data with a dashboard in order to monitor them consistently. By enhancing senior officials' interests in programmes' performance, and monitoring performance consistently, more relevant performance information could be made available and applied to programme management. The advantages of the US

GPRAMA and the possibility of applying the system to SABP in Korea need to be considered in a future study.

6.8.2 Options for Short-Term Policy Change

A More Active Role for Related Organisations

Among short-term policy alternatives, a more active role for the MOSF, the National Assembly and the BAI was the most frequently indicated by the first interviewees (11 interviewees with 16 references). Although it was the MOSF, in 2010, that started to give penalties to spending ministries when the gap between a spending ministry's SABP score and the MOSF's review score was 20 points or more, and to create a systematic performance history for each programme (MOSF, 2010a, 2010g, 2011b), the experts emphasised the division of work between the related organisations. They commented that the MOSF should focus on increasing the appropriateness of performance indicators, such as outcome-related and representative indicators, while the BAI should focus on verifying the suitability of performance data by checking whether spending ministries used official and consistent statistics drawn from reliable statistical databases when submitting their performance reports to the MOSF and the National Assembly. In this context, the MOSF's distribution of the manual for developing appropriate performance indicators and target levels in various budgetary programmes made an important contribution (MOSF and KIPF, 2012). The BAI's checking could reduce misrepresentation by spending ministries very sharply, and the MOSF's manual could be used to evaluate spending ministries' performance target levels precisely, and this

would lead to a decrease in the ratchet effect. In addition, the National Assembly's strong monitoring of the performance reports of central government departments would also minimise various dysfunctional consequences of the SABP system.

An Open Evaluation System

In the first interviews, many respondents commented on the necessity for what is called an open evaluation system ranging from previous consensuses on performance indicators and target levels among spending ministries, the MOSF, the experts, and even the National assembly, to opening up performance assessment by related beneficiaries, experts, government-affiliated organisations, spending ministries and the MOSF. The NSTC had already announced that an open assessment system for research and development (R&D) programmes[28] would be introduced from 2012, to increase the transparency and reliability of assessment (NSTC, 2011). In particular, consensuses on performance indicators and performance target levels among related organisations would decrease the dysfunctional consequences at the performance planning stage, such as tunnel vision, myopic management, suboptimisation and the ratchet effect.

A Relative Evaluation System

The MOSF also recognised spending ministries' persistent optimism bias, so it announced it was introducing a "relative evaluation system" into the spending ministries'

[28] On the commencement of the NSTC as a new control body for R&D policies in Korea, on March 2011, responsibility for performance assessment on R&D programmes was transferred from the MOSF to the new body.

self-assessment stage from 2011, on a trial basis. This would make it compulsory for the proportion of programmes graded “effective” to be less than 20% and those graded “poor” to be more than 10% (MOSF, 2011b). Although only three interviewees indicated the necessity for this policy change in the first interviews, I wanted to know whether the system would prove feasible or not, so the supplementary interviews asked for opinions on it. In the supplementary interviews, ten interviewees (six MOSF, three spending ministry, one expert) agreed with the policy change and another ten respondents (four MOSF, five spending ministry, one expert) partly agreed with it, while nine interviewees (two MOSF, four spending, three expert) disagreed with it. That is, the number of respondents holding each of the three different opinions was nearly same. The first group recognised that compulsory rationing seemed to be inevitable, and there were no alternatives for reducing spending ministries’ optimism bias sharply and rapidly. The second group agreed with the contribution made by the system; but they also mentioned the potential risk it posed, indicating the possibility of spending ministries’ strategic behaviours in the process of satisfying the proportions of effective and poor grade programmes. For example, it was likely that spending ministries would deliberately self-assess programmes the MOSF or the National Assembly were interested in as poor, although they knew these programmes were implemented effectively, because they knew very well that these programmes were too important for the MOSF or the National Assembly to reduce their budgets. In addition, the assignment of an effective or poor grade can be determined by the working of interests within each spending ministry, not by objective criteria. The third group, including three experts, focused only on the negative effects of the relative evaluation system. Three experts stated the necessity of preventing unexpected side effects and tactical behaviours by

spending ministries, in order to operate the policy change consistently. With respect to the appropriateness of the proportions – 20% and 10% – most respondents could not answer whether this number was suitable or not, because the MOSF had decided on those proportions rather arbitrarily. Moreover, the average level of perceived importance for this policy was not so high compared with those for the previous two policy alternatives. It was just 3.1 points (3.4 MOSF, 2.9 spending ministry), a modest level. Overall, considering the different opinions voiced by groups of similar supplementary interviewees and the trial basis implementation of the policy change, it will take more time to evaluate the feasibility of this policy.

Others

Besides these policy alternatives, some interviewees indicated that the SABP results should be reported publicly every year, in order to enable spending ministries to consider ordinary citizens' satisfaction and to pursue genuine performance of programmes (two interviewees with three references in the first interviews). In addition, reducing unnecessary workload related to SABP was also suggested by two interviewees with two references in the first interviews.

6.8.3 Options for Medium- or Long-Term Policy Change

Elevating the Assessment Level from "Sub-programme" to "Programme"[29] and Focusing on Core Programmes

With respect to medium- or long-term policy alternatives across the three case study areas, elevating the assessment level and focusing on the core programmes of spending ministries were the most mentioned policy alternatives (12 interviewees with 18 references in the first interviews). Many interviewees indicated that the present assessment level of SABP was too small to consider the long-term and overall objectives of an organisation. This policy alternative reflected the Obama administration's initiative in the US, focusing on just a few HPPGs in each spending ministry (KIPF, 2011).

Against this background, the supplementary interviews asked for opinions on elevating the assessment programme from the present "sub-programme" level to "programme" level in the programme budgeting structure mentioned in Section 3.2.3. Fourteen interviewees (four MOSF, eight spending ministry, two expert) agreed with this policy alternative, while 15 (eight MOSF, six spending ministry, one expert) disagreed with it. In addition, two experts offered rather different opinions. The main reason for agreeing with the policy was that many sub-programmes of spending ministries were closely

[29] In this policy alternative, I distinguish between the terms "sub-programme" and "programme". The present SABP system was implemented at the "sub-programme" level in the programme budgeting structure. Thus, strictly speaking, the term "programme" as it has been used in the study until now means "sub-programme". However, in other parts of the study I do not use the term "sub-programme", in order not to cause confusion, because "sub-programme" is a concept belonging to the programme budgeting structure, and is not used generally.

related to each other within the same programme level, and each sub-programme was implemented in systematic cooperation with other sub-programmes. Therefore, some outcomes of one sub-programme were produced jointly with those of other sub-programmes, and so it was necessary to consider their performance more comprehensively. In addition, a sub-programme's performance indicator was generally too narrow and small to be understood by ordinary citizens. Consequently, sub-programme level's performance information is difficult to consider in terms of long-term or whole performance. On the other hand, other interviewees argued that performance information at the programme level cannot give useful feedback on budget allocation, because the programme consists of several different sub-programmes. Moreover, a programme can include several non-associated performance indicators of several sub-programmes, and so it is difficult to reflect each sub-programme's characteristics well and to decide on its real performance at a programme level. They explained that this is why the MOSF prefers to assess budgetary programmes at sub-programme, not programme level. Two experts commented that setting up an exact relationship between sub-programme and programme, before discussing the appropriate level of SABP and assessing programme level, is the role of the in-depth evaluation system (IES), which is one of the Korean budgeting systems, rather than of SABP. The importance level of this policy alternative was just 3.4 points (3.3 MOSF, 3.4 spending ministry).

Increasing Expertise

Increasing the expertise of both spending ministries and the MOSF would also be

important in order to decrease the information asymmetry between spending ministries and the MOSF. By reducing this information asymmetry, the MOSF would be able to ask spending ministries to submit exact performance data and to evaluate the appropriateness of their performance target levels, and spending ministries would be able to understand the genuine purposes of performance management, such as increasing the responsibility and accountability, as well as the efficiency and effectiveness of budgetary programmes. Thus, increasing the expertise of both spending ministries and the MOSF would reduce the possibility of dysfunctional consequences of SABP, particularly the ratchet effect and misrepresentation.

Building Trust and Giving More Discretion to Spending Ministries

In the first interviews some interviewees (three MOSF, six spending ministry) indicated that changing the SABP system from a MOSF-driven to a spending ministries-driven system based on mutual trust might be necessary in order for SABP to be implemented in a substantially self-assessing way. It has not been easy for spending ministries to focus voluntarily on the qualitative, long-term, non-measurable and whole performance of a programme in a situation where the MOSF did not trust spending ministries' self-assessment results. Greater trust might lead spending ministries and the MOSF to reduce their differences over SABP results and the possibility of dysfunctional consequence created by spending ministries.

In the supplementary interviews, 15 interviewees (three MOSF, 12 spending ministry) agreed, seven (three MOSF, one spending ministry, three expert) agreed conditionally,

and nine (six MOSF, one spending, two expert) disagreed with this policy alternative. Interestingly, most spending ministry officials unconditionally agreed with this change, while most MOSF officials disagreed or conditionally agreed with it. Issues raised in the supplementary interviews were more profound and detailed. The first main reason for agreement was that spending ministries have more information about their programmes and each programme's relative importance within the ministry, so they can set the direction for how to evaluate even the invisible performance of a programme, and how to operate SABP in the long run. The second main reason was that SABP can ultimately be successful only if it is implemented on the basis of mutual trust between spending ministries and the MOSF. On the other hand, several reasons for disagreement or conditional agreement were also mentioned. Firstly, if the MOSF gives more discretion to spending ministries, then the optimism bias of spending ministries will become larger, and so this change needs to proceed steadily, taking into consideration how great spending ministries' optimism bias will be. Secondly, although the MOSF gives some discretion in terms of developing performance indicators and setting performance target levels, related organisations, such as the MOSF, the National Assembly and the BAI, have continuously indicated the lack of quality and reliability of this performance information. Thus, instead of unconditionally increasing the discretion allowed to spending ministries, strengthening cooperation or consultancy with the MOSF might be needed when it comes to developing performance indicators and setting performance target levels. Thirdly, various dysfunctional consequences of SABP come not from the lack of discretion allowed to spending ministries but from differing perspectives on the appropriateness of performance indicators and performance target levels, so increasing discretion cannot be a suitable policy alternative for minimising

these dysfunctional consequences. The interesting point, regardless of the number of interviewees who agreed or disagreed, was that the perceived importance of the MOSF was greatly lower than that of spending ministries (2.8 MOSF, 3.6 spending ministry). This coincided with the first interview results, and it might mean that the MOSF still thinks that spending ministries lack expertise, objectivity, reliability and responsibility in producing their self-assessment results, and moreover there is still distrust between spending ministries and the MOSF. Also, many respondents answered that the MOSF's ability to compare performance between programmes would not necessarily be decreased if spending ministries' discretion were increased through the above policies. They also indicated the necessity for improving the expertise of the MOSF, regardless of the extent of spending ministries' discretion, in order to increase the ability to compare the differences between programmes.

Abolishing the Spending Ministries' Self-assessment Stage

Although this policy alternative was indicated by just two interviewees in the first interviews, I thought that it needed to be investigated more thoroughly in the supplementary interviews, because the change could have a great influence on implementing the SABP system in the future. In the supplementary interviews, the number of agreements and disagreements was nearly same for spending ministry officials and MOSF officials. Fourteen interviewees (five MOSF five, seven spending ministry, two expert) agreed with the policy and 15 interviewees (six MOSF, seven spending, three expert) disagreed with it. Nearly half of all interviewees agreed, not only because spending ministries' optimism bias was so great that the self-assessment

stage might be meaningless but also because the MOSF had not accepted spending ministries' self-assessment results as they were, and this could reduce unnecessary administrative costs as well as increase administrative efficiency. However, the other half of the interviewees disagreed, because spending ministries' position on many programmes could be delivered to some extent through self-assessment results, and spending ministries had an opportunity to increase their responsibility for their programmes as well as to enhance their capacity to manage programmes' performance. In addition, some interviewees were of the opinion that the fact that spending ministries' self-assessment results were not accepted by the MOSF could not be a reason for abolishing the self-assessment stage, and that finding out the reason why spending ministries' self-assessment results were cut drastically at the MOSF's review stage was more important. Moreover, the MOSF considered this change as more important than the spending ministries did. The MOSF's perceived importance level was 3.8 points, while that of the spending ministries was 3.1. This might imply that the MOSF was playing a more active role in designing and revising the SABP system, while the spending ministries were adopting a rather passive position on this fundamental change. On the other hand, eight interviewees replied that the MOSF's confidence in its ability to evaluate performance without spending ministries' self-assessment results would fall, whilst 16 respondents answered that this change would not necessarily decrease the MOSF's confidence. Thus, the MOSF's confidence in its evaluations seemed not to be a critical factor when it came to deciding whether or not to abolish the spending ministries' self-assessment stage. Rather, the more important thing seemed to be to consider seriously the purpose of SABP.

Others

The other two policy alternatives were related to the restructuring of governments or the reengineering of related governments' roles. Interestingly, the necessity for integrating the MOSF's Performance Management Bureau into the Ministry's Budget Office, on the lines of the US OMB model, was suggested by three MOSF officials in the first interviews as a means of increasing the expertise of the MOSF. As for similar performance management systems, the PMO assessed spending ministries' major programmes, regardless of whether they were budgetary or non-budgetary programmes, based on the Government Performance Assessment Act passed in 2006. However, the MOSF assessed spending ministries' budgetary programmes on the basis of the National Finance Act implemented in 2007. The PMO and the MOSF agreed to match the assessment level of the PMO with that of the MOSF in order to avoid conflicts between two Acts. On the other hand, the responsibility for assessing R&D programmes was transferred from the MOSF to the NSTC with the setting up of the NSTC in 2011. Although the purposes, assessment level and application of results among these performance management systems were rather different, spending ministries complained about the excessive and unnecessary workloads related to performance management systems. However, these policy alternatives should be decided cautiously after analysing the advantages and disadvantages of them in the medium and long terms.

The number of references and respondents on policy implications for improving SABP in the first interviews can be seen in Table 6-9.

Table 6-9 Number of Codings and Respondents for Policy Implications for Improving SABP

(Unit: number of coding and person)

Policy	Public housing			Youth employment			SOC			Total		
	Coding	MOSF	Spending	Coding	MOSF	Spending	Coding	MOSF	Spending	Coding	MOSF	Spending
Options for Urgent Policy Change												
Refine SABP checklist	15	3	5	15	3	5	21	4	5	51	10	15
Changing penalty-oriented operation	1	-	1	14	4	5	2	1	1	17	5	7
Senior officials' interest	5	2	1	4	2	2	3	2	1	12	6	4
Options for Short-term Policy Change												
More active role (monitoring, penalty, education, etc.)	2	1	1	9	3	2	5	4	-	16	8	3
Open evaluation system	7	2	2	2	2	-	4	1	3	13	5	5
Relative evaluation system	2	1	1	1	1	-	-	-	-	3	2	1
Reporting SABP results publicly	2	-	1	-	-	-	1	-	1	3	-	2
Reducing unnecessary workload	-	-	-	1	-	1	1	-	1	2	-	2
Options for Medium- or Long-term Policy Change												
Elevating assessment level and focusing on main programmes	5	2	1	6	3	2	7	2	2	18	7	5
Increasing expertise	4	1	2	7	2	4	4	2	1	15	5	7
Building trust and giving more discretion	2	-	1	8	2	4	2	1	1	12	3	6
Abolishing self-assessment stage	-	-	-	2	1	1	-	-	-	2	1	1
Integrating the Budget Office and Performance Mgt Bureau	1	1	-	2	2	-	-	-	-	3	3	-
Integrating performance management system	1	1	-	1	-	1	1	-	1	3	1	2
Others	2	1	1	2	-	2	1	-	1	5	1	4

A summary of the differences in responses to the six policy alternatives between the first and the supplementary interviews can be seen in Table 6-10. The feasibility of these six policy alternatives will be discussed further in Chapter 7.

Table 6-10 Summary of Differences Between First and Supplementary Interviews

(Unit: number of interviewees, points)

Policy	First interviews			Supplement interviews				Perceived importance
	Coding	MOSF	Spending	Agree		Disagree		
				MOSF	Spending	M	S	
Refining SABP checklist	51	10	15	12	14	2	2	4.0
Changing penalty-oriented operation	17	5	7	9	12	3	2	3.9
Relative evaluation system	3	2	1	10 (4 is partly)	8 (5 is partly)	2	4	3.1
Elevating assessment level	18	7	5	4	8	8	6	3.4
More discretion to spending ministries	12	3	6	6 (3 is conditional)	13 (1 is conditional)	6	1	3.2
Abolishing self-assessment stage	2	1	1	5	7	6	7	3.5

6.9 Conclusion

This chapter has shown how the research found that the SABP system had been subject to various dysfunctional consequences, some of which were unintended but others of which were intentional on the part of the spending ministries, agents in the principal-agent relationships. Although the perspective on the existence of each dysfunctional consequence might not be the same in the spending ministries as it was in the MOSF, it was possible to say that various dysfunctional consequences, both large and small, of SAPB had led to a strengthening of spending ministries' optimism bias in their self-assessment. This could be the reason why significant differences between the SABP results of spending ministries and those of the MOSF persisted. However, I would not contest the usefulness of SABP. This is because, as previous researchers have found, SABP has increased both the efficiency of budgetary programmes and the accountability of spending ministries. The important thing is not the existence of

various dysfunctional consequences but the need to solve them. In implementing policy, pursuing the second best solution, with every effort made to minimise attendant problems, may be more desirable than seeking an optimum solution with no faults. To this end, the Korean government, especially the MOSF, has implemented many policy modifications indicated by the National Assembly and the BAI. All of these have been related to the urgent, short-, medium- or long-term policy aims mentioned in this chapter. From the several policy changes suggested by interviewees, I drew the conclusion that refining the SABP checklist, changing the penalty-oriented operation of the SABP system, and involving senior government officials such as the Minister and Vice Minister were the policy changes most likely to improve the SABP system rapidly. The next chapter will discuss several issues raised by the findings of both the quantitative (Chapter 5) and the qualitative (Chapter 6) analysis, including the feasibility of policy alternatives suggested by the first and the supplementary interviews.

CHAPTER 7

DISCUSSIONS

7.1 Introduction

This chapter interprets and discusses in detail several issues raised by the findings from the quantitative and qualitative research approaches, and distinguishes this study's unique contribution from the contributions and limitations of previous studies. Some issues involve findings from both the quantitative and qualitative analysis. The chapter firstly interprets the factors which affect the DR and ROR of spending ministries, and then explains the reasons for these phenomena in the context of the Korean performance budgeting system. Secondly, with respect to the qualitative analysis, it discusses five main issues, including the appropriateness of categorising unintended and intended dysfunctional consequences by spending ministries; the impacts of these; and feasible policy alternatives. On the other hand, as part of the discussion of the data, findings, and interpretations of the quantitative and qualitative analysis, the chapter evaluates the validity and reliability of these. Finally, the chapter will discuss four contributions (theoretical, empirical, methodological and practical) to knowledge in three main areas.

7.2 Implications of the Quantitative Analysis

7.2.1 Factors which Affect the DR and ROR

Relationships Between Two Independent Variables and the DR and ROR

According to the quantitative analysis, "the MOSF's review score" had a negative relationship with the DR and ROR of spending ministries, while "specific programme's self-assessment score minus the spending ministry's average score in that year (specific score minus average score)" had a positive relationship with the DR and ROR in a statistically significant way, at the level of 0.01, from 2005 to 2010. These findings have the following implications. Firstly, they imply that over these six years there was indeed not only a high and consistent optimism bias by spending ministries in the self-assessment stage of SABP, but also drastic cutting by the MOSF of spending ministries' generous self-assessment results. When we consider that the main purpose of SABP is to increase budget efficiency and effectiveness, as well as to improve spending ministries' performance management capability, these findings might suggest that SABP has not accomplished its aims so successfully and there have been various unintended dysfunctional consequences by the MOSF. In addition, they suggest the necessity for feasible policy alternatives to minimise dysfunctional consequences. This means that these findings from the quantitative analysis shed meaningful light on exploring dysfunctional consequences of SABP and searching for useful policy alternatives that will refine and redesign the SABP system by employing qualitative analysis.

Secondly, although the two independent variables had statistically significant relationships with the DR and ROR for three sections (performance planning, results and whole section) in the three time periods, when comparing the adjusted R squares, we can see that the adjusted R square in the performance planning section in the three time periods was generally higher than that in the other two sections. For example, in the case of the DR, the adjusted R square in the performance planning section in the three time periods was 0.363, 0.389 and 0.346 respectively, and in the case of the ROR it was 0.166, 0.312 and 0.213 respectively. Except for 0.166 in the case of the ROR in the first cycle, the other adjusted R squares were higher than those for the results and the whole section. This suggests that in each time period spending ministries' optimism bias and the MOSF's drastic cutting might be better explained in relation to the performance planning section of SABP rather than to other sections. In other words, it means that spending ministries did their best to get a "Yes" answer in the performance planning section and at the same time the MOSF also considered spending ministries' self-assessment results in the performance planning section more carefully. These results can be explained not only by the fact that SABP Questions, especially Questions 2-1, 2-2 and 4-1, are logically linked to each other, but also by the fact that spending ministries have more discretion and tend to self-assess more generously in performance planning, for example when they develop performance indicators and set performance target levels. Thus, these results mean that preventing dysfunctional consequences in the performance planning section might be a priority in order to reduce the DR and ROR. Also, they may justify the effort of examining various dysfunctional consequences in the separate categories of performance planning, performance measurement and

performance reporting. Moreover, policy makers should pay greater attention to developing appropriate performance indicators and target levels before implementing programmes, although this might reduce spending ministries' discretion to some extent. This implication effectively supports the policy alternative, which will be discussed later, of strengthening cooperation between spending ministries and the MOSF when developing performance indicators and target levels, while reducing spending ministries' discretion at this stage.

Thirdly, the adjusted R square for the three sections generally decreased from the first SABP cycle to the second SABP cycle. For example, in the whole section the DR fell from 0.362 in the first SABP cycle to 0.228 in the second SABP cycle, and the ROR was reduced from 0.311 to 0.176. This means that if other conditions are equal, the extent of the explanatory power of spending ministries' optimism bias and the MOSF's drastic cutting has been decreased. In other words, it is possible to estimate that spending ministries' optimism bias and the MOSF's drastic cutting were weakened as the SABP process was repeated, due to learning by doing. This interpretation is supported by the negative relationship between the number of SABP implementations and the ROR in the following.

Relationships Between Control Variables and the DR and ROR

Eight control variables had relationships of different kinds with the DR and ROR. Among the relationships between the eight control variables and the DR, only programmes directly managed by the government and large budget programmes had

relationships with the DR in expected ways. These findings mean that spending ministries can have greater influence on directly managed programmes than on indirectly managed ones, so they have paid more attention to obtaining bigger budgets for directly managed programmes, and this tendency is likely to increase spending ministries' overestimation of these programmes in the self-assessment stage. Similarly, spending ministries tend to have more concerns over large budget size programmes, because these programmes generally have a closer relationship with important policies within the ministry. At the same time, the MOSF tends to use large budget size programmes more efficiently when changing budget structures. These interpretations can also be applied to the relationship between the two control variables and the ROR.

However, small budget size programmes and economy-related programmes had relationships with the DR that were the opposite of what might be expected. It seems that spending ministries also paid attention to small programmes, which is different from our initial expectation, because it might generally be easier to obtain suitable budgets for small programmes than for large budget programmes, despite spending ministries putting less effort into these. In addition, since small programmes often include the operating costs of related organisations, such as interested institutions, spending ministries are as concerned about these programmes as about large budget size programmes. Consequently, spending ministries might have an incentive to give as generous SABP scores as possible, despite these being small programmes. This might lead to strengthening spending ministries' optimism bias and the positive relationships between small programmes and the DR. This estimation can be also applied to the relationships between small programmes and the ROR. Thus, we can say that spending

ministries have focused on obtaining bigger budgets regardless of the budget size of a programme. Economy-related programmes showed negative relationships with the DR, which was different from the initial expectation, while the expected relationships between them and the ROR were proved correct. Economy-related programmes' budgets were initially expected to fall, due to increasing welfare expenditure following the Asian financial crisis in the latter 1990s; so they might be more affected by SABP results, and this would lead to a big difference between spending ministries and the MOSF. Thus, the unexpected relationship between this variable and the DR might come from ambiguity in the definition of economy-related ministries. In reality, economy-related ministries often implement welfare-related programmes as well, so simply regarding economy-related ministries' programmes as economy-related programmes might not be accurate. The other possible reason is the opposite effects produced by "Yes but No" type and "No but Yes" type disagreements, because unlike measuring the impact of economy-related programmes on the ROR, the impact of these on the DR did not consider the opposite meaning of the two types of disagreements. If economy-related programmes are operationalised more specifically in future research, then the results might be different.

The relationships between four control variables – budget changes for the current year, ministry programmes, welfare-related programmes, and the number of times SABP has been implemented – and the DR was looked at and it was not clear whether or not these variables had the relationships with the DR that would have been expected. However, interestingly, these four control variables had the relationships that would have been expected with the ROR. The possible reason for the difference might be the opposite

effects produced by the two types of disagreement. Although "No but Yes" type disagreements account for only about 9% of all disagreements, this might affect the relationships between these control variables and the DR and ROR. This means that it is necessary to consider the impact of the opposite effects of the two types of disagreements on the DR and ROR in order to examine the relationships between variables more accurately. Interpreting the findings based on the relationships between the four control variables and the ROR has several implications. Firstly, when budget increases for programmes for the current year are low, spending ministries try to obtain bigger budgets for the following year, whilst the MOSF evaluates these programmes' validity as low. These conflicting behaviours may lead to a widening of the difference between spending ministries and the MOSF. Secondly, in the case of ministries' programmes, which are more complex than agencies' programmes, the difference between spending ministries and the MOSF may generally be wider. Thirdly, by comparison with economy-related programmes, welfare-related programmes have been less affected by budgetary reforms, building a stronger social safety net in preparation for a low-birth rate and aging society. Finally, as SABP has been implemented several times, as time goes by, the difference between spending ministries and the MOSF may decrease due to the learning effect. The existence of a positive learning effect may give a good indication that the aims of the SABP system can be achieved.

Taken together, the findings about the relationships between control variables and the DR and ROR tell us that, regardless of its budget size, if a programme is directly managed by the government, is complex (a ministry's programme rather than an agency's), has had a low budget percentage change for the current year, and is

economy-related, then the difference between the spending ministry's score and that of the MOSF may be higher. On the other hand, this difference may be reduced when the SABP is implemented repeatedly, due to the learning effect. Therefore, the MOSF needs to consider carefully the greater possibility of an optimism bias by spending ministries in the case of directly managed programmes, ministries' programmes, programmes that have low budgets compared to the previous year, and economy-related programmes.

7.2.2 Dynamic Pattern of the RORs of Spending Ministries

From the analysis of the programmes which had been assessed by SABP two or three times, the study found that the MOSF had paid greater attention to reducing a high ROR by raising its review score, whilst spending ministries had focused more on increasing a low ROR by increasing their self-assessment score in the following SABP round. This means that the MOSF is more concerned by a high ROR than spending ministries, not only because the MOSF regards a high ROR among spending ministries as an unintended dysfunctional consequence of SABP, but also because a persistently high ROR among spending ministries can be criticised for not achieving the purposes of SABP by the National Assembly, the BAI and the news media. Moreover, the MOSF seems to have more responsibility for creating a successful SABP system as it is the designer and controller of the system, as well as the principal in the principal-agent relationship. However, spending ministries, as agents in the principal-agent relationship, seem to have more interests in high SABP results and bigger budgets than in the success of the SABP system itself.

In addition, analysis of the dynamic pattern of the ROR suggests that it might be possible to find an algorithm of the optimum response by spending ministries to any given level, or the exact pattern of the ROR, because spending ministries try to adjust specific programme's self-assessment score in the next SABP round after comparing the ROR of the specific programme to the average ROR of the spending ministry. If SABP is to be implemented repeatedly in the future, then finding an algorithm of the optimum response by spending ministries might be possible.

7.2.3 Validating the Data, Findings and Interpretations

This section discusses the quality of the quantitative data, the results and the interpretations. Regarding the validation of the data and findings, the study focuses on the internal and external validity, the possibility of causal reference, and the reliability of data.

Internal and External Validity

Cook and Campbell (1979, p.37) and Dooley (2001, p.268) divide the validity of quantitative social research into four categories: "statistical conclusion validity, internal validity, construct validity and external validity[30]". However, because internal validity includes statistical conclusion validity and external validity contains construct validity,

[30] If there is a relationship between two variables, statistical conclusion validity is accepted. Construct validity is concerned with the extent of consistency between the particular cause and effect constructs involved in the relationship (Cook and Campbell, 1979)

this section corroborates the internal and external validity of the study. According to Cook and Campbell (1979), internal validity is related to the exactness of the relationship between the independent and dependent variables. In addition, external validity is associated with the extent to which a particular result may be generalised to other persons, settings, and times.

With respect to internal validity, Campbell and Stanley (1966) classify the threats to internal validity as maturation, history, selection, mortality, regression, testing, and instrumentation, and these are related to the passage of time, additional events during the study period or the extent to which a study is based on experiment (Namkoong, 2010). However, these kinds of threats to internal validity do not appear pertinent to the quantitative analysis in this study, because the study uses numerical data that are drawn from the past and already fixed (i.e., SABP results and budget allocations from 2005 to 2010) and does not divide the subjects (budgetary programmes) into control and experimental groups, as experimental studies do. In addition, in order to establish the internal validity of variables, this study chose the dependent, independent and control variables from the relevant literature, such as Gilmour & Lewis (2005, 2006a), Park (2005, 2008), Bang (2009), Cho (2010), Chang &Yoon (2002), Kong et al. (2007), Yoon (2001). Moreover, the finding that all F ratios were statistically significant at the level of 0.01 means that the regression line was satisfied with the goodness-of-fit, and therefore the interpretations of the results mentioned above can be said to be valid. However, the analysis might be vulnerable to queries on other types of internal validity. For example, it may not fully consider extraneous variables that may affect the DR and ROR, such as the additional political characteristics of a programme (e.g., presidential

agenda programme, programme of interest to the National Assembly). Consequently, it might be said that consideration of more extraneous variables is necessary in order to improve the internal validity of the analysis. However, adding “presidential agenda programme” and “programme of interest to the National Assembly” as control variables was very difficult, because it needed a great deal of work and plenty of help from the MOSF officials. Adding these variables as control variables needs to be considered in further research.

As for the external validity, although the quantitative data in the study were drawn from 2,920 budgetary programmes from 2005 to 2010, a consistent policy trend towards performance budgeting will be followed by the Korean government in the future. Therefore, the findings and interpretations of the quantitative analysis can be applied to other budgetary programmes at different times. In particular, the results of the study could well be applied to the assessment of Research and Development (R&D) programmes and other types of performance assessment. As a result, it can be said that the data, findings and interpretations of the quantitative analysis have relatively high internal and external validity.

Possibility of Causal Reference

One of the main aims of quantitative analysis is to explain the relationships between variables. This leads to a consideration of whether the findings are also intended to verify the causal relationship between two variables. There have been many controversies concerning cause, effect, and the characteristics of relationships. Until

recently, the conditions of causal reference suggested by J. S. Mill have been generally accepted among scientists. J. S. Mill (1834) suggests three principles of causal reference: temporal precedence, constant conjuction, and exclusion of rival hypotheses. According to Johnson et al. (2008) and Belli (2008) three conditions are necessary in order to argue that a variable X causes a variable Y. These are: (i) variable X and variable Y must be related (the relationship condition), (ii) changes in X must happen before observed changes in Y (the temporal antecedence condition), and (iii) there is no possible alternative explanation for the relationship between X and Y (the lack of alternative explanation condition). Firstly, the relationship condition can easily be ascertained from the preliminary analysis results. According to the relationship analysis in Chapter 5, we can see that correlation coefficients and t-test results between variables were all statistically significant. Thus, the first condition for causal reference is fully satisfied. Secondly, the time order condition between the independent and dependent variables is also fully satisfied, because the DR and ROR, the dependent variables, were determined by the results of the independent variables: “the MOSF’s review score” and “specific score minus average score”. This might mean that the independent variable always precedes the dependent variable. However, the analysis has a disadvantage because it was difficult to find out and operationalise all the extraneous variables which might affect the DR and ROR, such as the presidential agenda and interest from the National Assembly, as mentioned above. Consequently, the findings of the quantitative analysis seemed not to succeed in verifying the causal reference between variables. Rather, the results were simply enough to explain meaningful associations between variables and to verify the expectations in question.

Reliability of Data

Quantitative reliability means that other researchers can reach consistent and stable results by using the same data (Creswell and Plano Clark, 2011; Namkoong, 2010). The study employed the SABP results from 2005 to 2010 published by the government, so there might be no difference among researchers when it comes to measuring variables. Thus, the quantitative analysis established the reliability of the data and findings.

7.3 Implications of the Qualitative Analysis

Although the qualitative analysis for the research successfully explored the types of, extent of, and reasons for dysfunctional consequences of the SABP system, the impacts of these consequences, and feasible policy alternatives, there may still be doubts about its findings. For example, (1) the influence of political factors on the MOSF's decision making; (2) the role of interviews with MOSF officials as colleagues or supervisors regarding the existence and extent of dysfunctional consequences; (3) the appropriateness of the categorisation of unintended and intended dysfunctional consequences by spending ministries; (4) the impact of dysfunctional consequences; (5) and the feasible policy alternatives for reducing or preventing these consequences. This section discusses these issues one by one.

7.3.1 Influence of Political Factors on the MOSF's Decision Making

According to the first interviews, the MOSF seemed to focus on economic rationalisation and to keep to the principles of SABP as much as possible. Conversely, although the performance indicators for P1 and P2 related simply to outputs – construction rate and budget implementation respectively – the MOSF changed its decision from a first "No"' to a final "Yes" in Question 2-1, as mentioned in 6.5.1. Thus, it is possible to argue about whether the MOSF's strategy is really economic rationalisation and whether its decision is not influenced by political pressure.

One possible reason may be that P1 and P2 are the government fund programmes, while the other six programmes are all general account programmes. In general, government fund programmes[31] are subject to weaker control by the National Assembly than general account programmes. Consequently, the MOSF might change its decision without too much trouble. The other, more plausible, reason may be the excessively penalty-oriented operation of SABP. That is, even though a programme is very important, if the programme is given a poor grade, then the MOSF has to reduce its budget for the following year by more than 10 percentage points. This strict ten percent rule may affect the MOSF's decision making. However, considering the interviews with spending ministry officials, as well as those with MOSF officials, the case studies suggest that the MOSF eagerly tries to obey the principles of SABP, especially in terms of strictly

[31] Government fund is established to accomplish specific purposes in specific fields, and the main revenue source for it is a surcharge based on a special law. However, the main revenue source for general account programmes is the tax people pay. Consequently, those running government fund programmes have more discretion than those running general account programmes, and the former are less controlled than the latter. In 2012, there were 65 government funds in Korea.

reviewing generous self-assessment results. For example, in the similar cases of P6 and P8, the MOSF gave a “Yes” answer to Question 2-1 on the condition that spending ministries would develop outcome-related and representative performance indicators later. Moreover, the results of the quantitative analysis also support the statement that the MOSF drastically cut the results of spending ministries’ optimism bias. From these interpretations, I can draw out two important implications. Firstly, the fact that P1 and P2 finally received a “Yes” answer in Question 2-1 from the MOSF does not support the spending ministries’ pursuit of non-quantifiable and long-term performance of a programme. Rather, it is the result of programme characteristics (government fund programme rather than general account one) or the excessively penalty-oriented operation of SABP. Secondly, in order to increase the effectiveness of the SABP system, the excessively penalty-oriented operation of SABP surely needs to be changed in some ways although the MOSF does its best to keep the principle of economic rationalisation and not to be affected by political pressure.

7.3.2 The Role of Interviews with MOSF Officials

It is possible to argue about whether exploring dysfunctional consequences of SABP, or the differences between spending ministries and the MOSF, using interviews with spending ministry officials is appropriate or not, not only because this issue is very sensitive and complex but also because we cannot be sure that spending ministry officials answered questions honestly. However, considering the attitude of spending ministry officials during the interviews and the results of the case studies, spending ministry officials may have responded very honestly, although the questions in the

interviews were very sensitive and difficult to answer. For example, responding to questions about three types of dysfunctional consequences – tunnel vision, myopic management and the ratchet effect – spending ministry officials strongly agreed that these existed. In the case of the ratchet effect, 17 out of 19 spending ministry officials answered honestly that they set the performance target levels as conservatively as possible because they were considering the next year, and there was no denial of this. Although we accept the honesty of spending ministry officials, however, some questions still remained unanswered, especially regarding how to interpret the difference in opinions on the existence of the other four dysfunctional consequences. In particular, in the case of cherry-picking, the MOSF officials partly agreed on the existence of this phenomenon whilst spending ministry officials strongly denied it. Here, we need to consider the role of the MOSF officials' interviews. The MOSF officials' opinions on the behaviours of spending ministry officials play a particularly important role in exploring dysfunctional consequences of SABP, because spending ministry officials know well that their behaviours relating to dysfunctional consequences do not follow the SABP manual and do not contribute to achieving a genuinely good performance from a programme. Thus, the spending ministry respondents might be likely unconsciously to underreport their actual perceptions of dysfunctional consequences, which could lead to methods bias. In other words, it is always possible that spending ministry officials gave socially desirable answers, even though they remained anonymous (Sackett et al., 1989). Under these circumstances, evaluations from other people, such as their colleagues or supervisors, by providing external criteria, can give a useful insight into the nature of dysfunctional consequences (Kwok et al., 2005). In particular, MOSF officials working at a similar level to spending ministry

officials may give more accurate opinions and evaluations of dysfunctional consequences of spending ministries, emphasising “task relevant abilities and competencies in appraising others” (Mumford, 1983, p. 873). As a result, although spending ministry officials answered honestly, there is still a possibility of unconscious underreporting of dysfunctional consequences, especially in the case of intended ones such as misrepresentation and cherry-picking. Consequently, policy makers need to consider possible underestimations by agents, especially regarding intended dysfunctional consequences, when it comes to refining the SABP system.

7.3.3 Appropriateness of Unintended and Intended Dysfunctional Consequences

Some may argue whether the categorisation of dysfunctional consequences into unintended and intended ones by spending ministries, and all unintended ones by the MOSF, based on the principal-agent relationship, is appropriate.

With respect to four unintended dysfunctional consequences by spending ministries, firstly, the case studies indicated that spending ministry officials voluntarily made some efforts to prevent or minimise them, as mentioned in 6.6. This result may confirm that four dysfunctional consequences are unintended by spending ministries. Secondly, the reasons given for unintended dysfunctional consequences in the interviews were not related to the intentions or tendencies of spending ministries, but rather associated with the characteristics of programmes in the public sector and factors which are difficult to control in the short-term. For example, in the case of tunnel vision, most interviewees indicated the ease with which it was possible to evaluate and increase the reliability and

fairness of some assessments as the reason for this phenomenon. Similarly, Smith (1995) points out that the existence of diverse objectives among various stakeholders, the difficulty of quantifying many important performances appropriately, and various unanticipated externalities in a public sector programme might lead to focusing on a small number of quantifiable performance features. De Bruijn (2002) also comments that measuring performance in the public sector includes many aspects and so requires professionalism. In the case of myopic management, most interviewees cited the characteristics of public programmes and the unique features of Korean government culture, such as one-year rotation of personnel and the short tenure of ministers. These problems cannot be solved by spending ministries in the short-term. Smith (1995) also indicates that the myopia problem can be worsened by the short-term careers of workers and that many outputs in the public sector can be achieved by long-term efforts and strategies, so the current performance indicators can result in managerial myopia. In the case of measure fixation, many interviewees cited the characteristics of public programmes, and the limitations of time and cost. Similarly, Smith (1995) indicates the impossibility of fully representing all associated objectives as the reason for this phenomenon. In the case of suboptimisation, many interviewees suggested the unique features of Korean government culture, such as sectionalism, one-year rotation of personnel and the short tenure of ministers. Similarly, Smith (1995) cites the strict hierarchy of organisations and the difficulty of combining the joint outputs of several agencies as leading to suboptimisation. De Bruijn (2002) also states that excessive competition among organisations makes them less willing to share best practice with each other. Consequently, considering the case studies and relevant literature, these four dysfunctional consequences are unintended by spending ministries.

On the other hand, with regard to the three intended dysfunctional consequences, firstly, the case studies confirmed that spending ministries did not make any effort to prevent or minimise them, as mentioned in 6.6. This evidence may confirm that the three dysfunctional consequences are intended by spending ministries. Secondly, the reasons for intended dysfunctional consequences given in the interviews and the literature were closely related to the intentions or tendencies of spending ministries. In the case of the ratchet effect and misrepresentation, most interviewees pointed out the information asymmetry between spending ministries and the MOSF, and the excessively penalty-oriented operation of SABP, as the reasons. As for the ratchet effect, Smith (1995) also indicates the controlling body's inability to produce comparable evidence. Similarly, Hood (2006) and Bevan and Hood (2006) comment that information asymmetry between the Treasury and spending departments can be the reason for a ratchet effect. In the case of misrepresentation, Smith (1995) points out the lack of professional judgement in describing an event and the limited external audit. In the case of cherry-picking, Kelman and Friedman (2009) cite the job-training organisation's deliberate gaming in which they made no contribution to jobseekers' prospects. Thirdly, researchers' definitions of the ratchet effect, misrepresentation and cherry-picking included deliberate intent by agents, such as "the tendency of managers" and "the deliberate manipulation by agents" (Simth, 1995; Bevan and Hood, 2006; LeGrand and Bartlett, 1993). These findings by other researchers indicate that the categorisation of dysfunctional consequences suggested by this research may be appropriate; and they give useful insights for exploring complex phenomena based on the principal-agent relationship more dynamically.

7.3.4 Impacts of Dysfunctional Consequences on SABP Results

This section discusses the impacts of dysfunctional consequences on the behaviours of spending ministries and the MOSF through both the qualitative and the quantitative analysis results. The results of the supplementary interviews indicated that various dysfunctional consequences might lead both to spending ministries' optimism bias and the MOSF's drastic cutting of spending ministries' self-assessment results. These results are also well supported by the results of regression analysis in Chapter 5. The quantitative analysis shows that the MOSF's review score has negative relationships with both the DR and ROR, and it means that there is an optimism bias by spending ministries when it comes to the self-assessment stage, because if spending ministries self-assessed their programmes as strictly and objectively as the MOSF, it is not certain that the DR or ROR would have negative relationships with the MOSF's review score in a statistically significant way. In addition, the finding that specific programme's self-assessment score minus and the spending ministry's average score in that year has positive relationships with the DR and ROR accounts for the MOSF's drastic cuts to spending ministries' generous self-assessment results. Furthermore, the interviews indicated that the difference between spending ministries and the MOSF mainly stems from incongruence in Question 2-1 (the appropriateness of performance indicators), because if a programme receives a "No" in Question 2-1, the programme cannot achieve a score of more than 65 points. These comments verify the argument of this study, which insists that dysfunctional consequences in performance planning need to be considered as more important than those in the performance measurement and

performance reporting stages. Although the MOSF started in 2010 to give penalties in cases where the SABP score gap between spending ministries and the MOSF was over 20 points, we need to watch the effectiveness of this policy for some time, while remaining alert to the possibility that score gaps converge around 20 points. Smith, (1995, p.305) similarly, indicates that "in practice, many of the outputs of the public sector are difficult to identify and measure; many are produced jointed with other organisations; and many unfold over a long period", and so, if performance measurement systems in the public sector do not manage and seriously consider these characteristics, the systems must fail. This is why seeking more feasible policy alternatives for minimising various dysfunctional consequences of SABP, which will be discussed in the next section, is so important.

7.3.5 Feasible Policy Alternatives

This section mainly discusses the feasibility of policy alternatives for preventing or minimising dysfunctional consequences of the SABP system, especially focusing on six policy alternatives which both the first and supplementary interviewees were asked about, and it then summarises feasible policy changes.

Refining the SABP Checklist

The findings from the first and supplementary interviews as well as previous studies on dysfunctional consequences of performance budgeting indicate the necessity of refining the SABP checklist by adding or revising questions in order to reduce various

dysfunctional consequences. In particular, most interviewees in both interviews suggested the addition of questions asking spending ministries to develop performance indicators relating to the non-quantifiable, long-term, non-measurable and whole performance of a programme, and the submission of evidence of spending ministries' efforts to achieve these performances.

However, unlike in the first interviews, in the supplementary interviews seven interviewees (two MOSF, two spending ministry, three expert) explicitly disagreed with the policy alternative for the following reasons. Firstly, they insisted that the non-quantifiable, long-term, non-measurable and whole performance of a programme can be considered under the present SABP checklist. That is, Question 2-1 requires the spending ministries to develop "representative" and "outcome-related" performance indicators, and Question 4-2 asks them to implement "comprehensive and objective programme evaluation" for measuring the qualitative as well as quantitative performance of a programme. Secondly, requiring evidences of spending ministries' efforts might just lead to seriously aggravating the ministries' work burden without any big improvement in performance. These results of supplementary interviews provide a more diverse perspective from interviewees than the first interviews, and they help us understand more thoroughly the perceptions of government officials on these complex and sensitive issues.

These opinions seem to be reasonable, but they are missing some important points. Firstly, although the present SABP checklist contains questions that can help prevent or reduce tunnel vision and measure fixation (Questions 2-1 and 4-2), it certainly does not

contain direct questions asking about the long-term and whole performance of a programme, and so it might lead to myopic management and suboptimisation. In particular, one expert who was involved in developing the SABP system mentioned that the question of whether spending ministries should be required to submit information about the medium- or long-term performance of a programme was discussed at the initial stage of devising the SABP system, but this policy was finally not accepted by the MOSF. In addition, the Public Institutions Policy Bureau of the MOSF recently reported that the guidance for assessing public institutions' performance had a new question asking public enterprises to submit their medium- or long-term performance indicators to the MOSF (MOSF, 2012b). This policy change in the assessment of public institutions has useful implications for further improving SABP in the future. Consequently, in order to prevent or minimise myopic management and suboptimisation, adding direct questions asking about the long-term and whole performance of a programme might be inevitable. Secondly, regarding spending ministries' work burden, I think that increasing the workload in the way suggested can be justified. The most important thing is to reduce any unnecessary work burden on spending ministries, not to avoid necessary work that would improve the SABP system. However, policy makers in the MOSF need to consider ways of keeping down the work burden on spending ministries as much as possible.

Changing Penalty-Oriented Operation

Most interviewees in the first and supplementary interviews agreed that the penalty-oriented operation of SABP needed to be reconsidered. However, in the supplementary

interviews nine interviewees (three MOSF, two spending ministry, four expert) among 31 disagreed with this policy change. In addition, five interviewees (four spending ministry, one expert) mentioned that the policy change would not improve the genuine performance of a programme. In order to assess the feasibility of the policy change, we need to consider these opinions carefully.

The main reasons for differing perspectives are as follows. Firstly, the purpose of SABP is to increase the efficiency of budget allocation, so penalty-oriented operation is inevitable. Secondly, spending ministries still do not understand the purpose of performance budgeting well, so a performance culture has not been widespread until now. Thirdly, in order to prevent spending ministries' overestimation and moral hazard, penalty-oriented operation might be inevitable. Fourthly, this policy change could bring other negative effects without enhancing the genuine performance of a programme.

However, these opinions seem to ignore the following three points. Firstly, the acceptance level among spending ministries for the present SABP system is quite low, and the main reason for this is the system's excessively penalty-oriented operation. Spending ministry officials argue that there is no reason for making an effort to increase the performance of a programme, because the incentive for an effective programme is relatively small compared to the penalty for a poor programme. A low level of acceptance has led to various dysfunctional consequences and to not achieving either efficiency of budget allocation or the improvement of performance capability among spending ministries. Secondly, penalty-oriented operation tends to lead to an increase in spending ministries' optimism bias and their moral hazard, rather than reducing these.

Thirdly, the most important thing we need to consider is that even the MOSF has changed its mind when it comes to reviewing spending ministries' self-assessment results in the case of politically important programmes, such as P1 and P2, just in order to avoid poor grades and a 10 percent cut. Taken together, these points suggest it is time to reconsider the excessively penalty-oriented operation of SABP.

In addition, the fourth divergent opinion on this policy change shed light on useful implications for policy makers when refining the SABP system. That is, policy makers in the MOSF should consider cautiously other unexpected side effects that might be caused by any policy change, because it is always possible for the spending ministries, as agents in the principal-agent relationship, to try to play some tactics under the new policy situation. This implication can be also applied to other policy alternatives explained in this section.

Relative Evaluation System

Only three interviewees mentioned this policy change in the first interviews and their answers were not specific. However, the supplementary interviews treated this policy change as one of the main issues, and 29 interviewees answered the questions. The opinions of these respondents could be divided into three types: agree (ten interviewees), partly agree (ten interviewees) and disagree (nine interviewees). Considering these three categories of opinion, I ascertained that 19 interviewees were concerned at the possibility of other unexpected side effects resulting from the policy change. They indicated that the policy change could lead to spending ministries' strategic behaviour

when it comes to deciding on which programmes to grade as poor, and the proportion of programmes graded as effective or poor could be determined by the internal dynamics of each ministry, not by the objective assessment of performance. In addition, they also stated that the policy change had been temporarily implemented from 2011 on a trial basis, so the figures of 20% and 10% were decided rather arbitrarily by the MOSF without any reasonable evidence.

Overall, it seems to be difficult to say at this point whether the relative evaluation system is feasible or not. It will take more time to evaluate the feasibility of the policy alternatives for reducing dysfunctional consequences of SABP. However, the opinions of interviewees raised plenty of controversial issues in terms of adopting policy changes and the necessity of resolving unexpected dysfunctional consequences that might be caused by them. Consequently, further research relating to this policy change needs to include the following issues. Most of all, it needs to address how to prevent spending ministries' other strategic behaviours, such as deliberately labelling the programmes the MOSF or the National Assembly are interested in poor grade programmes. It also needs to address whether compulsory operation by the MOSF or voluntary operation by the spending ministries is the more suitable option; whether using the ratio of the number of programmes or that of expenditure, as Canada's Strategic Spending Review does, is a suitable option; and what the appropriate proportion of effective and poor grade programmes should be, and how to achieve this.

Elevating the Assessment Level from "Sub-programme" to "Programme"

In the first interviews, 12 among 38 interviewees mentioned this policy change, especially for minimising myopic management and suboptimisation. However, as this policy alternative was one of the main issues dealt with in the supplementary interviews, all 31 interviewees had the opportunity to think over the advantages and disadvantages of such a policy change and to think over the answers they had made to the questions, and consequently two groups, each with a similar number of respondents, expressed opposite opinions on it. However, this result may support the idea that repeated interviews on the same issues with the same interviewees at intervals of time are a more effective way to explore the exact perceptions of participants, and this kind of interview method may improve the reliability of findings.

The interviewees who agreed with the policy change in both the first and supplementary interviews had similar reasons for accepting the necessity of elevating the present assessment level of SABP. They focused on the general limitations of the SABP system based on the characteristic of performance budgeting in the public sector, such as the difficulty of assessing joint outcomes in several sub-programmes. However, a similar number of respondents to the supplementary interviews paid greater attention to more realistic problems that would be brought about by elevating the assessment level. They were concerned over the difficulty of applying SABP results to budget allocation effectively.

The reasons of each group seemed to be very persuasive for each other, so I think that it would be better to revisit the purposes of SABP when dealing with these controversial arguments. As mentioned earlier, the SABP system was introduced to increase the efficiency and effectiveness of budgetary programmes (at sub-programme level) by reflecting SABP results in budget allocations in response to spending ministries' greater autonomy under the "top-down" budgeting system (MOSF, 2011b; KIPF, 2008). That is, the SABP system needs to focus on the appropriate assessment level – a level which can provide useful performance information when it comes to budget allocation and performance management. In this regard, the present sub-programme level seems to be more appropriate as the SABP assessment level. That is also why many OECD countries apply performance budgeting as an analytic tool for better budget allocation (Shick, 2007).

However, although we accept sub-programme as a more appropriate assessment level for the present SABP system, some questions raised by other groups of interviewees still remain unsolved – e.g., how to consider the long-term and whole performance of a programme, and how to consider joint outcomes associated with several sub-programmes. I think that the first question can be solved by refining the SABP checklist, changing the penalty-oriented operation of the system, and by other policy alternatives which will be discussed later. As for the second question, the IES can consider joint outcomes of cross-cutting programmes. As a result, elevating the assessment level from sub-programme to programme, as suggested by many interviewees in the first interviews, seems not to be feasible any more. However, it is possible that, as mentioned by one expert, the policy alternative could be considered again after distinct

relationships between the sub-programme and programme levels have been established in the near future.

Building Trust and Giving More Discretion to Spending Ministries

As with the policy discussed above, there were big differences in opinion concerning this policy alternative between the first and the supplementary interviews. In the first interviews, nine interviewees mentioned broadly the necessity for this policy change in order to reduce various dysfunctional consequences of SABP. However, when dealing with policy change as one of the main issues, all 31 interviewees answered questions about the proposal, and the responses fell into three groups: 15 interviewees unconditionally agreed with the proposal, seven conditionally agreed, and nine disagreed. By interviewing more than once on this policy alternative, I was able to explore the exact perceptions of the participants on these complex and sensitive issues.

There may not have been so much difference between the first and supplementary interviews regarding the necessity for building trust between spending ministries and the MOSF in order to reduce various dysfunctional consequences of the SABP. De Bruijn (2002) also indicates that mutual trust and the rules of the game regarding product definitions, functions and meanings need to be embedded in the process of performance measurement in order to reduce various perverse effects. Similarly, Kelman and Friedman (2009) and Heinrich (2007) suggest encouraging the public service motivation or stewardship of public employees towards cooperative behaviour to prevent negative effects of performance management.

However, interviewees in the supplementary interviews expressed different opinions about whether the MOSF should give more discretion to the spending ministries. Considering these opinions and the importance of reducing dysfunctional consequences in the performance planning stage, cooperation between the spending ministries and the MOSF when it comes to developing performance indicators and performance target levels might be more important than increasing spending ministries' discretion unconditionally. Although it will eventually be necessary to increase the spending ministries' discretion, in the long-term, it is also essential to require spending ministries to be more accountable, responsible, objective and reliable in the short-term. As a result, increasing the appropriateness of performance indicators and performance target levels through cooperation between the spending ministries and the MOSF, as well as enhancing spending ministries' discretion on other procedures of the SABP, such as programme management and specific methods for achieving performance targets, might be necessary at the same time. It is possible that this policy alternative became more precisely specified between the first and the supplementary interviews.

Abolishing Spending Ministries' Self-Assessment Stage

With respect to this policy alternative, the supplementary interviews provide a more complete understanding, because only two interviewees in the first interviews suggested the policy change. Two groups containing similar numbers of respondents gave opposite opinions in the supplementary interviews. Interestingly, the number of interviewees from the MOSF, the spending ministries and the experts in each group was

nearly same. More interestingly, and contrary to initial expectation, seven spending ministry officials disagreed with the policy change, while seven agreed with it. At first glance, all spending ministry officials seemed to agree with the policy change, because the MOSF changed nearly all the self-assessment results of the spending ministries, and so most spending ministry officials seemed to regard the self-assessment stage as useless. However, nearly half of spending ministry officials in the supplementary interviews put an emphasis on the opportunity for increasing their responsibility over the programme and enhancing their performance management capability. Considering the purposes of SABP, which are to improve the spending ministries' ability to manage performance and to establish and spread a performance management culture among spending ministries, as well as to increase budgeting efficiency, these opinions seem to be very reasonable. As a result, abolishing the spending ministries' self-assessment stage might not be feasible, not only because the self-assessment stage can lead to increasing interest in SABP as well as strengthening spending ministries' responsibility and accountability, but also because spending ministries have a useful opportunity to experience a learning process through the self-assessment stage.

Summary of Six Policy Alternatives

Out of the six policy alternatives, I was able to confirm the feasibility of two: refining the SABP checklist; and changing the penalty-oriented operation of the system. Building trust and giving more discretion to spending ministries were found to need more detail, both on strengthening cooperation regarding performance indicators and performance target levels and on enhancing discretion as to other parts of the procedure.

The creation of a relative evaluation system needed to be reconsidered after reviewing the results of a trial implementation of such a system. However, elevating the assessment level and abolishing the self-assessment stage did not seem to be feasible policy alternatives when the purposes of SABP were taken into consideration. Before elevating the assessment level from the present sub-programme to programme level, it might be necessary to establish a more explicit relationship between the sub-programme and programme levels. Abolishing the self-assessment stage of spending ministries was also found not to be feasible. A summary of discussions regarding the six policy alternatives is given in Table 7-1.

Table 7- 1 Summary of Discussions Regarding Six Policy Alternatives

Policy alternative	Feasibility
Refine the SABP checklist	Feasible
Change the penalty-oriented operation	Feasible
Relative evaluation system	Needs to be reconsidered later
Elevate assessment level	Not feasible
Build trust and give more discretion to spending ministries	Needs to be more clearly specified
Abolish self-assessment stage	Not Feasible

Besides the interviewees' evaluation of the feasibility of these policy changes, it is notable that the MOSF officials rated the importance of most of the changes more highly than did the spending ministry officials. This might imply that the MOSF officials took potential changes to the SABP system more seriously, as the principal in the principal-agent relationship, and this result is also well supported by the implications of the dynamic pattern of ROR of spending ministries mentioned in Section 7.2.2.

Summary of Feasible Policy Alternatives for Preventing or Reducing Dysfunctional Consequences of SABP

First I extracted 14 policy alternatives for preventing or reducing various dysfunctional consequences of SABP from the first interviews, and then, through the supplementary interviews and reviews of relevant studies, I evaluated the feasibility of six policy alternatives which seemed to be thought worth discussing. Among 14 suggested policy alternatives, Table 7-2 summarises nine feasible policy alternatives. Five policy alternatives need to be reconsidered later or are not feasible. In particular, the impacts of two policy alternatives that related to the restructuring and reengineering of government departments will need to be looked at in future research, because they need more in-depth analysis and it will take long time to evaluate these kinds of policies. Moreover, Table 7-2 indicates the main actors involved in each policy alternative, in order to suggest more customised policy alternatives for related organisations.

Table 7- 2 Summary of Feasible Policy Alternatives

Policy alternatives	Unintended dysfunctional consequences by agents				Intended dysfunctional consequences by agents			Main actors
	Tunnel vision	Myopic management	Suboptimisation	Measure fixation	Ratchet effect	Misrepre Sentation	Cherry-picking	
Options for urgent policy change								
Refine SABP checklist	√	√	√	√				MOSF
Change penalty-oriented operation	√	√	√	√	√	√	√	MOSF
Senior officials' interest	√	√	√	√	√	√	√	Both
Options for short-term policy change								
More active role (monitoring, penalty, education, etc.)	√	√	√	√	√	√	√	All related organisations
Open evaluation system	√	√	√	√	√	√	√	All related organisations
Report SABP results publicly						√	√	MOSF

Reduce unnecessary workload		√		√				MOSF
Options for medium- or long-term policy change								
Increase expertise	√	√	√	√	√	√		Both
Build trust, strengthen cooperation and give more discretion	√	√	√	√	√	√	√	Both

Notes 1. Modified Smith (1995) 2. The symbol '√' indicates a putative positive effect 3. "Both" in main actors indicates the MOSF and spending ministries

From the interviews and the review of relevant literature, it can be said that refinement of the SABP checklist may reduce four unintended dysfunctional consequences by the spending ministries, and the other two options for urgent policy change are likely to address all the dysfunctional consequences of SABP, as mentioned earlier.

As for options for short-term policy change, firstly, a more active role for related organisations, such as the MOSF, the National Assembly and the BAI, and external actors like the ombudsman, grassroots organisations and client panels mentioned by Thiel and Leeuw (2002), might reduce all dysfunctional consequences as well. This policy alternative includes very extensive policies: not just more fundamental approaches, such as developing and distributing a detailed SABP manual for appropriate performance indicators and performance target levels, or educating spending ministry officials, but also less fundamental approaches, such as strengthening the audit of performance data, and even giving penalties for deliberate manipulation of performance. Similarly, Smith (1995) indicates the necessity of keeping the system under constant review, maintaining a careful audit of data, and developing independent benchmarks. Bevan and Hood (2006) also suggest specifying performance targets and filling the audit gap for reducing the ratchet effect and output distortion. Heinrich (2007)

comments that in order to increase the effectiveness of a performance management system, evaluating and recognising performance achievements are necessary.

Secondly, an open evaluation system reaching from the initial stage to the assessment stage of a programme could also solve all dysfunctional consequences of SABP. This policy alternative involves prior consensus on performance indicators and target levels, and expert interpretation of performance indicators, as mentioned by Smith (1995). Similarly, Bevan and Hood (2006) suggest increasing face-to-face communication between governors and governed, and Heinrich (2007) indicates the necessity of increasing openness and fairness regarding the use of performance information to reduce the side effects of performance management. Reducing disagreement in the performance planning stage by an open evaluation system could be helpful in preventing subsequent dysfunctional consequences in the performance measurement and reporting stages. In addition, if there is no disagreement between spending ministries and the MOSF over performance indicators and their target levels, then the MOSF can just focus on evaluating whether spending ministries have achieved their performance target levels as expected, and the BAI and the National Assembly can just pay attention to verifying whether the grounds for performance achievement submitted by spending ministries are appropriate or not. That is, a more desirable division of roles between related organisations might be possible under an open evaluation system.

Thirdly, reporting SABP results publicly could be expected particularly to reduce the phenomena of misrepresentation and cherry-picking, because it could put pressure on spending ministries not to distort their performance and not to intentionally prioritise

groups that will produce favourable results. In addition, reducing unnecessary workload might give spending ministries more time to wait for the long-term and non-measurable performance of a programme, because many interviewees indicated lack of time and high costs as the main reasons for myopic management and measure fixation.

Among the options for medium- or long-term policy change, firstly, increasing the expertise of both spending ministries and the MOSF could be expected to reduce information asymmetry and enhance the MOSF's capability in evaluating the appropriateness of performance target levels and performance results. Thus, this policy alternative might address the ratchet effect and the misrepresentation caused by information asymmetry. In addition, this policy change is likely to address four unintended dysfunctional consequences, by increasing spending ministries' capability in performance management. Secondly, building trust between spending ministries and the MOSF, strengthening cooperation in terms of developing performance indicators and setting performance target levels, as well as giving more discretion to spending ministries in other processes of SABP, might also have a positive effect in reducing all dysfunctional consequences, especially those in the performance planning stage, such as tunnel vision, myopic management, suboptimisation and the ratchet effect.

Considering all feasible policy alternatives, more fundamental approaches, such as refining the SABP checklist, developing a detailed SABP manual for appropriate performance indicators and target levels, and educating government officials, mainly seem to address unintended dysfunctional consequences by agents, whilst less fundamental approaches, such as strengthening audits, sometimes giving penalties for

fraud, and reporting SABP results publicly, might generally resolve intended dysfunctional consequences by agents.

Moreover, these policy alternatives could be customised, according to the main actors in each programme. It is the MOSF that can be expected to play a key role in changing policies related to refining the SABP checklist, changing penalty-oriented operation, reporting SABP results publicly and reducing unnecessary workload. On the other hand, obtaining senior officials' interest, increasing expertise, building trust, strengthening cooperation and giving more discretion to spending ministries are roles both the MOSF and spending ministries can undertake. The remaining two policy alternatives are associated with all the related organisations, such as the MOSF, spending ministries, the National Assembly, the BAI, experts, citizens' groups, and the press. Here, though, there is one thing which needs to be carefully noted by policy makers: all policy makers should always consider the possibility that other unexpected dysfunctional consequences may be caused by these policy changes when it comes to redesigning the SABP system.

7.3.6 Validity of the Data, Findings and Interpretations

This section evaluates the validity of the data, findings and interpretations of the qualitative analysis. The coding of qualitative data was implemented by one person, me, so the reliability problem does not arise. Thus, the section focuses on the validity, determining that the account provided by the study is "accurate, can be trusted, and is credible" (Creswell and Plano Clark, 2011, p.211). Creswell and Plano Clark (2011)

indicate several strategies to check for qualitative validity, such as member checking, triangulation of the data, and asking for the opinions of others. Firstly, I sent the conference paper, which included my main findings, such as case study results, major themes and theoretical framework, to the participants in the first interviews and asked them whether my findings reflected their answers accurately, as well as whether I had respected their privacy. Most participants gave me feedback positively and some experts gave useful comments. Secondly, the study employed several sources, many participants and multiple case studies, as well as merging the qualitative and quantitative analysis results. That is, by not only interpreting multiple sources of evidence, such as documentation, archival records, and first and supplementary interviews from eight budgetary programmes in three case areas, but also merging related findings from both qualitative and quantitative analysis, the study can provide a more accurate and complete understanding of dysfunctional consequences of SABP. Finally, the transcripts of the interviews, the theoretical framework of the analysis, and the major findings were examined by my PhD colleagues, my supervisors and many experts at the conference. Also, I tried to explain the major findings to individuals who were not familiar with SABP system, such as government officials who had no experience of SABP. Consequently, it can be said that the validity of the data, findings and interpretations of the qualitative analysis is very high.

7.4 Contributions to Knowledge

This study contributes to academic knowledge in three areas of research. Firstly, it contributes to studies on dysfunctional consequences of performance budgeting in the

public sector. Secondly, it contributes to the application of the principal-agent approach in a Korean context. It has demonstrated that the principal-agent theory developed by Western researchers can be well applied to analysing dysfunctional consequences of the Korean performance budgeting system, SABP. Thirdly, it offers support to performance budgeting policy makers in analysing various dysfunctional consequences and considering the refinements that might bring a better performance budgeting system.

7.4.1 Contribution to Performance Budgeting Studies

The research focused on dysfunctional consequences of performance budgeting in the public sector in order to provide a theoretically-informed and empirically-rich account of them, and moreover it suggested feasible policy alternatives for minimising or preventing dysfunctional consequences. The study was mainly built on the work of Smith (1990, 1993, 1995, 2005), one of many researchers on the dysfunctional consequences of performance management in the public sector. It particularly supports his work (Smith, 1995) related to the existence of unexpected dysfunctional consequences of performance management and the necessity for policy makers to be aware of these as well as of feasible policy alternatives, while developing a more detailed theoretical framework to analyse various dysfunctional consequences. Furthermore, the results of the study can well be applied to examine not only other countries' performance budgeting but also other performance assessment systems in Korea, such as R&D programme assessment, regional development programme assessment and public institution evaluation. The specific contributions to performance budgeting studies are as follows.

Firstly, Smith (1995) suggests eight types of unintended dysfunctional consequences of publishing performance data in the public sector based on the principal-agent theory. However, this research attempts to suggest a more detailed typology of dysfunctional consequences based on the principal-agent relationship, some of them unintended by both principal and agents (tunnel vision, myopic management, measure fixation, suboptimisation), others unintended by the principal, but intended by the agents (ratchet effect, misrepresentation, cherry-picking), and to underpin it by extensive empirical evidence. Apart from Smith's work (1995), there have been few studies which categorise dysfunctional consequences into unintended and intended ones by agents, as this research does.

Secondly, the research highlights the relationship between the categorisation of dysfunctional consequences and the reasons for them. Although many researchers have indicated diverse reasons for dysfunctional consequences, and the reasons for each dysfunctional consequence are similar to those in this study, they have not explicitly linked the reasons with the categorisation of dysfunctional consequences (Smith, 1995; de Bruijn, 2002; Hood, 2006; Bevan and Hood, 2006; Kelman and Friedman, 2009). However, the findings from the interviews suggest that whether spending ministries make efforts to prevent dysfunctional consequences is closely related to their intentions with regard to dysfunctional consequences. In addition, the specific reasons for each dysfunctional consequence suggested by the interviews are also associated with the categorisation of dysfunctional consequences. That is, the reasons for unintended ones

are more related to the characteristics of public programmes, while those for intended ones are more associated with spending ministries' deliberate strategic behaviours.

Thirdly, this research sheds light on exploring the difference in the extent of dysfunctional consequences between principal and agents. There have been few studies describing the difference in the extent of perverse effects on the basis of the principal-agent relationship. Moreover, the research pays attention to the possibility of underreporting of intended dysfunctional consequences by spending ministries, and suggests that MOSF officials need to consider this tendency in spending ministry officials when it comes to refining the system.

Fourthly, this research focuses on finding feasible policy alternatives for reducing or preventing various dysfunctional consequences through two rounds of interviews and an extensive review of the relevant literature. Furthermore, the study divides options for policy change into urgent, short-term, medium-term or long-term, based on the number of respondents and perceived importance level of interviewees, in order to increase the timeliness of policy implementation. There have been few studies suggesting policy alternatives in the way this study does.

7.4.2 Contribution to the Application of the Principal-Agent Approach in a Korean Context

Firstly, this research has attempted to apply the principal-agent theory developed by Western researchers to the Korean performance budgeting system and to provide empirically-rich explanations for dysfunctional consequences of SABP. Although many

Korean researchers have analysed SABP, most of them have focused on examining the relationship between SABP results and budget allocation, or explaining the factors which affect budgeting or performance results (Park, 2005, 2008; Bang, 2009; Cho, 2010; Yoon, 2001; Kang, 2007). Even though some studies have indicated possible dysfunctional consequences of SABP, they have not applied the principal-agent theory in detail (Bang, 2009; Cho, 2010; Park, 2005; Keum and Lee, 2009; Shin, 2012). Consequently, there have been few studies which have analysed dysfunctional consequences of SABP on the basis of the principal-agent theory. This has provided the researcher with a good opportunity to broaden the scope of analysing dysfunctional consequences of performance budgeting based on the principal-agent theory from a focus on Western countries such as the UK and the US to findings from Korea.

Secondly, with respect to the methodology, there have been few studies of performance budgeting in either western countries or Korea which have employed both the quantitative and the qualitative approach. However, this research uses both approaches for analysing the quantitative and qualitative data provided by Korean government departments and officials. In order to explain and explore the difference in SABP results between spending ministries and the MOSF in Korea, it also adopts various kinds of research methods, such as correlation analysis, t-test, ANOVA and multiple regression analysis, analysis of documents and archival records, and in-depth interviews. In addition, the study has attempted to merge the results from two different kinds of approach, so the major findings of the quantitative approach support and strengthen those of the qualitative approach, and vice versa. For example, statistically significant relationships between two independent variables and two dependent variables support

the findings about spending ministries' optimism bias and the MOSF's drastic cutting of self-assessment results provided by the qualitative analysis; and the use of a higher adjusted R square in the performance planning section than in other sections strengthens the finding about the necessity for careful consideration in the performance planning stage in order to avoid subsequent dysfunctional consequences in the performance measurement and reporting stages. Furthermore, the research implemented interviews with almost the same group of participants twice with a time interval, and this led to a more complete and exact exploration of the perceptions of the participants concerning very complex and sensitive issues. For example, four out of six controversial policy alternatives produced different results between the first and supplementary interviews.

7.4.3 Contribution to Policy Makers

Although many studies on dysfunctional consequences of performance budgeting have applied the principal-agent theory to suggest useful policy alternatives, few studies have tried to match each policy alternative with suitable actors in detail. Also, there have been few studies carried out to evaluate each policy alternative's feasibility concretely, through repeated interviews, as this study has done. In addition, there may be main actors who play a key role in implementing each policy alternative, so the research has tried to indicate suitable actors for this. During this process, a better understanding of the complex interests and incentives among related organisations provides a good opportunity to discover feasible policy alternatives for diverse policy makers.

Firstly, policy makers of all related organisations should pay attention to the following. The research points out that making an effort to prevent or reduce dysfunctional consequences in the performance planning stage is important in order to avoid dysfunctional consequences in the performance measurement and performance reporting stages. In particular, prior consensus on the appropriate performance indicators and performance target levels between spending ministries and the MOSF might be essential for minimising dysfunctional consequences in the performance planning stage. This will especially reduce unintended dysfunctional consequences and the ratchet effect.

Secondly, at the same time, in order to reduce intended dysfunctional consequences more efficiently, policy makers in the BAI and the National Assembly should strengthen their auditing of spending ministries' performance reports. Also, policy makers in the spending ministries should develop an official database of data about their performance and use these data when submitting the account of their performance to the MOSF, the National Assembly and the BAI.

Thirdly, policy makers in both the MOSF and spending ministries should build trust with each other and try to increase their expertise in performance budgeting. Mutual trust and higher expertise can reduce information asymmetry and various dysfunctional consequences. Also, senior managers in both the MOSF and spending ministries should pay greater attention to SABP. Their interest may be essential to minimise all kinds of dysfunctional consequences of SABP.

Fourthly, policy makers in the MOSF should carefully consider changing the excessive penalty-oriented operation of SABP, refining the SABP checklist, reporting SABP results publicly, and reducing unnecessary workload, in order to prevent various perverse effects. In addition, MOSF officials, as mentioned earlier, need to be aware of possible underestimations of intended dysfunctional consequences when refining the SABP system, because spending ministry officials have a natural tendency to give socially desirable answers on these negative effects. Moreover, they need to consider more seriously the possibility of spending ministries' generous self-assessment results when it comes to reviewing directly managed programmes, ministries' programmes rather than agency ones, programmes that have lower budgets compared to the previous year, and economy-related programmes rather than welfare-related ones.

Fifthly, policy makers in spending ministries should try to reduce opportunistic behaviours based on the information asymmetry between them and the MOSF. Considering the nature of SABP, although it may not be possible to completely abolish spending ministries' deliberate and intentional behaviours, spending ministry officials themselves should make an effort to prevent intended dysfunctional consequences.

Finally, the study suggests that all policy makers should be cautious about designing and refining the performance budgeting system, considering both the unintended and the intended dysfunctional consequences already evident and those that changes might bring.

7.5 Conclusion

This chapter has discussed and interpreted several issues raised by the findings from the quantitative and qualitative analysis. For example, statistically significant relationships between two independent variables and the DR and ROR in the quantitative analysis supported well the findings of the qualitative analysis, especially the impacts of various dysfunctional consequences on the behaviours of both spending ministries and the MOSF. “No but Yes” type disagreements affect the relationship between control variables and the DR, and that may be the reason why some control variables did not have clear relationships with the DR. From the clear relationships between control variables and the ROR, several useful implications were extracted. Five major issues in the qualitative analysis were discussed, which were closely related to the research questions: the type of, the extent of, and the reasons for dysfunctional consequences of SABP; the impact of these dysfunctional consequences; and some feasible policy alternatives. These detailed discussions and interpretations were intended to confirm the answers to the four research questions. Also, the chapter has evaluated the validity and reliability of the data, findings and interpretations of both the quantitative and qualitative approaches, and concluded that both approaches have high validity and reliability. Finally, it has indicated that the study makes a contribution to knowledge in three key areas: to the study of performance budgeting (theoretical contribution); to the application of the principal-agent approach in a Korean context (empirical and methodological contributions); and to the work of policy makers (practical contribution).

CHAPTER 8

CONCLUSION

8.1 Introduction

In the public sector the concept of performance is often elusive, and there are arguments about the best way to implement a performance budgeting system. This thesis has not set out to focus on these arguments. Instead, it has attempted to examine possible dysfunctional consequences of a performance budgeting system, focusing on the process of SABP in South Korea, and suggesting feasible policy alternatives for refining the SABP system.

Both quantitative and qualitative approaches, with various kinds of research methods, were used for this study. The quantitative analysis showed that there is indeed a high and persistent optimism bias by spending ministries, as well as drastic cutting by the MOSF of spending ministries' generous self-assessment results. The case studies on which the qualitative analysis was based indicated that there are "unintended" and "intended" dysfunctional consequences of SABP by spending ministries, and that these can lead to both an optimism bias by spending ministries and a strict review of spending

ministries' results by the MOSF. In addition, the findings of both approaches shed light on developing feasible policy alternatives for improving the SABP system.

This chapter summarises the main answers to the research questions by drawing on the findings of both approaches, and then considers the study's limitations, as well as future research areas, before offering some closing remarks.

8.2 Answers to the Research Questions

This study dealt with the following four research questions by employing both quantitative and qualitative approaches: What are the factors which affect the differences in SABP results between spending ministries and the MOSF? What are the types of, extent of, and reasons for dysfunctional consequences of SABP? How do these dysfunctional consequences impact on both spending ministries' self-assessment scores and the MOSF's review scores? How can the SABP system in Korea be improved and refined? The first research question is related to the quantitative analysis, and the second and third questions are associated with the qualitative analysis, while the fourth is related to both the analyses. However, as we saw in Chapter 7, the results of both approaches were closely related to each other. The answers to the research questions examined from Chapter 5 to Chapter 7 are presented below.

In order to address the first research question, by employing quantitative analysis, this study firstly defined and operationalised the dependent variables (DR and ROR), the independent variables, and the control variables through a review of relevant literature,

followed by preliminary analyses – correlation analysis, scatter diagram, t-test, and multiple regression analysis – to verify the relationship between variables. Regression analysis showed firstly that two independent variables, such as the MOSF's review score and the specific programme's self-assessment score minus the spending ministry's average score for the year (specific score minus average score), had statistically significant relationships with the DR and ROR for all six years being studied, at the level of 0.01. The effective directions were negative and positive respectively. The former relationship implied that there is indeed optimism bias by spending ministries in their self-assessment stage, while the latter one indicated that the MOSF does drastically cut spending ministries' generous self-assessment scores. These findings provided a useful insight into the necessity for exploring potential dysfunctional consequences of SABP and the impacts of these on the behaviours of both spending ministries and the MOSF, using the qualitative approach, as well as for suggesting feasible policy alternatives for refining the SABP system. Secondly, regarding the relationship between control variables and the ROR, the study found that differences in SABP scores between spending ministries and the MOSF may be higher if a programme is directly managed by the government, is a ministry rather than an agency programme, was given a low budget percentage change for the current year, and is an economy-related rather than a welfare-related programme, regardless of programme budget size. In addition, the differences may be lower if SABP has been implemented on a programme several times. These findings implied that policy makers in the MOSF need to consider spending ministries' possible optimism bias more carefully when it comes to reviewing these programmes' self-assessment results, and that repeated SABP implementation brings positive learning effects. Thirdly, the study analysed the dynamic pattern of RORs

through the programmes which had been assessed by SABP two or three times from 2005 to 2010. The findings implied that the MOSF, as controller and designer of the SABP system, has seen itself as responsible for improving the SABP system, whilst the spending ministries, as agents in the principal-agent relationship, have paid greater attention to high SABP scores and the bigger budgets they bring.

With the second research question, the case studies (eight budgetary programmes in three case study areas) highlighted the types of, extent of, and reasons for various dysfunctional consequences of SABP based on the principal-agent relationship. To do this, the study extracted seven meaningful potential dysfunctional consequences of SABP through the review of the relevant literature, and then categorised them into "unintended" dysfunctional consequences (tunnel vision, myopic management, measure fixation, suboptimisation) and "intended" ones (ratchet effect, misrepresentation, cherry-picking) by the spending ministries, while accepting that all dysfunctional consequences are unintended by the MOSF, which is the principal in the principal-agent relationship. The case studies suggested that, in the case of unintended dysfunctional consequences, spending ministry officials, as the agents in the relationship, make some efforts to prevent them, whilst in the case of intended ones, they do not make any efforts to prevent them. In addition, the case studies indicated that the reasons for unintended dysfunctional consequences are closely related to intrinsic characteristics of public programmes, such as intangibility, indirectness, externality and complexity (Smith, 1995; Wilson, 1989; Thiel and Leeuw, 2002; de Bruijn, 2002), and factors which are difficult to control in the short-term, such as unique features of Korean government culture, whilst those for intended ones are more associated with the intentions or

tendencies of spending ministries to obtain high SABP scores and bigger budgets by using the information asymmetry between themselves and the MSOF or avoiding the excessively penalty-oriented operation of SABP. Moreover, the case studies confirmed that in the case of tunnel vision, myopic management and the ratchet effect, both spending ministries and the MOSF strongly agreed on the existence of these; but in the case of measure fixation, suboptimisation and misrepresentation, the MOSF strongly agreed on their existence while the spending ministries only partly agreed. And in the case of cherry-picking, the case studies suggested that the MOSF partly agreed on the existence of this, but spending ministries strongly denied its existence. However, after discussing the role of MOSF officials as colleagues or supervisors of spending ministry officials in Chapter 7, the study reached the conclusion that although we could agree on the honesty of spending ministry officials in terms of answers in the interviews, there was always the possibility of them unconsciously underreporting dysfunctional consequences, especially intended ones, as agents in pursuit of socially desirable answers (Sackett et al., 1989). Policy makers need to consider these possible underestimations when it comes to redesigning the SABP system.

With respect to the impact of these dysfunctional consequences on spending ministries' self-assessment scores and the MOSF's review scores, the topic guide in the first interviews did not include direct questions about this impact, and so its exploration of this was initially somewhat dependent on the interpretation of the first interviews and the documentary analysis. Consequently, in the supplementary interviews, the study directly asked about the impact of dysfunctional consequences on SABP scores, and the interviews suggested that various dysfunctional consequences led both to optimism bias

by spending ministries and to drastic cutting of these ministries' generous self-assessment results by the MOSF. The study also confirmed that these results were well supported by the quantitative analysis results, verifying the statistically significant relationships between two independent variables and the DR and ROR for six years. Furthermore, during the supplementary interviews many respondents indicated that the differences in SABP scores between spending ministries and the MOSF mainly came from incongruence in Question 2-1 (the appropriateness of performance indicators), which is logically related to subsequent questions such as Questions 2-2 and 4-1. Thus, the study confirmed that dysfunctional consequences need to be considered more seriously in the performance planning stage than in the performance measurement or reporting stages. These results shed light on the importance of seeking more feasible policy alternatives for improving the SABP system.

In order to suggest feasible policy alternatives for refining the SABP system, the study firstly extracted 14 policy changes from the first interviews, and then examined in detail, using the supplementary interview results and reviews of relevant studies, six policy alternatives that had aroused particular discussion. The study finally suggested nine feasible policy alternatives in the form of options for urgent, short-term, medium-term or long-term policy changes, specifying the main actors for each policy change. The six policy alternatives that aroused particular discussion were: refining the SABP checklist, changing the penalty-oriented operation of SABP, creating a relative evaluation system, elevating assessment from "sub-programme" to "programme" level in the programme budgeting structure, building trust and giving more discretion to spending ministries, and abolishing the spending ministries' self-assessment stage. The study found that

interviewees' perceptions changed between the first and supplementary interviews, and the reason might be that the interviewees had had an opportunity to rethink the advantages and disadvantages of each policy alternative in more detail, because in the supplementary interviews the six policy alternatives were dealt with as main issues. These results shed light on how repeated interviews on the same issues, with the same interviewees, and at intervals of time, can be an effective way to understand and explore the exact perceptions of participants on very complex and sensitive issues, and to increase the reliability of findings. The study looked at how, of the six policy alternatives, opinion on four changed from "feasible" in the first interviews to "not feasible", "need to be reconsidered later", or "need to be more clearly specified" in the supplementary interviews. After discussion, the study suggested refining the SABP checklist (MOSF), changing excessively penalty-oriented operation (MOSF), and encouraging senior officials' interest (MOSF and spending ministries) as options for urgent policy change; more active roles, such as developing a detailed SABP manual, educating, auditing or monitoring (all by related organisations), an open evaluation system (all by related organisations), reporting SABP results publicly (MOSF) and reducing unnecessary workload (MOSF) as options for short-term policy change; and increasing expertise (MOSF and spending ministries), and building trust and strengthening cooperation between spending ministries and the MOSF when developing performance indicators and target levels, while giving more discretion to spending ministries in the other procedures (MOSF and spending ministries), as options for medium- or long-term policy change. The study also suggested that more fundamental approaches, such as refining the SABP checklist, developing a detailed SABP manual, and educating government officials, seemed to be more closely related to addressing

unintended dysfunctional consequences by agents, while less fundamental approaches, such as strengthening the audit of performance data, sometimes giving a penalty for manipulation of data, and reporting SABP results publicly, might be more associated with reducing intended dysfunctional consequences by spending ministries. Finally, the study recommended that all policy makers should always be aware of the unintended and intended dysfunctional consequences which these policy changes might bring, as well as the dysfunctional consequences which are already evident.

8.3 Future Research Areas

This research has been successful in explaining and exploring the differences between spending ministries and the MOSF in the process of arriving at SABP results, and it has done this through both quantitative and qualitative approaches. However, further research with similar research topics on other policy areas, in Korea and other countries, could enrich the findings of this study. In particular, a cross-national comparison could provide researchers with an opportunity to find out the different dimensions of various dysfunctional consequences of performance budgeting in the public sector. For example, a further study might develop another useful typology for analysing perverse effects of performance management and exploring the reasons why agents need to game.

Secondly, further research is needed to examine whether the differences between the SABP results of spending ministries and the review results of the MOSF are high and persistent over time. Longer time-series data on SABP results might be useful to analyse whether SABP can achieve its initial purposes. In addition, it might enable

researchers to discover some algorithms for the optimum response by spending ministries to any given level of ROR through the analysis of dynamic ROR patterns. The MOSF also needs to prepare to reduce spending ministries' strategic behaviours by calculating the optimum response.

Thirdly, future research needs to increase the internal validity of the quantitative analysis by considering not only additional extraneous variables which may affect the difference in SABP results between spending ministries and the MOSF, such as the presidential agenda programmes and the programmes of interest to the National Assembly, but also more outcome-oriented variables, such as citizen's satisfaction or happiness level. In addition, more delicate operationalisation of the economy-related programmes might be helpful in explaining the unexpected relationship between this variable and the DR.

Finally, future research needs to consider the advantages of the US GPRAMA and the possibility of applying the provisions of this act to SABP in terms of monitoring several high priority performance goals and linking performance information with a dashboard for constant review. In addition, there may be plenty of scope for further research related to the issues of relative evaluation systems and elevating the assessment level from sub-programme to programme level. For example, with the relative evaluation system, several issues, such as whether compulsory operation by the MOSF or voluntary operation by spending ministries is preferable; whether the criterion of the ratio of the number of programmes or that of the ratio of expenditure, as in Canada's Strategic Spending Review, is preferable; what the appropriate proportions are and how

to create these; and how to prevent spending ministries' other strategic behaviours, need to be considered. Moreover, it would be worthwhile future research considering the impact of the restructuring or reengineering of related organisations on the dysfunctional consequences of SABP mentioned in the study.

8.4 Closing Remarks

Many countries from the OECD to South America, Africa, and developing regions have introduced various kinds of performance budgeting systems in order to increase the efficiency and effectiveness of budgetary programmes in the public sector, although these countries have different histories, cultures and political environments. These efforts have actually improved programmes' performance, governments' accountability and transparency, and the democracy of the policy process. However, performance budgeting in the public sector cannot be a panacea if policy makers do not consider seriously the potential dysfunctional consequences caused not only by the intrinsic characteristics of public programmes, such as multiple objectives and difficulties in measurement, but also by agents' deliberate intention of gaming, due to the information asymmetry between the principal and agents when it comes to designing and refining the system (Smith, 1995; Hood, 2006; Dixit, 1997; Thiel and Leeuw, 2002).

Against these backgrounds, this research provides a better and more complete understanding of the differences between spending ministries and the MOSF in the SABP system in Korea through both a quantitative and a qualitative approach; and it mainly reveals that there has indeed been both optimism bias by spending ministries in

their self-assessment stage and drastic cutting of spending ministries' generous self-assessment results by the MOSF, and that this is the result of unintended as well as intended dysfunctional consequences caused by spending ministries.

However, this research does suggest that a performance measurement regime which has produced some dysfunctional consequences has not necessarily failed to improve the performance of participating organisations (Smith, 1995; Kelman and Friedman, 2009; Pollitt, 1990). Rather, the research argues that the most important thing is not the existence of dysfunctional consequences themselves, but the careful consideration of them, as well as the pursuit of feasible policy alternatives for minimising or preventing them when it comes to implementing or revising the performance budgeting system. Kelman and Friedman (2009, p.942) interestingly comment that, "The appropriate comparison is between an organization's performance level with performance measurement and the dysfunctional responses, and the counterfactual performance level with no measurement, not between an organization's performance level assuming performance measurement with no dysfunctional responses and the level with some dysfunctions." Although the findings and discussions of the research cannot resolve all concerns over performance budgeting in the public sector, it is to be expected that they can give useful insights for understanding the process of performance budgeting and they can be effectively applied to other performance assessment systems, such as R&D programmes and public enterprise evaluations, in Korea, as well as to other countries' performance budgeting systems.

Appendix 1: Interviewees for the Case Studies

1-1. List of Interviewees in the Public Housing Area

1. Manager in the MLTM

(1) Main career path: Senior Deputy Director, Public Housing programme division

(2) Role in SABP: Overall handling of SABP results of the MLTM

(3) Time & Place: 10:30-11:10, 14th November 2011, Cafeteria at the Government Complex in Gwacheon, Korea; 5th November 2012 (written reply)

2. Manager in the MLTM

(1) Main career path: Senior Deputy Director, Budget division of the MLTM

(2) Role in SABP: Overall handling of SABP results of the MLTM

(3) Time & Place: 11:20-13:00, 31st October 2011, MLTM office in Pyungchon, Korea
17:00-17:30, 2nd November 2012 (telephone)

3. Senior manager in the MOSF

(1) Main career path: Director of Performance Management division of the MOSF

(2) Role in SABP: Policy making on SABP, coordinating welfare policy

(3) Time & Place: 14:20-15:10, 3rd November 2011, MOSF office in Gwacheon, Korea

4. Manager in the MLTM

(1) Main career path: Deputy Director of the Public Housing Fund division

(2) Role in SABP: Self-assessment on the Public Housing programme in the MLTM

(3) Time & Place: 11:00-11:40, 11th November 2011, Cafeteria at the Government Complex in Gwacheon, Korea

5. Manger in the MLTM

(1) Main career path: Deputy Director of Public Housing Fund division

(2) Role in SABP: Self-assessment on the Public Housing programme in the MLTM

(3) Time & Place: 13:30-14:10, 27th October 2011, Meeting room in the MLTM
8th November 2012 (written reply)

6. Manager in the MOSF

(1) Main career path: Deputy Director, Budgeting of Social Welfare field

(2) Role in SABP: Reviewing the MLTM's self-assessment results

(3) Time & Place: 11:00-12:00, 9th November 2011, Cafeteria at Gwacheon Complex
17:00-17:30, 6th November 2012 (telephone)

7. Manager in the MOSF

(1) Main career path: Deputy Director, Budgeting of Social Welfare and SOC

(2) Role in SABP: Reviewing the MLTM's self-assessment results

(3) Time & Place: 15:10-16:00, 11th November 2011, Meeting room in the MOSF
17:30-18:00, 25th October 2012 (telephone)

8. Manager in the MLTM

(1) Main career path: Deputy Director, Public Housing programme division

(2) Role in SABP: Self-assessment on the Public Housing programme in the MLTM

(3) Time & Place: 10:50-11:50, 1st November 2011, MLTM office

9. Manager in the MLTM

(1) Main career path: Director, Public Housing programme division

(2) Role in SABP: Coordinating self-assessment on the Public Housing programme

(3) Time & Place: 10:00-10:45, 1st November 2011, MLTM office
17:00-17:40, 5th November 2012 (telephone)

10. Senior manager in the MOSF

(1) Main career path: Director General of Fiscal Policy Bureau in the MOSF

(2) Role in SABP: Policy making on SABP

(3) Time & Place: 18:20-19:00, 1st November 2011, Cafeteria in Seoul City

11. Manager in the MOSF

(1) Main career path: Director of Performance Management division of the MOSF

(2) Role in SABP: Policy making on SABP, coordinating SABP results

(3) Time & Place: 17:00-17:30, 13th November 2012 (telephone)

12. Senior manager in the MOSF

(1) Main career path: Director General of Fiscal Policy Bureau in the MOSF

(2) Role in SABP: Policy making on SABP

(3) Time & Place: 5th November 2012 (telephone)

1-2. List of Interviewees in the Youth Employment Area

1. Assistant manager in the MOEL

(1) Main career path: Youth Employment programme implementation

(2) Role in SABP: Self-assessment of youth employment programme

(3) Time & Place: 16:30-17:10, 1st November 2011, Meeting room in the MOEL

2. Assistant manager in the MOEL

(1) Main career path: Budget division of the MOEL

(2) Role in SABP: Overall handling of SABP results of the MOEL

(3) Time & Place: 16:40-17:40, 31st October 2011, Meeting room in the MOEL;
29th October 2012 (written reply)

3. Manager in the MOSF

(1) Main career path: Deputy Director, Performance Management division in the MOSF

(2) Role in SABP: Reviewing the MOEL's self-assessment results

(3) Time & Place: 16:20-17:00, 28th October 2011, Meeting room in the MOSF;
17:40-18:15, 2nd November 2012 (telephone)

4. Manager in the MOSF

(1) Main career path: Deputy Director, Performance Management division in the MOSF

(2) Role in SABP: Reviewing the MOEL's self-assessment results

(3) Time & Place: 13:00-14:20, 2nd November 2011, Chinese restaurant, Gwangwhamun;
18:00-18:40, 3rd November 2012 (telephone)

5. Manger in the MKE

(1) Main career path: Global Trade Specialist Training programme implementation

(2) Role in SABP: Self-assessment on the Youth Employment programme in the MKE

(3) Time & Place: 15:10-15:50, 28th October 2011, Meeting room in the MKE;
17:00-17:30, 8th November 2012 (telephone)

6. Manager in the MKE

(1) Main career path: Global Trade Specialist Training programme implementation, Management of Public Enterprise in the MKE

(2) Role in SABP: Self-assessment on the Youth Employment programme in the MKE

(3) Time & Place: 11:00-12:00, 27th October 2011, Meeting room in the MKE;
16:00-16:40, 7th November 2012 (telephone)

7. Senior manager in the MKE

(1) Main career path: Director of Budget division of the MKE

(2) Role in SABP: Overall handling of the MKE's self-assessment results

(3) Time & Place: 18:00-18:30, 31st October 2011, Office in the MKE;
17:00-17:35, 12th November 2012 (telephone)

8. Manager in the MOSF

(1) Main career path: Budget Allocation in the Industrial field

(2) Role in SABP: Reviewing the self-assessment results of the MKE

(3) Time & Place: 17:15-18:00, 1st November 2011, Meeting room in the MOSF;
17:00-17:30, 30th October 2012 (telephone)

9. Assistant manager in the MEST

(1) Main career path: Junior College Student Training programme implementation

(2) Role in SABP: Self-assessment of Youth Employment programme

(3) Time & Place: 14:30-15:10, 2nd November 2011, Cafeteria in the MEST;
31st October 2012 (written reply)

10. Assistant manager in the MEST
(1) Main career path: Budget division of the MEST
(2) Role in SABP: Overall handling of the MEST's self-assessment results
(3) Time & Place: 15:20-16:20, 2nd November 2011, Meeting room in the MEST

11. Manager in the MOSF
(1) Main career path: Budget Allocation of the Education field
(2) Role in SABP: Reviewing of the MEST's self-assessment results
(3) Time & Place: 14:40-15:30, 1st November 2011, Meeting room in the MOSF;
17:45-18:20, 12th November 2012 (telephone)

12. Manager in the MOSF
(1) Main career path: Deputy Director, Performance Management division of the MOSF
(2) Role in SABP: Policy making on SABP, reviewing the MEST's self-assessment results
(3) Time & Place: 16:30-17:10, 27th October 2011, Cafeteria, Gwacheon Complex;
17:40-18:15, 30th October 2012 (telephone)

1-3. List of Interviewees in the Social Overhead Capital Case Study

1. Manager in the MLTM
(1) Main career path: Senior Deputy Director, Road Construction programme
(2) Role in SABP: Self-assessment of road construction programme
(3) Time & Place: 15:15-16:00, 31st October 2011, Meeting room in the MLTM;
17:00-17:30, 7th November 2012 (telephone)

2. Manager in the MLTM
(1) Main career path: Road Construction programme implementation
(2) Role in SABP: Self-assessment of road construction programme

(3) Time & Place: 16:05-16:35, 31st October 2011, Meeting room in the MLTM;
20th November 2012 (written reply)

3. Senior manager in the MOSF

(1) Main career path: General Director of Budget division of the MLTM

(2) Role in SABP: Overall coordinating of the MLTM's self-assessment results

(3) Time & Place: 17:15-18:10, 11th November 2011, Office in the MLTM

4. Assistant manager in the MOSF

(1) Main career path: Budgeting the SOC field (road, railway construction)

(2) Role in SABP: Reviewing the MLTM's self-assessment results

(3) Time & Place: 15:20-16:20, 27th October 2011, Meeting room in the MOSF
18:20-18:55, 25th October 2012 (telephone)

5. Manager in the MOSF

(1) Main career path: Deputy Director, Performance Management division in the MOSF

(2) Role in SABP: Reviewing the MLTM's self-assessment results

(3) Time & Place: 17:20-18:20, 28th October 2011, Meeting room in the MOSF;
17:00-17:30, 31st October 2012 (telephone)

6. Manager in the MOSF

(1) Main career path: Budgeting the SOC field (road construction)

(2) Role in SABP: Reviewing the MLTM's self-assessment results

(3) Time & Place: 18:20-19:10, 11th November 2011, Meeting room in the MOSF;
17:00-17:25, 20th November 2012 (telephone)

7. Manager in the MLTM

(1) Main career path: Railway Construction programme implementation

(2) Role in SABP: Self-assessment on the Railway Construction programme

(3) Time & Place: 16:15-17:00, 11th November 2011, Meeting room in the MLTM;
2nd November 2012 (written reply)

8. Manager in the MLTM

(1) Main career path: Railway Construction programme implementation

(2) Role in SABP: Self-assessment on the railway construction programme

(3) Time & Place: 12:20-13:00, 11th November 2011, meeting room in the MLTM;
16:00-16:35, 13th November 2012 (telephone)

9. Manager in the MOSF

(1) Main career path: Budgeting the SOC field (Railway Construction)

(2) Role in SABP: Reviewing the MLTM's self-assessment results

(3) Time & Place: 19:00-20:00, 27th October 2011, Meeting room in the MLTM

10. Manager in the MLTM

(1) Main career path: Budget division of the MLTM

(2) Role in SABP: Overall handling of the MLTM's self-assessment results

(3) Time & Place: 13:10-14:00, 31st October 2011, Office in the MLTM;
17:00-17:35, 9th November 2012 (telephone)

1-4. List of Interviewees from among Experts

1. Executive officer in the KIPF

(1) Role in SABP: SABP planning, SABP manual development, reviewing the spending ministries' self-assessment results

(2) Time & Place: 15:30-16:30, 4th November 2011, Office in the KIPF

2. Executive officer in the KDI

(1) Role in SABP: Reporting the SABP system, Consultant to the SABP

(2) Time & Place: 15:00-16:00, 8th November 2011, Office in the KDI;
17:00-17:30, 29th October 2012 (telephone)

3. Professor in the University

(1) Role in SABP: Overall handling of the SABP report in the National Assembly, reporting the SABP system, Consultant to the SABP

(2) Time & Place: 14:40-15:20, 9th November 2011, Office in the University;
17:40-18:20, 6th November 2012 (telephone)

4. Researcher in the KIPF

(1) Main career path: Budgeting of Social Welfare, SOC

(2) Role in SABP: Reviewing the MLTM's self-assessment results

(3) Time & Place: 11:00-12:00, 4th November 2011, Office in the KIPF;
17:00-17:30, 26th October 2012 (telephone)

5. Researcher in the KIPF

(1) Role in SABP: Reviewing spending ministries' self-assessment results, educating government officials in SABP

(2) Time & Place: 14:20-15:20, 4th November 2011, Office in the KIPF;
17:40-18:10, 26th October 2012 (telephone)

6. Deputy Director of the Board of Audit and Inspection of Korea

(1) Role in SABP: Reporting the SABP system, audit and inspection on the MOSF's Performance Management division

(2) Time & Place: 19:00-21:00, 2nd November 2011, Restaurant in Seoul City;
18:00-18:45, 20th November 2012 (telephone)

Appendix 2: Topic Guide for Government Officials of Spending Ministries

(First interview)

Dear Participant,

First of all, I would really appreciate your help, despite your busy schedule. I am a doctoral research candidate at the University of Birmingham in the UK. I am doing research on the features of the Korean performance budgeting system, Self-Assessment of Budgetary Programme (SABP). I am going to give you information and invite you to be part of this research.

The purpose of this research is to explore the differing perspectives on the performance of programmes between spending ministries and the Ministry of Strategy and Finance, which controls the SABP system, and to suggest policy alternatives for a better performance budgeting system in Korea.

You are being invited to take part in this research because you are/were one of the civil servants who are/were in charge of SABP, or you are an expert in performance budgeting or the SABP system.

The research will involve your individual participation in an interview that will take about an hour. If you are willing to take part in this research, I will arrange a convenient time and place to interview you through telephone or e-mail. Before the interview, you will have an opportunity to ask any questions about the research, and then I will ask you to sign a consent form. During the interview, if you do not wish to answer any of the questions, you may say so and I will move on to the next question. The entire interview will be tape-recorded because it will help to make sure that nothing is forgotten, but no-one will be identified by name on the tape and the tape will be destroyed after the research is finished. After the interview, you will have an opportunity to confirm the

accuracy of the analyses and findings of the research as it concerns your answers.

Your participation in this research is entirely voluntary. It is your choice whether to participate or not. Even if you do decide to take part in this research, you are still free to stop at any time. If you want to stop, you do not have to give any reason.

I declare that all your answers will be kept confidential and will be quoted anonymously. I will not be sharing information about you with anyone. The findings will be presented as part of my thesis, but no identifying information will ever be available to anyone.

If you wish to make any additions or amendments to your answers or the research, please do not hesitate to contact me. Thank you again for your time and consideration.

Sanghoon Shin, Ph.D Candidate
University of Birmingham, UK
Contact Telephone: 070-4045-3949
E-mail: shoon117@gmail.com

I have read the foregoing information, or it has been read to me. I have had the opportunity to ask questions about it and any questions I have asked have been answered to my satisfaction. I consent voluntarily to be a participant in this study.

Name of Participant _____

Signature of Participant _____

Date _____

Day/month/year

1. Would you give me a short professional resume and tell me about your job, including your role in SABP?

2. How do you think SABP has changed the performance and management of budgetary programmes? In particular, how has the linking of SABP results and the budget for a programme influenced the performance and management of programmes?

3. Does/did your ministry have any strategies for obtaining high SABP grades and scores? Is it possible to divide these strategies into formal and informal ones?

4. What do/did you think about the quantifiable and non-quantifiable performance of your programmes? Which type of performance do you think more important? Why? Do you think that the more important type of performance is/was well reflected in SABP grades and scores? If so, why? If not, why not?

5. To what extent do you agree with the statement that SABP has a tendency to focus on the quantifiable performance of your programme? Why? How could the SABP system be changed to reduce this?

6. In your division, which is/was regarded as more important – long-term strategy (e.g. ordering the priority of policy directions) or short-term achievement of the performance index? Why? What are/were the effects of this when you come/came to assess the performance of your programme in SABP? What could be done to increase consideration of the long-term or medium-term perspectives of programmes in SABP?

7. How much attention do/did you give to objectives which are not measured on the performance index? Do/did you make efforts not to focus only on the performance index? (e.g. increasing the number of items on the performance index, emphasizing the other outcomes of a programme) Which non-measurable objectives should be considered in SABP?

8. Do you think that in government as a whole the tendency to pursue narrow objectives which can easily be achieved, or the second best solution, rather than the overall objectives of the organisation, or the first best solution, has increased since the introduction of SABP? How about in your ministry? How about in your programme? What should be done to reduce this?

9. What are/were the main criteria for setting the performance targets of a programme? To what extent do/did you consider the next year's performance target level when deciding this year's targets? What would be/would have been the reaction of your division or ministry if you suggested a rather ambitious performance target? Do you

think your programmes' performance targets are/were highly ambitious? If so, why? If not, why not? Any solutions?

10. Have you ever found that your performance has been underestimated or overestimated? If you discovered some disadvantageous performance data when you were assessing the programme, what would you do/have done? Is/was it possible for you to present these as they were? If not, any solutions?

11. To what extent do/did you consider the opinions of related experts, interested parties, and ordinary citizens when you assess/assessed the performance of a programme? Do you think performance data fully reflect/reflected the need to satisfy citizens or the genuine outcome of a programme? If not, any solutions?

12. What do you think the most important criteria are when it comes to selecting high priority groups in relation to SABP? (e.g. trainees, in the case of an employment training programme) To what extent do/did you consider the attainability of performance targets when you choose/chose policy objects? Why? Any solutions?

13. Have there been/were there any differences in SABP results between your ministry and the MOSF? How do/did you cope with cases where there are/were differences in SABP scores and grades between your ministry and the MOSF? Why do you think these differences in SABP score and grade between your ministry and the MOSF happen/happened? Any solutions?

14. How do you think the unique features of Korean government culture have impacted on SABP? Do you think these cultural features could lead to the differences between spending ministries and the MOSF? If so, why? If not, why not? Any solutions?
(e.g. general one-year rotation of personnel system, ministers' short tenure of office, power relations between spending ministries and the MOSF, priority given to the presidential agenda etc.)

15. What do you think about the budget division's recommendations on your self-assessment results?
* In the case of budget division's officials: What do you think about the programme divisions' self-assessment results?

16. Do you have any further comments on disagreements between your ministry and the MOSF? What do you think should be done to fully achieve the objectives of SABP?

Appendix 3: Topic Guide for Government Officials of the MOSF and Experts on Performance Budgeting (First interview)

1. Would you give me a short professional resume and tell me about your job, including your role in SABP?

2. How do you think SABP has changed the performance and management of budgetary programmes? In particular, how has the linking of SABP results and budgets influenced the performance and management of programmes?

3. Does/did your ministry have any strategies for reviewing spending ministries' self-assessment results? Is it possible to divide these strategies into formal and informal ones?
* For experts: Do you think either spending ministries or the MOSF have any particular strategies for SABP? Is it possible to divide these strategies into formal and informal ones?

4. What do/did you think spending ministries regard/regarded as more important: the quantifiable performance of programmes, or the non-quantifiable performance? Which kind of performance do you think more important? Why? Do you think that the more important kind of performance was well reflected in SABP grades and scores? If so, why? If not, why not?

5. To what extent do you agree with the statement that SABP has a tendency to focus on the quantifiable performance of a programme? Why? How could the SABP system be changed to reduce this?

6. What do you think spending ministries regard as more important - long-term strategy (e.g. ordering the priority of policy directions) or short-term achievement of the performance index? Why? What could be done to increase consideration of the long-term or medium-term perspectives of programmes in SABP?

7. How much attention do you think spending ministries give to non-measurable objectives and how much to objectives that can be measured on a performance index? Why? What non-measurable objectives should be included in SABP?

8. Do you think spending ministries have/had a tendency to pursue narrow objectives which can/could easily be achieved, or the second best solution, rather than the overall objectives of the organisation, or the first best solution? What should be done to reduce this?

9. What do you think spending ministries consider/considered most important when deciding the performance targets of a programme? Do you think spending ministries' performance targets are/were highly ambitious? If so, why? If not, why not? Any solutions?

10. Do you think there is/was any distortion, underestimation or overestimation in the self-assessed results of spending ministries? Or do you think spending ministries are/were faithfully presenting their self-assessment results, including disadvantageous performance data? If not, any solutions?

11. Do you think spending ministries consider/considered fully the opinions of related experts, interested parties, and ordinary citizens when they assess/assessed the performance of a programme? In particular, do you think spending ministries' performance data fully reflect/reflected the need to satisfy citizens or the genuine outcome of a programme? If not, any solutions?

12. What do you think spending ministries consider/considered the most important criteria when it comes/came to selecting high priority groups? (e.g. trainees in the case of an employment training programme) To what extent do/did spending ministries consider the attainability of performance targets when they choose/chose high priority groups? Why? Any solutions?

13. Are there any differences in SABP results between spending ministries and the MOSF? How do/did you cope with cases where there are/were differences in SABP scores and grades between spending ministries and the MOSF? Why do you think these differences in SABP scores and grades between spending ministries and the MOSF happen/happened? Any solutions?
* For experts: What do you think about the differences in SABP results between spending ministries and the MOSF? Why do you think these differences in SABP scores and grades between spending ministries and the MOSF happen? Any solutions?

14. How do you think the unique features of Korean government culture have impacted on SABP? Do you think these cultural features could lead to the differences between spending ministries and the MOSF? If so, why? If not, why not? Any solutions?
(e.g. general one-year rotation of personnel system, ministers' short tenure of office, power relations between spending ministry and the MOSF etc.)

15. Do you think the budget divisions of spending ministries have a different perspective on SABP from that of programme divisions? If so, why? If not, why not?

16. Do you have any further comments on disagreements between spending ministries and the MOSF? What do you think should be done to fully achieve the objectives of SABP?

Appendix 4: Topic Guide for Government Officials and Experts

(Supplementary interview)

Dear Participant,

First of all, I'd like to say again that I really appreciated your help last year. Having studied the results of the first interviews, I have realised that there are widely differing perspectives on the performance of a programme between spending ministries and the Ministry of Strategy and Finance (MOSF).

I should therefore like to ask you some further questions. In this interview, I want to ask your opinions on what might lead to these differing perspectives between ministries, and how they can be reduced. Your valuable answers will enable me to put forward more realistic and feasible policy alternatives for a better SABP system.

I once more declare that all your answers will be kept confidential and be quoted anonymously. I will not be sharing information about you with anyone. The findings will be presented as a part of my thesis, and no identifying information will ever be available to anyone.

If you wish to make any additions or amendments to your answers or the research, please do not hesitate to contact me. Thank you again for your time and consideration.

Sanghoon Shin, Ph.D Candidate
University of Birmingham, UK
Contact Telephone: 070-4045-3949
E-mail: shoon117@gmail.com

1. The results of the first interviews suggested that spending ministries have a tendency to focus on the quantifiable, short-term, measurable and narrow performance of a programme rather than on the non-quantifiable, long-term, non-measurable and whole performance, such as the government's or the organisation's objectives. In addition, spending ministries have also tried to set performance target levels as low as possible, considering their attainability and the next year's target level. Furthermore, sometimes they have hidden disadvantageous performance data and even manipulated these, and they have tended to select favourable priority groups. The case study indicated that these tendencies might lead both to spending ministries' optimism bias in terms of self-assessment results and to the MOSF's drastic cutting of spending ministries' self-assessment results. What do you think about this?

2. The analysis of the first interviews indicated that both spending ministries and the MOSF have tried to reduce or prevent the phenomena mentioned above. For example, the MOSF has revised the SABP manual every year to reflect spending ministries' opinions, and has tried to give more incentives to effective programmes as well as penalties to poor programmes. In the following section, the researcher will ask your opinions on these policy alternatives and on other policy alternatives suggested through the first interviews.

2-1. According to the first interviews, many people were concerned that SABP did not fully take account of the outcomes of programmes, such as (a) qualitative performance (b) long-term performance (c) difficult-to-measure performance and (d) overall achievement of programmes.

(1) Do you agree with this statement? If so, why? If not, why not?

(2) Do you believe that SABP should have more questions about (a), (b), (c) or (d)?

(3) Do you believe that spending ministries should submit evidence that they have pursued (a), (b), (c) or (d) for the MOSF to assess spending ministries' efforts properly?

(4) Overall, how important do you think changes like those mentioned in (2) and (3) are? (Please select one of the following: very important, important, modest, not important, not very important.)

2-2. Many interviewees indicated that SABP needed to change its operational direction, for example its "penalty-oriented operation" or the "strict relationship" it enforces between SABP results and budgets, especially budgets for poor grade programmes.

(1) Do you agree with this statement? If so, why? If not, why not?

(2) Do you think that increasing incentives to effective programmes is more important than giving penalties to poor programme?

(3) Do you think such a change might lead to increasing the genuine performance of a programme?

(4) Overall, how important do you think it is to make changes like this? (Please select one of the following: very important, important, modest, not important, not very important)

2-3. From the year 2011 the MOSF introduced what is called a relative evaluation system at some spending ministries' self-assessment stage, making it mandatory for effective programmes to make up less than 20% of a ministry's programmes and poor programmes to make up more than 10%. This was done to reduce spending ministries' optimism bias.

(1) Do you favour making this compulsory policy in principle?

(2) Do you believe the proportions of 20% and 10% are appropriate?

(3) Overall, how important do you think it is to make this change? (Please select one of the following: very important, important, modest, not important, not very important)

2-4. Some interviewees also suggested that the assessment level might be elevated from "sub-programme" to "programme" level to consider the long-term and whole (overall) performance, and at the same time constant monitoring of "programme" level should be implemented.

* The programme budgeting structure of Korea: Organisation → Programmes → Sub-programmes (present assessment level) → Activities → Cost elements

(1) Do you agree with this statement? If so, why? If not, why not?

(2) Overall, how important do you think it is to make this change? (Please select one of the following: very important, important, modest, not important, not very important)

2-5. The necessity of giving more discretion to spending ministries in terms of self-assessment methods, selecting performance indicators and setting performance target levels was mentioned by many people as a means of increasing the trust between spending ministries and the MOSF and implementing SABP as a real self-assessment system by spending ministries.

(1) Do you agree with this statement? If so, why? If not, why not?

(2) To what extent would this policy decrease the MOSF's ability to compare performance between programmes?

(3) Overall, how important do you think it is to make this change? (Please select one of the following: very important, important, modest, not important, not very important)

2-6. Some interviewees argued that spending ministries' self-assessment stage should be abolished and spending ministries' role should be limited to submitting performance data for the MOSF's performance evaluation of a programme, considering the MOSF's continuous drastic cutting of spending ministries self-assessment score.

(1) Do you agree with this statement? If so, why? If not, why not?

(2) How is it possible to trust in the MOSF's performance evaluation without knowing spending ministries' self-assessment results?

(3) Overall, how important do you think it is to make this change? (Please select one of the following: very important, important, modest, not important, not very important)

3. Besides the policy alternatives mentioned above, if you have any other ideas for improving the SABP system, please give details.

BIBLIOGRAPHY

Ammons, D. N. (1995a) **Accountability for Performance: Measurement and Monitoring in Local Government.** Washington D.C.: The International City/County Management Association.

Ammons. D. N. (1995b) Overcoming the Inadequacies of Performance Measurement in Local Government: The Case of Libraries and Leisure Services. **Public Administration Review**, 55(1): 37-47

Andrews, M. and Hill, H. (2003) The Impact of Traditional Budgeting Systems on the Effectiveness of Performance-Based Budgeting: A Different Viewpoint on Recent Findings. **International Journal of Public Administration**, 26(2): 135-155

Archer, S. (1988) "Qualitative research and the epistemological problems of the management discipline." In Pettigrew, A. M. (ed) **Competitiveness and the Management Process**. Oxford: Basil Blackwell Inc. pp. 265-302

Arksey, H. (2004) "Semi-structured and Unstructured Interviewing" In Becker, S. and Bryman, A. (eds) **Understanding Research For Social Policy and Practice: Themes, Methods and Approaches**. Bristol: The Policy Press.

Asiaeconomy (2009) **Spending ministries overestimated their SABP score, questioning the effectiveness** [online]. Available from: http://stoo.asiae.co.kr/news/stview.htm?idxno=2009121310493055040 [Accessed 4th May 2012]

BAI (2011a) **The result of auditing: Inspection of the MLTM's performance report in the fiscal year 2010**. Seoul: Board of Audit and Inspection of Korea.

BAI (2011b) **BAI news release: Ascertainment of government's settlement of accounts in the fiscal year 2010 and Report of Audit results** [online]. Available from: www.bai.go.kr [Accessed on 20th April 2012]

Baker, G. P. (1992) Incentive contracts and performance measurement. **Journal of Political Economy**, 100: 598-614

Bang, M. (2009) **Analysis of the impact of results from the self-assessment of budgetary programmes in budget decisions**. Doctoral dissertation, Sungkunkwan University, Seoul: Korea.

Bang, M. and Yoon, K. (2006) Analysis of the SABP results in 2006. **Korean Society and Public Administration**, 2006-5, 409-424

Bates, R. H. (1995) "Social Dilemmas and Rational Individuals" In Harriss, J. et al. (eds) **The New Institutional Economics and Third World Development**. London: Routledge. pp. 27-48

Bazeley, P. (2007) **Qualitative Data Analysis with Nvivo**. London: Sage.

Becker, S. and Bryman, A. (2004) "The Nature of Quantitative and Qualitative Research" In Becker, S. and Bryman, A. (eds) **Understanding Research For Social Policy and Practice: Themes, Methods and Approaches**. Bristol: The Policy Press.

Behn, R. D. (2003) Why Measure performance? Different purposes require different measures. **Public Administration Review,** 63(5): 586-606

Bell, J. (2006) **Doing your research project: A guide for first-time researchers in Education, Health and Social Science**. 4th ed. Berkshire: McGraw-Hill Education.

Belli, G. (2008) **Nonexperimental Quantitative Research** [online]. Available from: http://media.wiley.com/product_data/excerpt/95/04701810/0470181095-1.pdf [Accessed 24 March 2011]

Bevan, G. and Hood, C. (2006) What's measured is what matters: targets and gaming in the English public health care system. **Public Administration**, 84(3): 517-538

Board of Audit and Inspection of Korea (2011) **Comprehensive policy inspection: Concerns over the supply and demand, and price of public housing construction**. Seoul: BAI.

Boland, T. and Fowler, A. (2000) A systems perspective of performance management in public sector organisations. **International Journal of Public Sector Management**, 13(5): 417-446

Bouckaert, G. and Balk, W. (1991) Public Productivity Measurement: Diseases and Cures. **Public Productivity & Management Review**, 15(2): 229-235

Bouckaert, G. and Halligan, J. (2008) **Managing performance: international**

comparisons. Oxon: Routledge.

Bouckaert, G. and Peters, G. (2002) Performance Measurement and Management: The Achilles's Heel in Administrative Modernisation. **Public Performance & Management Review**, 25(4): 359-362

Bovaird, T. (1996) "The Political Economy of Performance Measurement." In Halachmi, A. and Bouckeart, G. (eds) **Organizational Performance and Measurement in the Public Sector: Towards Service, Effort and Accomplished Reporting**. Westport, CT: Quorum Books. pp. 145-166

Bovaird, T. and Loffler, E. (2003) **Public management and governance**. London: Routledge.

Breul, J. D. (2007) Three Bush Administration Management Reform Initiatives: The President's Management Agenda, Freedom to Manage Legislative Proposals, and the Program Assessment Rating Tool. **Public Administration Review**, 67(1): 21-26

Bryman, A. (2008) **Social Research Methods**. 3rd ed. Oxford: Oxford University Press.

Bryman, A. (2006) Integrating quantitative and qualitative research: how is it done? **Qualitative Research**, 6(1): 97-113

Bryman, A. (1988) **Quantity and Quality in Social Research**. London: Routledge.

Buchanan, D. R. (1992) An Uneasy Alliance: Combining Qualitative and Quantitative Research. **Health Education Quarterly**, 19(1): 117-135

Campbell, D. T. and Stanley, J. C. (1966) **Experimental and Quasi-Experimental Designs for Research**. Boston: Houghton Mifflin Company.

Carter, N., Klein, R. and Day, P. (1995) **How Organisations Measure Success: The Use of Performance Indicators in Government**. London: Routledge.

Chang, M. and Yoon, S. (2002) Empirical Analysis of Determining Factors of Budget Deliberation of the National Assembly. **Korean Policy Studies Review**, 11(2): 100-119

Cherryholmes, C. H. (1992, August-September) Notes on pragmatism and scientific realism. **Educational Researcher**, 13-17

Cho, I. (2010) **The impact of Korean performance budgeting on budgetary programmes**. Doctoral dissertation, University of Birmingham, Birmingham: UK.

Choi, S. and Park, M. (2008) A Study on the Ratchet Effect of Performance Management in the Public Sector. **Korean Journal of Public Administration**, 46(4): 135-161

Cook, T. D. and Campbell, D. T. (1979) **Quasi-Experimentation**. Boston: Houghton Mifflin.

Courty, P. and Marschke, G. (2007) Making Government Accountable: Lessons from a Federal Job Training Program. **Public Administration Review**, 67(5): 904-916

Courty, P. and Marschke, G. (2008) An Empirical Investigation of Gaming Response to Explicit Performance Incentives. **Journal of Labor Economics**, 22(1): 23-56

Creswell, J. W. and Plano Clark, V. L. (2011) **Designing and Conducting Mixed Methods Research**. 2nd ed. Thousand Oaks, CA: Sage.

Creswell, J. W. (2009) **Research design: Qualitative, Quantitative, and Mixed methods approaches**. 3rd ed. Thousand Oaks, CA: Sage.

Creswell, J. W. (2007) **Qualitative inquiry and research design: Choosing among five approaches.** 2nd ed. Thousand Oaks, CA: Sage.

Crotty, M. (1998) **The foundations of social research: Meaning and perspective in the research process**. London: Sage.

De Bruijn, H. (2002). Performance Measurement in the Public Sector: Strategies to Cope With the Risks of Performance Measurement. **The International Journal of Public Sector Management**, 15(7): 578-594

De Vaus, D. (2001) **Research Design in Social Research**. London: Sage.

Denscombe, M. (2003) **Good Research Guide: For Small-Scale Research Projects**. Berkshire: McGraw-Hill.

Diamond, J. (2005) Establishing a Performance Management Framework for Government. **IMF Working Paper.** IMF.

Dixit, A. (1997) Power of Incentives in Private versus Public Organizations. **American Economic Review**, 87(2): 378-382

DOE (1995) **How to Measure Performance: A Handbook of Techniques and Tools**. Washington D.C.: DOE

Dooley, D. (2001) **Social Research Methods**. 4th ed. NJ: Prentice-Hall.

Downs, G. W. and Larkey, P. D. (1986) **The Search For Government Efficiency: From Hubris to Helplessness** (1st ed.). NY: Random House.

Dull, M. (2006) Why PART? The Institutional Politics of Presidential Budget Reform. **Journal of Public Administration Research and Theory**, 16(2): 187-215

Eggertsson, T. (1990) **Economic Behavior and Institutions: Principles of Neoinstitutional Economics**. Cambridge: Cambridge University Press.

Epstein, P. D. (1992) Get ready: The Time for Performance Measurement is finally coming. **Public Administration Review**, 52(5): 513-519

Ferris, J. M. and Graddy, E. A. (1998) A Contractual Framework for New Public Management Theory. **International Public Management Journal**, 1(2): 225-240

Fox, M. (2007) "Australia" In Teresa, C. (ed.) **Performance Budgeting in OECD Countries**. Paris: OECD. pp. 89-102

Flynn, N. (1997) **Public Sector Management**. 3rd ed. London: Prentice-Hall/Harvester Wheatsheaf.

GAO (1996) **Executive Guide Effectively Implementing the Government Performance Results Act**. GAO/GGD-96-118. Washington D.C.: GAO.

GAO (1997a) **The Government Performance and Results Act: 1997 Government Implementation will be uneven**. GAO/GGD-97-109. Washington D.C.: GAO.

GAO (1997b) **Performance Budgeting: Past Initiatives Offer Insights for GPRA Implementation**. GAO/AIMD/GGD-99-67. Washington D.C.: GAO.

GAO (1999) **Performance Budgeting: Initial Experiences Under the Results Act in Linking Plans with Budgets**. GAO/AIMD-99-67. Washington D.C.: GAO.

GAO (2001) **Managing for Results: Federal managers' views on key management issues vary widely across agencies**, GAO-01-592

GAO (2003) **High-Risk Series: Strategic Human Capital Management**. GAO-03-120. Washington D.C.: GAO.

GAO (2004a) **Performance-Based Initiatives Are Evolving**. GAO-04-422. Washington D.C.: GAO.

GAO (2004b) **Results-Oriented Government: GPRA Has Established A Solid Foundation For Achieving Greater Results**. GAO-04-38. Washington D.C.: GAO.

GAO (2011) **GPRA Modernization Act Provides Opportunities to Help Address Fiscal, Performance, and Management Challenges**. GAO-11-466T. Washington D.C.: GAO.

Gillham, B. (2005) **Research Interviewing: The Range of Techniques**. Berkshire: McGraw Hill Education.

Gilmour, J. B. and Lewis D. E. (2005) Assessing performance budgeting at OMB: The influence of politics, performance, and programme size. **Journal of Public Administration Research and Theory**, 16: 169-186

Gilmour, J. B. and Lewis, D. E. (2006a) Does performance budgeting work? An examination of the office of management and budget's PART scores. **Public Administration Review**, 66(5): 742-752

Gilmour, J. B. and Lewis, D. E. (2006b) Political Appointees and the Competence of Federal Program Management. **American Politics Research**, 34(1): 22-50

Gilmour, J. B. (2007) Implementing OMB's Program Assessment Rating Tool (PART): Meeting the Challenges of Integrating Budget and Performance. **OECD Journal on Budgeting**, 7(1): 1-40

Goddard, M., Mannion, R. and Smith, P. (2000) Enhancing performance in health care: a theoretical perspective on agency and the role of information. **Economics of Health Care Systems**, 9(2): 95-107

Greene, J. C., Caracelli, V. J. and Graham, W. F. (1989) Toward a conceptual framework for mixed-method evaluation designs. **Educational Evaluation and Policy Analysis**, 11(3): 255-274

Grix, J. (2010) **The Foundations of Research**. 2nd ed. Hampshire: Palgrave Macmillan.

Ha, Y. (2002) The new trend of Institutional Theory: Theoretical self-innovation and convergence. **Korean Review of Public Administration,** 36(4): 339-359

Hall, P. A. and Taylor, Rosemarym C. R. (1996) Political Science and the Three New Institutionalisms. **Political Studies**, 44: 936-957

Harriss, J., Hunter, J. and Colin, M. L. (1995) **Introduction: Development and Significance of NIE**. London: Routledge. pp. 1-13

Heinrich, C. J. (2007) False or fitting recognition?: The use of high performance bonuses in motivating organizational achievements. **Journal of Policy Analysis and Management**, 26(2): 281-304

Heinrich, C. J. and Marschke, G. (2010) Incentives and Their Dynamics in Public Sector Performance Management Systems. **Journal of Policy Analysis and Management**, 29(1): 183-208

Holmstrom, B. and Milgrom, P. (1991) Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. **Journal of Law, Economics and Organization**, 7: 24-52

Hood, C. (1991) A Public Management for all Seasons? **Public Administration**, 69(Spring): 3-19

Hood, C. (2006) Gaming in Targetworld: the targets approach to managing British public services. **Public Administration Review**, July/August: 515-521

Hood, C. (2007) Public Service Management by Numbers: Why Does it Vary? Where Has it Come From? What Are the Gaps and the Puzzles? **Public Money and Management**, 27(2): 95-102

Hyde, A. C. (1992) **Government Budgeting: Theory·Process·Politics**. 2nd ed. California: Brooks/Cole.

James, O. (2001) New public management in the UK: Enduring legacy or fatal remedy? **International Review of Public Administration**, 6(2): 15-26

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

Johnsen, A. (2005) What Does 25 Years of Experience Tell Us About the State of Performace Measurement in Public Policy and Management? **Public Money & Management**, 25(1): 9-17

Johnson, R. B. and Christensen, L. (2008) **Educational Research: Quantitative, Qualitative, and Mixed Approaches**. 3rd ed. Thousand Oaks, CA: Sage.

Joongangilbo (2009) **Spending ministries, Inflated the SABP score** [online]. Available from:http://article.joinsmsn.com/news/article/article.asp?ctg=10&Total_ID=3919961[Accessed 4th May 2012]

Jordan, M. M. and Hackbart, M. M. (1999) Performance budgeting and performance funding in the States: a status assessment. **Public Budgeting & Finance**, 19(1): 68-88

Joyce, P. G. and Sieg, S. (2000) Using performance information for budgeting: Clarifying the framework and investigation recent state experience. **2000 Symposium of the center for accountability and performance of the American Society for public administration**. Washington D.C., 11-12 February 2000. Washington D.C.: American Society for Public Administration.

Joyce, P. G. (2011) The Obama Administration and PBB: Building on the Legacy of Federal Performance-Informed Budgeting? **Public Administration Review**, 71(3): 356-367

Jung, M. (2012) **The impact of Financial Performance Information on Budget Decision: The Case of Self-Assessment of Budgetary Program in Korea**. Doctoral dissertation, Korea University, Seoul: Korea.

Kang, S. (2000) **Financial Administration**. Seoul: Parkyoungsa.

Kang, Y. (2007) An Evaluation of the Performance Budgeting System and the Factors Affecting its Results. **Korean Governance Studies Review**, 14(2): 291-323

KDI (2007) **A Guideline for the In-depth Evaluation System**. Seoul: KDI.

KDI (2008) **Programme budgeting structure**. Seoul: KDI.

Kelman, S. and Friedman, J. N. (2009) Performance Improvement and Performance Dysfunction: An Empirical Examination of Distortionary Impacts of the Emergency Room Wait-Time Target in the English National Health Service. **Journal of Public Administration Research and Theory**, 19: 917-946

Keum, J. and Lee, S. (2009) An Empirical Study on Dysfunctional Consequences of the Performance Management System: Focus on Perceptions of the Central Government Officials. **Korean Society and Public Administration**, 20(2): 347-377

King, G., Keohane, O. and Verba, S. (2001) **Designing Social Inquiry: Scientific Inference in Qualitative Research**. Princeton, NJ: Princeton Universtiy Press.

Kim, N. (1997) Institutions Do Matter: From the Perspective of New Institutional Economics. **Korean Policy Studies Review**, 6(1): 127-151

Kim, T. (2009) **Spending Ministries, Inflating their Self-Assessment Results** [online]. Joong-Ang Daily Newspaper. Available from: http://article.joinsmsn.com/news/article/article.asp?ctg=10&Total_ID=3919961[Accessed 18 October 2011]

KIPF (2012) **Guidance for Developing Performance Indicator of Budgetary Program.** Seoul: Korea Institute of Public Finance.

KIPF (2011) **Recent trends in the performance management systems of major countries**. Seoul: Korea Institute of Public Finance.

KIPF (2008) **International Comparison of Performance Budgeting**. Seoul: KIPF.

KIPF (2006) **Guidance for Developing Performance Indicator in Budgetary Program.** Seoul: Korea Institute of Public Finance.

Ko, Y. et al. (2004) **Performance Management in the Public Sector**. Seoul: KDI.

Kong, B. (2008) Performance Management & Government Evaluation under the Integrated Performance Management System. **Korean Policy Studies Review**, 17(1): 55-79

Kong, D. et al. (2009) A Study on the Problems and Polices of Performance Measurement and Behaviours in the Individual Performance Measurement. **Korean Public Administration Journal**, 18(2): 59-79

Kong, D. et al. (2007) What determines Program Performance Ratings in Korea. **Korean journal of policy analysis and evaluation**, 17(4): 27-56

Kwok, C., Au, W. T. and Ho, J. M. C. (2005) Normative controls and self-reported counterproductive behaviors in the workplace in China. **Applied Psychology: An International Review**, 54(4): 456-475

Kwon, S. and Kim, N. (1995) The Economic of Organisation for the Field of Public Administration: Focus on Principal-Agent Theory. **Korean Public Administration Review**, 29(1): 77-95

Kuhn, T. S. (1970) **The structure of scientific revolutions**. 2nd ed. Chicago: University of Chicago Press.

Lakatos, I. (1974) "Falsification and the Methodology of Scientific Research Programmes". In Lakatos, I. and Musgrave, A. (eds) **Criticism and the Growth of Knowledge**. Cambridge: Cambridge University Press, pp. 91-196

Lane, J. E. (1993) **The Public Sector: Concepts, Models, and Approaches**. Newbury Park: Sage.

Lauth, T. P. (1985) Performance evaluation in the Georgia budgeting process. **Public Budgeting and Finance**, 5(1): 67-85

Lebas, M. J. (1995) Performance measurement and performance management. **International Journal of Production Economics**, 41(1): 23-35

Lee, H. (1998) The Implication of New Institutional Economics for the Field of Public Administration. **The Annals of Public Administration Research**: 257-272

Lee, H. (2009) "Unexpected Consequences as an Object of Policy Study." In **Conference on Seoul Public Administration Society. Seoul, Winter 2009**. Seoul: Seoul Public Administration Society. pp. 229-252

Lee, J. (1993) **Microeconomic theory**. 2nd ed. Seoul: Bup Moon Sa.

Lee, M. (2006) **Institutionalizing Congress and the Presidency: The U.S. Bureau of Efficiency, 1916-1933**. College Station, TX, Texas A & M University Press.

Lee, N. (2003) **A Study for Activating Performance Budgeting System**. Seoul: KIPA.

Lee, Y. (2007) A Study on Improved Approaches to Evaluation for Performance Management in the Korean Government. **Korean Journal of Policy Analysis and Evaluation**, 17(3): 1-30

LeGrand, J. and Bartlett, W. (Eds.) (1993) **Quasi-markets and social policy**. London: Macmillan.

Lincoln, Y. S. and Guba, E. G. (2000) "Paradigmatic controversies, contradictions, and emerging confluences". In Lincoln, Y. S. and Guba, E. G. (eds) **Handbook of qualitative research**. Thousand Oaks, CA: Sage. pp. 163-188

Lowndes, V. (1996) Varieties of New Institutionalism: A Critical Appraisal. **Public Administration**, 74(2): 181-198

Mackay, K. (2007) **How to build M&E system to support better government**. Washington DC: World Bank.

Magee, Bryan (1985) **Philosophy and the Real World: An Introduction to Karl Popper**. Open Court.

McCormack, L. (2007) "Canada." In Teresa, C. (ed.) **Performance Budgeting in OECD Countries**. Paris: OECD. pp.103-120

McGill, R. (2001) Performance budgeting. **International Journal of Public Sector Management**, 14(5): 376-390

Melkers, J. and Willoughby, K. (2001) Budgeters' views of state performance-budgeting systems: distinctions across branches. **Public Administration Review**, 61(1): 54-64

Meyer, M. W. and Gupta, V. (1994) The performance paradox. **Research in Organizational Behavior**, 16: 309-369

Miles, M. and Huberman, A. (1994) **Qualitative Data Analysis: An Expanded Sourcebook**. London: Sage.

Mill, J. S. (1834) **System of Logic Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation**. 8th ed. London: Longmans, Green and Company.

Miller, G. J. et al (2001) **Performance-Based Budgeting**. Oxford: Westview Press.

Ministry of Land, Transport, and Maritime Affairs (2011) **Stabilisation of the common people's housing through housing trade normalisation**. Seoul: MLTM.

Ministry of Land, Transport, and Maritime Affairs (2010) **Housing transaction vitalisation**. Seoul: MLTM.

Ministry of Planning and Budget (2003) **Performance Management Handbook**. Seoul: MPB.

Ministry of Strategy and Finance (2012a) **Results of the SABP from 2005 to 2010**. Seoul: MOSF.

Ministry of Strategy and Finance (2012b) Guidance for public enterprise assessment in 2013 [online]. Available from: http://www.mosf.go.kr/news/news02.jsp?boardType=general&hdnBulletRunno=62&cvbnPath=&sub_category [Accessed 11th February 2013]

Ministry of Strategy and Finance (2011a) MOSF·OECD, Opening performance management international forum [online]. Available from: http://www.mosf.go.kr/news/news01.jsp?boardType=general&hdnBulletRunno=60&cvbnPath=&sub_category [Accessed 6th November 2012]

Ministry of Strategy and Finance (2011b) **Policy alternatives for improving the SABP system** [online]. Available from: http://www.mosf.go.kr/news/news02.jsp?cvbnPath=yes&boardType=general&hdnBulletRunno=62&hdnDiv=&hdnRecent=20&hdnCommand=new&cat=1 [Accessed 2nd April 2012]

Ministry of Strategy and Finance (2011c) **Guidance for Drawing Up Performance Planning Reports in the fiscal year 2012.** Seoul: Korea.

Ministry of Strategy and Finance (2010a) **A Guidance for Self-Assessment of Budgetary Programme in the fiscal year 2009**. Seoul: MOSF.

Ministry of Strategy and Finance (2010b) **Budgeting Guidelines**. Seoul: MOSF.

Ministry of Strategy and Finance (2010c) **SABP report: Ministry of Knowledge Economy in the fiscal year 2009.** Seoul: Korea.

Ministry of Strategy and Finance (2010d) **SABP report: Ministry of Land Transport and Maritime Affairs in the fiscal year 2009.** Seoul: Korea.

Ministry of Strategy and Finance (2010e) **SABP report: Ministry of Labor in the fiscal year 2009**. Seoul: Korea.

Ministry of Strategy and Finance (2010f) **SABP report: Ministry of Education, Science and Technology in the fiscal year 2009**. Seoul: Korea.

Ministry of Strategy and Finance (2010g) **Guidance for Self-Assessment of Budgetary Programmes in 2011**. Seoul: Korea.

Ministry of Strategy and Finance (2009a) **Guidance for Drawing Up Performance Planning Reports in the fiscal year 2010.** Seoul: Korea.

Ministry of Strategy and Finance (2009b) **Guidance for Self-Assessment of Budgetary Programmes in the fiscal year 2008**. Seoul: Korea.

Ministry of Strategy and Finance (2008a) **Government fund evaluation in the fiscal year 2007.** Seoul: Korea.

Ministry of Strategy and Finance (2008b) **Guidance for Self-Assessment of Budgetary Programmes in the fiscal year 2007**. Seoul: Korea.

Ministry of Strategy and Finance (2007) **Guidance for Self-Assessment of Budgetary Programmes in the fiscal year 2006**. Seoul: Korea.

Moore, M. (2002) **Creating Public Value: Strategic Management in Government**. Boston: Harvard University Press.

Morgan, D. (2007) Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. **Journal of Mixed Methods Research**, 1(1): 48-76

MOSF and KIPF (2012) **Manual for developing performance index in the budgetary programmes: Performance index development on-site**. Seoul: Korea.

Mumford, M. D. (1983) Social comparison theory and the evaluation of peer evaluations: A review and some applied implications. **Personnel Psychology**, 36: 867-881

Nabli, M. K. and Nugent, J. B. (1989) The New Institutional Economics and Its Applicability to Development. **World Development**, 17(9): 1333-1347

Namkoong, K. (2010) **Research Methods for Public Administration**. Seoul Korea: Bup Moon Sa.

National Assembly Budget Office (2011) **Evaluation of Performance Plans in 2012**. Seoul: NABO.

National Assembly Budget Office (2010a) **Evaluation of the Korean Government's Youth Employment Programmes**. Seoul: NABO.

National Assembly Budget Office (2010b) **Evaluation of the Korean Government Performance Report**. Seoul: NABO.

National Assembly Budget Office (2010c) **Evaluation of Performance Report in the fiscal year 2009**. Seoul: NABO.

National Assembly Budget Office (2010d) **Evaluation of Performance Plans in 2011.** Seoul: NABO.

National Assembly Budget Office (2009a) **Evaluation of Performance Plans in 2010**. Seoul: NABO.

National Assembly Budget Office (2009b) **Evaluation of Global Internship Programmes**. Seoul: NABO.

National Assembly Budget Office (2008) **Evaluation of Performance Plans in 2008**. Seoul: NABO.

National Assembly Budget Office (2007a) **Evaluation of Performance Plans in 2007**. Seoul: NABO.

National Assembly Budget Office (2007b) **Performance Planning Implementation, Development and Evaluation**. Seoul: NABO.

National Assembly Research Service (2011) **Issues of the Inspection of government offices: the problem and policy alternatives of the Bokeumjari public housing construction**. Seoul: NARS.

National Science and Technology Commission (2011) **The NSTC, Suggestions for Performance Assessment on R&D programmes for the next five years** [online]. Available from: http://www.nstc.go.kr/nstc/civil/report.jsp [Accessed 26th April 2012]

Newland, C. A. (1976) Policy/Program Objectives and Federal Management: The Search for Government Effectiveness. **Public Administration Review,** 36(1): 20-27

Noman, Z. (2007) "United Kingdom." In Teresa, C. (ed.) **Performance Budgeting in OECD Countries**. Paris: OECD. pp.191-208

North, D. C. (1992) Institutions and economic theory. **The American Economist**, 36(1): 3-6

North, D. C. (1995) “The New Institutional Economics and Third World Development.” In Harriss, J. et al. (eds) **The New Institutional Economics and Third World Development**. London: Routledge. pp. 17-26

OECD (2004) **Public Sector Modernisation: Governing for Performance**. Policy Brief. Paris: OECD.

OECD (2005a) **Using Performance Information for Managing and Budgeting: Challenges, Lessons and Opportunities**. Paris: OECD.

OECD (2005b) **Performance Information in the Budget Process: Results of OECD 2005 Questionnaire**. Final Report. Paris: OECD.

OECD (2005c) **Government Performance: Lessons and Challenges**. Paris: OECD.

OECD (2012) **OECD Economic Outlook**. No. 91. Paris: OECD.

OECD (2009) **Employment Outlook**. Paris: OECD.

OMB (2010) **Analytical Perspectives: Budget of the United States Government, Fiscal Year 2011**. Washington, D.C.: U.S. Government Printing Office.

OMB (2008) **Overview of ratings for all federal programs** [online]. Available from: http://georgewbush-whitehouse.archives.gov/omb/expectmore/about.html [Accessed 13th November 2012]

OMB (2006) **New PART Ratings Showcase More Contradictions of Program**. Washington, D.C.: OMB.

OMB (2005) **Guidance for Completing the Program Assessment Rating Tool**. Washington, D.C.: OMB.

Osborne, D. and Gaebler, T. (1992) **Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector**. Massachusetts: Addison-Wesley.

Osborne, S. P., Bovaird, T., Martin, S., Tricker, M., and Waterston, P. (1995) Performance management and accountability in complex public programmes. **Financial Accountability & Management**, 11(1): 19-37

Park, C. (2011) **Nvivo 9**. Korea: Hyung-seol.

Park, H (2009) **The Status of the SABP and its Policy Directions**. Seoul: NABO

Park, N. (2005) **A study of performance management and the introduction of economic incentives in the public sector**. Seoul: Korea Institute of Public Finance.

Park, N. (2008) Does more information improve budget allocations? Evidence and lessons from performance-oriented budgeting in Korea. **Working Paper to be presented at the Congress of the International Institute of Public Finance,** Maastricht: The Netherlands.

Pauly, M. (1974) Overinsurance and Public Provision and Insurance: The Role of Moral Hazard and Adverse Selection. **Quarterly Journal of Economics**, 88(1): 44-62

Pauly, M. (1968) The Economics of Moral Hazard: Comment. **American Economic Review**, 58(3): 531-537

Pawson, R and Tilly, N. (1997) **Realistic Evaluation**. London: Sage.

Pidd, M. (2005) Perversity in Public Service Performance Measurement. **International Journal of Productivity and Performance Management**, 54(5/6): 482-493

Pitches, D., Burls, A. and Fry-Smith, A. (2003) How to make a silk purse from a sow's ear – a comprehensive review of strategies to optimise data for corrupt managers and incompetent clinicians. **British Medical Journal**, 327: 1436-1439

Poister, T. H. (2003) **Measuring Performance in Public and Nonprofit Organization**. CA: Jossey-Bass.

Pollitt, C. (2001) Integrating financial management and performance management. **OECD Journal on Budgeting**, 1(2): 7-37

Pollitt, C. (1990) **Managerialism in the public services: The Anglo-American experience.** Oxford: Basil Blackwill.

Presidential Transition Committee (2008) **Presidential Agenda of New Administration**. Seoul: Korea.

Prime Minister's Office (2008) **Implementation plan of government performance assessment**. Seoul: PMO.

Primenews (2011) **The MLTM will continue to construct 1.5million 'Bokeumjari' public housing as planned** [online]. Primenews. Available from:

http://www.newsprime.co.kr/news/articleView.html?idxno=200598
[Accessed 20 October 2011]

Radnor, Z. J. (2008) Muddled, massaging, manoeuvring, or manipulated?: A typology of organisational gaming. **International Journal of Productivity and Performance Management**, 57(4): 316-328

Radnor, Z. J. and Barnes, D. (2007) Historical analysis of performance measurement and management in operations management. **International Journal of Productivity and Performance Management**, 56(5): 384-396

Ridgway, V. F. (1956) Dysfunctional Consequences of Performance Measurements. **Administrative Science Quarterly**, 1(2): 240-247

Ridley, C. E. and Simon, H. A. (1943) **Measuring Municipal Activities: A Survey of Suggested Criteria for Appraising Administration**. Chicago: The International City Management Association.

Robinson, M. and Brumby, J. (2005) **Does performance budgeting work? Analytical Review of the Empirical Literature**. International Monetary Fund.

Rouse, J. (1999) "Performance management, quality management, and contracts." In Horton, S. and Farnham, D. (eds) **Public management in Britain**. Basingstoke: Macmillan. pp. 76-93

Rubin, M. and Willoughby, K. (2011) Measuring Government Performance: The Intersection of Strategic Planning and Performance Budgeting. **Working Paper to be presented at the Conference of the 7th Transatlantic Dialogue**, Newark, New Jersey: U.S.

Sackett, P. R., Burris, L. R. and Callahan, C. (1989) Integrity testing for personnel selection: An update. **Personnel Psychology**, 42: 491–528.

Shah, A. (2007) **Budgeting and Budgetary Institutions**. Washington D.C.: World Bank.

Shea, R. T. (2008) Performance Budgeting in the United States. **OECD Journal on Budgeting,** 8(1): 1-13

Shick, A. (2007) Performance Budgeting and Accrual Budgeting: Decision Rules or Analytic Tools? **OECD Journal on Budgeting**, 7(2): 109-138

Shick, A. (1992) “Incremental Budgeting in a Decremental Age.” In Hyde, A. C. (ed.) **Government Budgeting: Theory·Process·Politics**. California: Brooks/Cole. pp. 410-425

Shin, M. (2010) **A Study on the Distortions and Gaming of Performance Measurement in the Public Sector**. Doctoral dissertation, Sungkunkwan University, Seoul: Korea.

Shin, S. (2012) Dysfunctional Consequences of the Korean Performance Budgeting System and their Policy Implications: A Multiple Case Study in Three Case Areas. **Working Paper to be presented at the Conference of the 8th Transatlantic Dialogue**, Radboud University, Nijmegen: The Netherlands.

Shin, S. and Cha, K. (2010) **A study of the inspection method for performance targets and the data quality of the performance plan**. Seoul: The Audit and Inspection Research Institute.

Simon, H. A. (1978) Rationality as Process and as Product of Thought. **The American Economic Review**, 68(2): 1-16

Smith, P. C. (1990) The use of performance indicators in the public sector. **Journal of the Royal Statistical Society**, 153(1): 53-72

Smith, P. C. (1993) Outcome-related performance indicators and organizational control in the public sector. **British Journal of Management**, 4: 135-151

Smith, P. C. (1995) On the unintended consequences of publishing performance data in the public sector. **International Journal of Public Administration**, 18(2/3): 277-310

Smith, P. C. (2002) Performance management in British Health Care: Will it deliver? **Health Affairs**, 21(3): 103-115

Smith, P. C. (2005) Performance indicators: good, bad, and ugly. **Royal Statistical Society**, 168(1): 1-27

Smith, P. C. (2008) England: Intended and Unintended Effects. In Wrsmar et al. (eds) **Health targets in Europe: Learning from experience**. European Observatory: WHO. pp. 63-81

Smith, P. C. et al. (2011) Leadership and governance in seven developed health systems. **Health Policy**, 106(2012): 37-49

Spiller, P. T. (2008) An Institutional Theory of Public Contracts: Regulatory Implications. **NBER Working Paper** No. 14152

Stewart, J. and Walsh, K. (1994) Performance Measurement: When Performance can Never be Finally Defined. **Public Money & Management**, 14(2): 45-49

Talbot, C. (2010) **Theories of performance: organizational and service improvement in the public domain**. Oxford: Oxford University Press.

Talbot, C. (2005) "Performance Management". In Ferlie, E., Lynn, L. E. and Pollitt, C. (eds) **The Oxford Handbook of Public Management**. Oxford: Oxford University Press. pp. 491- 520

Tashakkori, A. and Teddlie, C. (1998) **Mixed methodology: Combining qualitative and quantitative approaches**. Thousand Oaks, CA: Sage.

van Thiel, S. and Leeuw, F. L. (2002) The Performance Paradox in the Public Sector. **Public Productivity & Management Review**, 25(3): 267-281

Walsham, G. (1993) **Interpreting information systems in organizations**. Wiley: Chichester.

Wang, X. H. (2000) Performance measurement in budgeting: A study of county governments. **Public Budgeting & Finance**, 20(3): 102-118

Wang, X. H. (2002) Assessing Performance Measurement Impact: A study of U.S. Local Governments. **Public Performance & Management Review**, 26(1): 26-43

Weitzman, M. L. (1980) The "Ratchet Principle" and Performance Incentives. **The Bell Journal of Economics**, 11(1): 302-308

Wholey, J. S. (1999) Performance-Based Management: Responding to the Challenges. **Public Productivity & Management Review**, 22(3): 288-307

Wholey, J. S. and Hatry, H. P. (1992) The Case for Performance Monitoring. **Public Administration Review**, 52(6): 604-610

Wholey, J. S. and Newcomer, K. E. (1997) Clarifying Goals, Reporting Results. **New Direction for Evaluation**, 75: 91-98

Wildavsky, A. (1974) **The Politics of the Budgetary Process**. Boston: Little Brown.

Williamson, O. E. (1990) **Transaction Cost Economics and Organization Theory**. Oxford: Oxford University Press. pp. 207–256

Williamson, O. E. (1985) **The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting**. New York: The Free Press.

Williamson, O. E. (1975) **Markets and Hierarchies: Analysis and Antitrust Implications**. New York: Free Press.

Wilson, J. Q. (1989) **Bureaucracy: What government agencies do and why they do it.** New York: Basic Books.

Winstanley, D. and Stuart-Smith, K. (1996) Policing performance: the ethics of performance management. **Personnel Review**, 25(6): 66-84

World Bank (2010) **Results, Performance Budgeting and Trust in Government.** Washington D.C.: World Bank.

Yeonhapnews (2008) **MLTM's work plan in 2009** [online]. Yeonhapnews. Available from: http://blog.naver.com/dhjjh0153?Redirect=Log&logNo=100058965749 [Accessed 20 October 2011]

Yin, R. K. (2009) **Case Study Research: Design and Methods**. 4th ed. London: SAGE.

Yoon, C. and Lee, G. (1992) **Industrial Economic Theory**. 2nd ed. Seoul: Bup Moon Sa.

Yoon, S. (2001) **Government and Budget**. Seoul: Dae Young Sa.