THE IMPACT OF THE SOUTH KOREAN PUBLIC INSTITUTION MANAGEMENT EVALUATION SYSTEM ON THE ACCOUNTABILITY OF SOUTH KOREAN PUBLIC INSTITUTIONS

By

MINKYU JIN

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

Institute of Local Government Studies

School of Government

College of Social Sciences

University of Birmingham

April 2023

UNIVERSITYOF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

ABSTRACT

This thesis examines the effectiveness of performance management systems in strengthening accountability in the public sector. Public accountability can serve as a tool to induce public organisations to improve performance by providing rewards and sanctions related to the effects of their performance. Therefore, many countries have endeavoured to design and implement performance management systems to improve public accountability according to the international guidelines on performance management systems provided by the OECD. This study adopts a case study design using the South Korean Public Institution Management Evaluation System (PIMES) because the PIMES is designed in accordance with the international guidelines. The study examines the effectiveness of PIMES in incentivising public institutions to improve accountability and performance and mitigating the strategic behaviour of public institutions. A mixed methods research is employed in this study. A quantitative analysis using regression models is conducted with the PIMES data for 2014-2019. The quantitative analysis finds that the PIMES leads the public institutions to improve social accountability, but does not impact managerial accountability. It also finds citizens' role in improving quality of public services through performance management systems. A qualitative analysis drawing on 42 semi-structured interviews and documents suggests that non-financial incentive schemes can be more effective for improving the accountability of public organisations than financial incentive schemes. As for strategic behaviour and gaming, the quantitative and qualitative data provide the finding that the strategic behaviour of public institutions occurs frequently in the PIMES evaluation process. This thesis contributes to a new theoretical understanding of ratchet effects and threshold effects by identifying that '*irrelevant ratchet effects*' and '*individual threshold effects*' exist due to a target-setting method in performance management systems.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

CHAPTER 4 RESEARCH DESIGN AND METHODOLOGY

CHAPTER 5 -THE IMPACT OF PIMES RANK ON THE ACCOUNTABILITY AND STRATEGIC BEHAVIOUR OF PUBLIC INSTITUTIONS

CHAPTER 6 AN EXAMINATION OF INCENTIVE SCHEMES, PERFORMANCE INDICATORS AND STRATEGIC BEHAVOIOUR

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

ALIO   All Public Information In-One

MOEF   Ministry of Economy and Finance

NDPB   Non-Departmental Public Bodies

NHS   National Health Service

NPM   New Public Management

OECD   Organisation for Economic Co-operation and Development

PIMES   Public Institution Management Evaluation System

UKGI   UK Government Investments

Chapter 1 Introduction

1.1 Background and context

This thesis aims to examine the relationship between performance management systems and public accountability and performance. Throughout this thesis, performance management systems refer to those in the public sector. The thesis explores the effect of performance management systems on public accountability and performance with panel data, documents and interview data.

Performance management systems have been prevalent in the public sector across the globe in response to demand for performance and the problems of coordination (Pollitt, 2013; Jantz, Christensen and Lægreid, 2015; Taylor, 2021). The drive for performance has been a growing concern for governments in many countries since the global economic crisis of the second half of the 1970s (Pollitt and Bouckaert, 2017; Van Dooren, Bouckaert and Halligan, 2015). Many countries including Australia, New Zealand, the UK and the US launched programs of central government reform to make the public sector work better and cost less (Pollitt and Bouckaert, 2017; Hammerschmid, Van de Walle, Andrews and Mostafa, 2019). In addition, central government reform had fragmented public organisations and new agencies were positioned at arm's length from ministers in the Netherlands, New Zealand, Sweden, the UK and the US (Pollitt, Talbot, Caulfield and Smullen, 2004; Pollitt and Bouckaert, 2017). As a result, governments have faced a coordination problem of getting many agencies to pursue cooperatively the same policy objective and there has been a lack of accountability to public concern in the arm's length agencies (Mulgan, 2014; Pollitt and Bouckaert,

2017). With respect to performance management systems to increase performance and mitigate the problems of coordination, research has mainly examined the relationship between different performance management systems and performance improvement (Verbeeten, 2008; Dewettinck and van Dijk, 2013, Speklé and Verbeeten, 2014; Gerrish, 2016). However, empirical research on the effect of performance management systems on different types of public accountability is scarce (Jantz et al., 2015). This thesis aims to fill the gap in the literature by examining the effect of performance management systems on managerial and social accountability of public organisations.

The global appeal of performance management is based on the assumption that it has the capacity to improve accountability of public organisations (Jantz et al., 2015; Taylor, 2021). However, the relationship between performance management and public accountability is still contested (Christensen and Lægreid, 2015; Jantz et al., 2015; Jones and Bouckaert, 2017). Whilst some scholars argue that performance management would make public organisations more accountable (Van de Walle and Cornelissen, 2014; Yang and Torneo, 2015), others argue that performance management has limited capacity for strengthening accountability (Van Dooren et al., 2015; Jones and Bouckaert, 2017). This thesis employs the South Korean Public Institution Management Evaluation System (PIMES) as a single case because the PIMES aims to increase accountability of over 100 South Korean public institutions across various public services and has been in operation for over 35 years. The thesis contributes to the academic debate about the effect of performance management systems on public accountability by examining the impact of the PIMES on the managerial and social accountability of public institutions.

In performance management systems, performance information can be used for improving

accountability and performance of organisations by linking it to financial or non-financial incentive schemes (Propper and Wilson, 2003; Van Dooren et al., 2015). For example, performance-related pay has been adopted as a financial incentive scheme in the majority of countries in the Organisation for Economic Co-operation and Development (OECD) and many developing countries (Belle and Cantarelli, 2015; Weibel, Rost and Osterloh, 2010). Publication of performance information functions as a non-financial incentive scheme because it can affect the behaviours of employees in the public sector through individuals' pride in their league table's position or avoidance of being given the label of a failing organisation (Propper and Wilson, 2003).

Many researchers have examined how financial and non-financial incentive schemes in performance management systems affect accountability and performance of public organisations (Weibel et al., 2010; Mayumana et al., 2017). While some studies show a positive relation of incentive schemes on public accountability and performance (Kahn, Silva and Ziliak, 2001; Franco, Bennett, Kanfer and Stubblebine, 2004; Lee, 2021), others argue limited efficacy of incentive schemes on public accountability and performance (Perry, Engbers and Jun, 2009; Belle and Cantarelli, 2015). This thesis sheds light on the academic debate about the effect of incentive schemes on public accountability and performance by examining the effectiveness of financial and non-financial incentive schemes through a case study of Korea's PIMES.

Performance indicators accompany targets and are used to monitor performance of organisations and allocate performance-related pay in performance management systems (Carter, Klein and Day, 1995; Hood, 2006; Van Dooren et al., 2015). Target-setting is a pivotal point in performance management systems (Stringer and Shantapriyan, 2012). Design of performance indicators and target-setting in

performance management systems needs to be considered when implementing the systems (Bracci, Maran and Inglis, 2017). This thesis examines the effect of performance indicators and target-setting on the accountability and performance of public institutions. The examination contributes to research on operational methods of performance management systems in practice.

Performance management systems are based on the assumption that they improve the accountability of employees in the public sector by managing their behaviours (Carter et al., 1995; Propper and Wilson, 2003). However, many scholars argue that performance management systems tend to become ineffective over time due to strategic behaviour or gaming of actors (Meyer, 1997; Hood, 2012). This thesis contributes to literature on strategic behaviour or gaming under performance management systems by exploring the strategic behaviour or gaming of public institutions in a case of Korea's PIMES.

1.2 Research questions and methods

In order to explore these gaps in the literature, this thesis has set three main research questions:

1. Do performance management systems improve the accountability of public organisations?

2. How do the conditions and processes of performance management systems affect the accountability of public organisations?

3.   What is the impact of performance management systems on the strategic behaviour of public organisations?

The first research question is answered by examining how a performance management system has affected the accountability of public organisations over time. Public accountability can be divided into various types of accountability, however, this thesis focuses on managerial and social accountability that has been emphasised since the advent of New Public Management (NPM) (Ospina, Grau and Zaltsman, 2004; Lægreid, 2014). This research aims to examine whether the PIMES evaluation process improves the level of managerial and social accountability of public institutions.

This research employs linear regression models with a data set for 2014-2019, which consists of the ratio of profit to assets, customer satisfaction scores, welfare benefits, PIMES evaluation results, type and stock exchange listing of public institutions, number and wage of employees, and a head's past career. . The regression models are appropriate because the models can identify the factors to affect the managerial and social accountability of public institutions. Thus, the regression models are set out with independent variable to measure the PIMES evaluation results and control variables in relation to external factors.

The second research question will analyse the PIMES to identify the individual effects of operational methods of performance management systems on the accountability and performance of public organisations. The PIMES evaluation process consists of target-setting, evaluation of management performance reports and feedback of the evaluation results through rewards and sanctions. This thesis

analyses and interprets the effect of the target-setting method, the use of performance indicators and the feedback through incentive schemes in the PIMES on the accountability and performance of public institutions. Semi-structured interviews with 42 key participants in the evaluation process and documents such as the management performance reports and guidance on management evaluation are used to answer the second research question.

The third research question is addressed by exploring the strategic behaviour and gaming of public organisations in the evaluation process under performance management systems. The regression models are used to analyse the effect of the PIMES evaluation process on the strategic behaviour of public institutions. The finding from the analysis is interpreted in relation to the strategic behaviour and gaming of public organisations. Moreover, the interview results are analysed and interpreted to examine how the strategic behaviour of public institutions occurs when they produce management performance reports, conduct their tasks and participate in customer satisfaction survey.

The three research questions are closely connected to examine systematically the effect of performance management systems on the accountability and performance of public organisations. While the first research question examines how the PIMES evaluation process has affected the public accountability with a longitudinal data set for 2014-2019, the second and third research questions investigate how the current operational methods of PIMES influence the public accountability and performance individually. Literature on performance management practices and the consequences of different performance management system designs in public organisations is limited (Stringer, 2007). Therefore, this thesis contributes to literature on performance management systems by examining the impact of PIMES designs on the accountability and performance of public institutions through the

three research questions.

1.3 Contribution to knowledge

The purpose of this thesis is to examine the impact of performance management systems on different types of public accountability. Public accountability can be categorised into different types such as political, managerial, legal and social accountability and their scopes vary (Day and Klein, 1987; Bovens, 2010; Jantz et al., 2015). Whilst considerable research has been undertaken examining whether performance management systems strengthen public accountability (Kenk and Haldma, 2016; Jones and Bouckaert, 2017; Mizrahi and Minchuk, 2019a), there is limited research on the effect of performance management systems on different types of public accountability (Jantz et al., 2015). The effect of performance management systems on public accountability can be different across types of public accountability. It is useful to discern different effects of performance management systems on different types of public accountability for design of better performance management systems. Therefore, it is important to research the effect of performance management systems on different types of public accountability. This thesis contributes to research on the relationship between performance management systems and public accountability by suggesting the significant positive effect of PIMES on the social accountability of public institutions and insignificant effect on the managerial accountability.

Abundant research has been undertaken to examine whether financial incentive schemes such as performance-related pay improve accountability and performance of public organisations. However,

research on the effectiveness of non-financial incentive schemes is limited to areas of health and education (Perry et al., 2009; Ferreira and Otley, 2009; Schlechter, Thompson and Bussin; 2015; Burgess, Metcalfe and Sadoff, 2021). This thesis fills a gap in the research by examining the effectiveness of non-financial incentive schemes across various areas of public services provided by public organisations financed by central government as well as financial incentive schemes. Target-setting has a critical impact on performance management systems through its link with rewarding performance, however, there is limited research on the level of targets set by organisations or target-setting process (Ferreira and Otley, 2009; Stringer, 2007). This thesis contributes to research on target-setting in performance management systems by investigating the effect of current target-setting method in the PIMES on accountability and performance of public institutions.

Literature on strategic behaviour and gaming under performance management systems suggests ratchet effects, threshold effects, output distortion, effort substitution and sampling bias as major types of strategic behaviour and gaming (Bohte and Meier, 2000; Hood, 2006; Bevan and Hood, 2006; Taylor, 2021). In the PIMES evaluation process, the South Korean public institutions propose their new targets and performance indicators periodically and the South Korean government determines whether the new targets and performance indicators are adopted. This target-setting process in the PIMES evaluation is different from uniform target-setting or the incremental increase of the level of an existing target where previous research has found strategic behaviour and gaming of public organisation (Hood 2006; Taylor, 2021). This thesis focuses on the particular target-setting process in the PIMES evaluation, therefore, finds '*irrelevant ratchet effects*' and '*individual threshold effects*' that build on ratchet effects and threshold effects discussed in literature.

Lastly, this thesis contributes to ascertaining the impact of NPM reform. The main weight of research on the NPM has usually been placed on countries such as Australia, New Zealand, the UK and the US which are considered the originators and most vigorous proponents of the NPM (Pollitt and Dan, 2013; Funck and Karlsson, 2019). While the NPM has been partly superseded in European countries, NPM reforms have become an integral part of public sector reforms in developing countries since the 1990s (Van Dooren et al., 2015; Samaratunge, Alam and Teicher, 2008; Haque, 2020). However, most research has been undertaken about NPM reforms for western countries, whereas limited research has been undertaken about NPM reforms in Asia, South America and Eastern Europe for 1991-2016 (Funck and Karlsson, 2019; Samaratunge et al., 2008). The study of the South Korean performance management system, therefore, contributes to research on NPM reforms, looking at a non-western context for adding to knowledge on the impact of NPM reforms.

1.4 Structure of the thesis

The thesis consists of eight chapters. After this introduction (Chapter 1), Chapter 2 follows, which introduces different types and understandings of public accountability, and then provides definition of accountability deficit. The chapter reviews the doctrines of NPM theory, changes in public accountability under the NPM reforms and academic debate about relationship between performance management systems and public accountability. Subsequently, examples of incentive schemes and the relationship between incentive schemes and motivation are discussed to examine the effect of incentive schemes on public accountability and performance. After that, the definition, types and criteria of performance indicators are provided. Lastly, an overview of principal-agent theory is

presented to support analysis of strategic behaviour and gaming of public institutions under performance management systems.

Chapter 3 discusses international guidelines on performance management systems, which have been provided by the OECD. The chapter reviews the different performance management systems in Canada, France, Sweden and the UK, and compares the PIMES with them. The South Korean political systems and governance are explained and the influence of the principles of NPM on the PIMES is reviewed. After that, the definition, functions, subjects, history, evaluation process and performance indicators of the PIMES are explained.

Chapter 4 provides research design and methodology for this thesis, which consist of case study and mixed methods research design. The chapter develops the three research questions stemming from the literature review, and then shows how this study's research method is appropriate for addressing each research question. Subsequently, the chapter discusses this study's philosophical perspective of critical realism, and explains the quantitative and qualitative approaches in detail.

Chapter 5 analyses the relationship between the PIMES and the managerial and social accountability and strategic behaviour of public institutions with the regression model set out in Chapter 4. A discussion of the data set and a preliminary statistical analysis including correlation analysis and scattergrams are presented, and then three hypotheses regarding the relationship are tested at a specific significance level by using the P-values. The chapter analyses the test results with the interview data and the effect of the organisation, personnel and finance character-related factors on the managerial and social accountability and strategic behaviour of public institutions.

Chapter 6 analyses the effect of the financial and non-financial incentive schemes, performance indicators and target-setting method in the PIMES on the accountability and performance of public institutions. The empirical analysis of the effect presents insight into how design of performance management systems influences accountability and performance of public organisations. After that, the chapter analyses the strategic behaviour and gaming of public organisations under performance management systems.

Chapter 7 discusses the implications of findings from the empirical analysis and this study's contribution to knowledge in terms of theoretical, methodological and empirical contributions. The chapter discusses the significant positive effect of performance management systems on the social accountability of public institutions and the finding that non-financial incentive schemes are more effective than financial incentive schemes in performance management systems. In addition, the chapter discusses '*irrelevant ratchet effects*' and '*individual threshold effects*' beyond the strategic behaviour and gaming discussed in the literature.

Chapter 8 concludes this thesis by considering the implications of the effect of performance management systems on social accountability, the effectiveness of non-financial incentive schemes and various strategic behaviour and gaming including '*irrelevant ratchet effects*' and '*individual threshold effects*' under performance management systems. The chapter recommends use of different target-setting methods, expansion of non-financial incentives, responsible ministries' verification of customer lists and continuity in evaluators for refining the PIMES. The chapter considers the limitations of this study and suggests future research areas.

Chapter 2 Literature Review: Accountability, Performance Indicators, and Strategic behaviour

2.1 Introduction

The aim of this research is to examine the impact of performance management systems on public accountability and performance. Therefore, the focus here will be on introducing the concept of accountability found in the literature and a discussion of incentive schemes and performance indicators in performance management systems and strategic behaviour of public organisations.

In public administration, political science, international relations, social psychology and business administration, attention on accountability has grown and is a central focus of scholarly debate and analysis (Dubnick, 2014; Lindberg, 2013; Bovens, Goodin and Schillemans, 2014). For example, the frequency of the word 'accountability' in books based on a million scanned volumes published in English has increased by ten times from 0.0001% in the 1970s to 0.001% in the 2000s (Dubnick, 2014). This increasing use of accountability may result from its association with related public values such as democracy, efficiency and greater administrative performance (Dubnick, 2014). Scholars in different disciplines have focused on different types of accountability (Bovens et al., 2014). For example, political scientists often approach accountability in terms of a relationship between elected politicians and their voters and public administration studies of accountability tend to focus on forms of accountability in relation to public service provision and regulation (Bovens et al., 2014). Therefore,

different types of accountability need to be identified in order to examine accountability in the public sector.

Discussion of accountability in the public sector has been influenced by the onset of NPM principles and diverse accountability mechanisms have subsequently been applied. Scholars of public administration have tried to examine what factors can secure good accountability in practice due to 'accountability deficit' which refers to observable lack of accountability and an attitude of disapproval towards such a lack (Mulgan, 2014; Han, 2020). The first section of this chapter explains the definition, characteristics and types of public accountability and the definition of 'accountability deficit' and relates these concepts to the key elements of NPM.

The second section of this chapter reviews the literature on performance-based incentive schemes in performance management systems, which are used to reward or punish the actors to improve accountability (Brandsma and Schillemans, 2013; Jantz et al., 2015; Han, 2020). Subsequently, definition, typology and criteria of performance indicators are discussed. A major objective of NPM reforms is the improvement of performance in the public sector and hence administrative executives are held more accountable to their performance and the work of their agencies under NPM reform (Christensen and Lægreid, 2015). Requirements to account for performance have obliged public organisations to put performance indicators in place (Carter, Klein and Day, 1995). Therefore, the investigation of performance indicators contributes to analysing the accountability of public organizations.

The third section of this chapter discusses strategic behaviour and gaming of public organisations

under performance management systems. Many scholars suggest that strategic behaviour and gaming of public organisations hinder performance management systems in improving accountability of the organisations (Goodhart, 1984; Hood, 2012; Van Dooren et al., 2015). Therefore, it is important to review types of strategic behaviour and gaming discussed in the literature to explore the effect of performance management systems on public accountability. This thesis explores strategic behaviour and gaming of public organisations, using principal-agent theory which can provide insights into key dimensions of accountability within public services (Behn, 2001; Lindberg, 2013; Murphy, 2019). The theory presents a flexible and useful framework for analysing accountability of public decision-makers and comparing the potential of diverse accountability mechanisms for improving accountability of agents (Bovens et al., 2014; Murphy, 2019). In order to attain an understanding of principal-agent theory, the third section starts with an overview of this theory, and is followed by a reflection on strategic behaviour and gaming of agents, signalling and screening.

2.2 Accountability

2.2.1 Definition and characteristics of accountability

This thesis works from the assumption that increasing accountability can result in improvement in the performance of government agencies (Dubnick, 2014; Rana and Hoque, 2020). Based on this assumption, the thesis focuses on the relationship between accountability and performance management systems. It is not easy to define accountability in a single phrase, because accountability operates in different ways and is understood through various approaches, activities and techniques

14

(Boyne, Gould-Williams, Law and Walker, 2002; Bovens, 2010). Thus, this section describes the diverse definitions of accountability, notably of accountability as a value and as a mechanism and accountability under professionalism, and draws out key characteristics of accountability in order to enable empirical understanding and analysis within the broad field of public administration.

Simon et al. (1950, p. 513) argue that, "Accountability consists of the methods, procedures and forces that determine what values will be reflected in administrative decisions". According to Day and Klein (1987), accountability implies both a shared set of expectations and a common currency of justifications. Expectations include agreement about the reason why one actor owes explanation to another, since this sense of obligation translates the giving of account into accountability (Day and Klein, 1987). For example, when a delinquent is required to explain his or her behaviour, the delinquent's explanation of his or her behaviour would not amount to accountability but excuse, apology or pretext. This is because there is no agreement about the reason why the delinquent owes the explanation to another (Day and Klein, 1987). Justification constitutes good reason for explaining one actor's conduct when the actor defends his or her conduct (Day and Klein, 1987). Based on both expectation and justification, accountability can be defined as "a relationship between an actor and a forum, in which the actor has an obligation to explain and to justify his or her conduct, the forum can pose questions and pass judgement, and the actor may face consequences" (Bovens, 2007, p. 450).

Accountability can be divided into accountability as a virtue and accountability as a mechanism in order to provide some foundation for comparative and cumulative empirical research (Bovens, 2010). While accountability as a virtue is primarily used as a normative concept such as standards for the evaluation of the behaviour of agents, accountability as a mechanism refers to an institutional relation

in which an agent can be held to account by a principal (Bovens, 2010). Although it is very difficult to settle on a single definition of accountability as a virtue due to various standards of many public organisations and political systems, scholars have sought to clarify a broad notion of accountability as a virtue (Bovens, 2010). For example, accountability is seen in this sense as possessing three dimensions such as performance, finance and fairness, therefore, accountability is present when public services possess high quality and low cost, and are performed in a courteous manner (O'Connell, 2005).

Accountability as a mechanism encompasses a variety of accountability relations of actors such as civil servants and public institutions and forums such as audit office and parliament, in which actors have an obligation to explain and justify their conduct and forums pass judgement (Bovens, 2010; Bertelli, 2012; Peters, 2014). The accountability relations between the actors and forums consists of at least three elements including the provision of various types of information about performance, interrogation of the adequacy of information and judgement on the conduct of the actors (Bovens, 2010).

The growth of professionalism in the twentieth century has compounded the concept of accountability in complex modern society (Day and Klein, 1987; Vriens, Vosselman and Groβ, 2018). A central notion of professionalism is that the profession is accountable to its peers rather than parliament or the public (Day and Klein, 1987). The accountability under professionalism can be observed in a type of quasi non-governmental organisation (QUANGO) which is collectively given a monopoly of practice in return for policing the competence of their members (Day and Klein, 1987). In addition, professional accountability may be found in white-collar professional services (Watt, 1998). As

white-collar professional services are provided by experts, it is difficult for a local authority to judge the quality of the services (Watt, 1998). In this situation, there is a possibility of opportunistic behaviour by the professional services providers (Watt, 1998).

Literature says that accountability is related to capabilities of principals to judge the performance of their agents (Lindberg, 2013; Gailmard, 2014; Murphy, Ferry, Glennon and Greenhalgh, 2019). The capabilities of principals are also in part a function of institutions themselves (Gailmard, 2014). The relationship between accountability and institutional structure can be understood with 'principal-agent theory' (Murphy et al., 2019; Gailmard, 2014), which will be explained in more detail in the third section of this chapter. An agent is understood as an actor or institution who is to give an account, whilst a principal is an actor or institution who receives that account from the agent (Lindberg, 2013). A sanctions-based model can be enacted through the principal-agent theory. Such a model focuses on the right of the principal to sanction the agent due to failure to justify decisions with regard to responsibilities (Murphy et al., 2019; Lindberg, 2013). The principal's capacity to decide the extent to which the agent has properly performed their responsibilities and ability to impose sanctions are central to accountability (Stewart, 1984).

From these diverse perspectives on accountability, the main characteristics of accountability can be identified as:

- an obligation to explain one actor or institution's conduct between principal and agent

- the evaluation of the agent's performance

- sanction based on the performance

These characteristics are well reflected in accountability processes which consist of three major phased components; information, discussion and consequence (Brandsma and Schillemans, 2013). In the first phase of information, the agent provides an account of its actions and performance through formats such as annual reports, financial documents and interactive information on a website. Following this, the principal evaluates the information as the second phase in the form of discussion. The last phase is consequence, in which the principal passes judgement on the performance of the agent by accordingly punishing or rewarding them (Mulgan, 2003; Brandsma and Schillemans, 2013; Van Dooren et al., 2015).

2.2.2 Definition and importance of public accountability

As explained earlier, this thesis focuses on accountability in the public sector, which applies to public institutions, as exemplified through performance management systems. Public accountability is differentiated from private accountability (Bovens et al., 2014). While private accountability refers to people's obligation to provide parents, bosses and neighbours with answers in their daily lives, public accountability applies to civil servants or public organisations (Bovens et al., 2014; Schillemans, 2016). Public organisations are run in accordance with the laws, budgets and aims of different services set by political decisions (Bejerot and Hasselbladh, 2013). As a result, public organisations are engaged not only in a negotiation network with responsible ministries, the Ministry of Finance, and parliament but also in a consumption network with users (Jørgensen, Hansen, Antonsen and Melander, 1998). Public organisations include ministries, local government, state-

owned companies, agencies, the police, public schools, public health-care enterprises and public museums (Christensen, Lægreid and Røvik, 2020).

Public accountability is accountability in and about the public domain, it is seen in a variety of ways in terms of 'public' (Bovens et al., 2014). Firstly, 'public accountability' refers to openness or transparency. In public accountability, the information provided about an actor's conduct is generally accessible and debate about the information is open to the public. Moreover, members of parliament, organisational superiors and ombudsmen promulgate their judgement on the information to the public at large. Secondly, 'public accountability' regards matters of public concern such as the spending of public funds, the exercise of public powers and the conduct of public organisations. Thirdly, 'public accountability' can refer to accounting perspectives and accounting standards. In other words, public accountability implies accounting that is performed with a view to the public interest or to public responsibilities.

Public accountability is of crucial importance in democracies for different reasons (Bovens, 2010; Bovens et al., 2014; Schillemans, 2016). Firstly, public accountability provides legitimacy to public officials and public organisations (Bovens, 2010). This is because public accountability in the sense of transparent, responsive and responsible governance can assure public confidence in government and bridge the gap between governed and government (Aucoin and Heintzman, 2000; Mulgan, 2003). Secondly, public accountability is instrumental in achieving important elements of democratic governance (Bovens, 2010). This is because public accountability imposes the obligation of giving account to voters and media, which is crucial as a feedback mechanism to establish popular control over government (Mulgan, 2003). Thirdly, public accountability contributes to preventing corruption

and the abuse of public powers by providing checks and balances (Aucoin and Heintzman, 2000; Bovens, 2010). For example, the public has delegated to parliament the power to request government to give account over public expenditures, coercive powers and equal treatment of citizens through public accountability mechanisms (Bovens et al., 2014). Lastly, public accountability can serve as a tool to induce governments, agencies and individual officials to improve their performance, since it provides external feedback such as sanctions related to the intended and unintended effects of their policies (Aucoin and Heintzman, 2000; Bovens, 2010).

2.2.3 Types of public accountability

This thesis aims to examine the effect of performance management systems on different types of public accountability. With respect to types of public accountability, many research studies have been undertaken (Boyne et al., 2002; Jantz et al., 2015; Bovens et al., 2014). Public accountability can firstly be classified into two major categories: external accountability and internal accountability (Boyne et al., 2002). External accountability operates when public organizations give an account to external individuals or organisations such as the public or different levels of government (Boyne et al., 2002). For example, local government is accountable to central government through various mechanisms such as performance indicators, audit, inspection reports and budgetary controls (Power, 1997). In contrast, internal accountability is about the relationship between superiors and subordinates within an organization (Stewart, 1984). Internal accountability traditionally meant accountability for applying the rules in a hierarchy, however, with the advent of NPM it has been transformed into accountability for outputs and outcomes in executive agencies by notions of increased efficiency, competition and cost effectiveness (Osborne and Gaebler, 1992; Boyne et al., 2002; Christensen and

Lægreid, 2015).

Public accountability can be divided into another two categories: political and managerial accountability (Day and Klein, 1987; Jantz et al., 2015). Political accountability is about those who have delegated authority answerable for their actions to the people, for instance, via parliament or managerial hierarchy (Day and Klein, 1987). For example, ministers explain their policies to members of parliament and amend their departments' shortcomings in accordance with political accountability (Woodhouse, 1994; Lam, 2009). This accountability is regarded as a fundamental condition for a democratic political system and is emphasised under traditional public administration (Ospina, Grau and Zaltsman, 2004).

Managerial accountability is about monitoring outputs or results and holding those with delegated authority to use public resources answerable for their performance according to agreed criteria (Jantz et al., 2015). This accountability includes responsibilities for an efficient and effective use of resources (Ospina, Grau and Zaltsman, 2004), because managerial accountability refers to a manager's accountability for the achievement of defined and agreed resource allocation objective (Glynn and Murphy, 1996). Managerial accountability is a more neutral and technical exercise involving evaluation of whether tasks are performed efficiently than political accountability (Christensen and Lægreid, 2015). Therefore, managerial accountability has been emphasised under NPM reforms that aim to improve performance in the public sector through accountability (Ospina et al., 2004; Lægreid, 2014; Christensen and Lægreid, 2015). This thesis focuses on the managerial accountability of public institutions because it is assumed that performance management systems aim to increase performance in the public sector through the improvement of managerial accountability

(Christensen and Lægreid, 2015; Lapuente and Van de Walle, 2020).

Subsequently, public accountability can be classified on the basis of three questions (Pollitt, 2003; Bovens et al., 2014). The first question is to whom the actor is required to give account. According to the first question, accountability is classified into political, administrative, legal and social accountability. Whilst political accountability is to give account to political representatives, parliamentary committees and voters, administrative accountability is to give account to courts of audit, ombudsmen and regulatory agencies (Bovens et al., 2014). For example, public managers may have to appear before parliamentary committees to answer questions according to political accountability (Pollitt, 2003).

Social accountability is defined as the extent and capability of citizens to hold the state accountable and make it responsive to their needs (World Bank, 2012; cited in Grandvoinnet, Aslam and Raha, 2015). Social accountability includes citizen monitoring and oversight of public sector performance, user-oriented public information access and citizen participation in resource allocation decision-making such as participatory budgeting (Fox, 2015). In accordance with social accountability, citizens can state their preferences about specific issues (Ahmad, 2008). Social accountability impacts on improved provision of public service because articulation of citizens' preferences about public services leads to responsive public officials and service providers, and results in better policy design and public services (Grandvoinnet et al., 2015; Brinkerhoff and Wetterberg, 2016; Hickey and King, 2016). Therefore, this thesis examines the effect of performance management systems on social accountability to explore how performance management systems affect the quality of public services.

The second question is who should give account. Accountability is sorted into organisational, hierarchical and collective accountability pursuant to the second question. Organisational accountability means that an organisation as a corporate entity is required to give an account in legal procedure (Bovens, 2010). Hierarchical accountability works when only top of an organisation such as minister or CEO is called to account externally (Bovens et al., 2014). For example, a minister is accountable for civil servants in a parliamentary system (Bovens, 2010). Collective accountability refers to accountability of a group for a particular act or acts committed by individual members of the group (Darcy, 2007). In other words, collective accountability involves persons being made accountable to the acts of others (Darcy, 2007). The third question is why the actor feels compelled to give account. Based on the nature of obligation, accountability is categorised into vertical, horizontal, and diagonal accountabilities (Bovens, 2010). Vertical, horizontal and diagonal accountability separately arises from hierarchical, contractual, and voluntary relationships (Bovens, 2010; Schillemans, 2011; Grandvoinnet et al., 2015). Table 2-1 summarises the different types of public accountability which have been discussed with their definitions and examples.

In this thesis, empirical research is undertaken on accountability affected by performance management systems. Therefore, it is necessary to construct measures of levels such as ordinal or scale for the above types of public accountability (Lindberg, 2013; Brandsma, 2014). However, the measurement instrument, criterion or approach, used for one type of accountability, may not be useful for gauging the level of another type (Lindberg, 2013). This may result from the fact that principals are fundamentally different in various types of public accountability, and the degree of control that principals exercise over agents varies (Lindberg, 2013). Therefore, it is important to develop a separate measurement instrument for each type of public accountability (Lindberg, 2013; Brandsma, 2014). This thesis employs different measurement instruments to measure the level of managerial and social accountability of public institutions to examine the effect of performance management systems.

Table 2-1: Summary of types of public accountability

Type of public accountability		Definition	Example
External		Public organisations give an account to external individuals or agencies	Local government is accountable to central government through inspection reports
Internal		Subordinates give an account to superiors within an organisation	Subordinates are accountable for their outputs or outcomes
Political		Public servants are accountable to their minister, who must give account to parliament	Politicians explain and justify their actions to the public
Managerial		Those with delegated authority are responsible for the efficient use of public resources	School teachers are held accountable for improving student learning outcomes
To whom	Legal	The legal justification of an actor's performance in relation to others	To Court of Justice
	Administrative	An administrative agency is accountable to its supervisory constituencies for the task assigned to it	To courts of audit or regulatory agencies
	Social	Public agencies give account to citizens or interest groups	To citizens and civil society
Who	Organisational	An organisation as a legal entity is to give account	Organisation as a corporate entity
	Hierarchical	Only top of organisation is required to give account externally	Minister or CEO
	Collective	Every member of organisation is called upon to give account for act of the organisation	Every member of organisation
Why	Vertical (mandatory)	A forum formally has power over an actor due to a hierarchical relationship between the forum and actor	Minister to parliament
	Horizontal (contractual)	A public agency is obliged to give account by a board of stakeholders	Non-government service provider to professional evaluation committee
	Diagonal (Voluntary)	An actor gives account to various stakeholders in society basically on a voluntary basis	Corporate executive to the press

2.2.4 Accountability deficit

Although accountability can function in many ways, there is the potential for an 'accountability deficit' when ministers decouple themselves from accountability by means of outsourcing and the establishment of executive agencies (Mulgan, 2014; Rock, 2020; Van Osch, De Ruiter and Yesilkagit, 2023). 'Accountability deficit' can be defined as an observable lack or deficiency of some substance or quality of accountability and an attitude of regret or disapproval towards such a lack or deficiency (Mulgan, 2014). The notion of an 'accountability deficit' has stimulated both conceptual analysis and empirical investigation of accountability (Mulgan, 2014; Rock, 2020). For example, Han (2020) examines the effect of accountability deficits on agency performance by using the three major phased components of accountability processes: (i) information provision, (ii) assessment and (iii) consequence (Brandsma and Schillemans, 2013). Han (2020) argues that although accountability systems might work well theoretically, they may work badly in practice, which means deficits in accountability mechanisms.

Many scholars suggest that the notion of 'accountability deficit' has been commonly linked with the absence of political control by political representatives (Bovens, Schillemans and Hart, 2008; Mulgan, 2014; Rock, 2020; Van Osch et al., 2023). For example, under the Next Steps initiative enacted by the UK government in 1988, agencies were established at arm's length from ministerial control and received funding and responsibility to implement particular government polices (Pollitt and Bouckaert, 2017). Subsequently, agencies became formally decoupled from their ministers and the ministers were able to avoid accountability for administrative decisions, handing the accountability over to the agencies and their chief executives (Bovens et al., 2008; Mulgan, 2014). Whilst ministers

are accountable to the members of parliament over their actions and policies, agencies lack exposure to political direction in response to public concern (Mulgan, 2014). This accountability deficit may be magnified as the number of agencies providing public services increases.

The concept of accountability deficit can be widened beyond deficit in political accountability to all the different types of public accountability where a deficit is found (Schillemans, 2011; Mulgan, 2014; Rock, 2020). For example, public agencies are not obliged to provide accounts of their performance for citizens, interest groups or mass media, therefore, the social accountability deficit can occur in public agencies (Schillemans, 2011). This thesis examines whether performance management systems mitigate the managerial and social accountability deficit of public agencies. The examination sheds light on research on 'accountability deficit' by exploring different types of 'accountability deficit' beyond deficit in political accountability (Rock, 2020).

2.2.5 Accountability in New Public Management

New Public Management (NPM) doctrine has placed great attention on the need to enhance public accountability (Stone, 1995; Boyne, 2002; Ospina et al., 2004; Lægreid, 2014). NPM has been identified as a holistic model of managerialism which has pervaded all aspects of public organizations since the 1980s (Osborne and McLaughlin, 2002; Van Dooren et al., 2015; Pollitt and Bouckaert, 2017). According to Hood (1991), NPM comprises seven doctrines: (i) hands-on professional management in the public sector, (ii) explicit standards and measures of performance, (iii) greater emphasis on output controls, (iv) shift to the disaggregation of units, (v) shift to greater competition, (vi) stress on private sector styles of management practice and (vii) stress on greater discipline and

parsimony in resource use. In addition, NPM is summarized with three key components: disaggregation, competition and incentivisation (Dunleavy, Margetts and Tinkler, 2005; Lapuente and Van de Walle, 2020). Disaggregation means a separation of provision through agencification of departmental units or contracting out production to non-profit organizations. Competition refers to marketization either with internal markets in the public sector or contracting out to the private sector. Incentivisation focuses on rewarding for specific performance, rather than diffuse performance, through a performance-related pay system.

NPM was first developed based on importing market-type mechanisms such as competitive tendering and performance-related pay into the public sector (Dunleavy et al., 2005). For example, a performance-related pay system has been adopted to improve performance in the public sector in Finland, the Netherlands, Sweden and the UK (Pollitt and Bouckaert, 2017; Lapuente and Van de Walle, 2020). However, the focus of NPM reforms has shifted to reinventing government within the context of a plural state and to the management of complex networks in public governance (Osborne and McLaughlin, 2002). NPM reforms intended to enhance efficiency of public organisations and accountability to customers for the quality of service by making the public sector more 'business-like' and giving managers more 'freedom to manage' (Power, 1997; Christensen and Lægreid, 2015; Lapuente and Van de Walle, 2020; Hammerschmid et al., 2019). As a result, the focus of accountability was changed from processes, compliance with rules and inputs to outputs and results under NPM reforms (Glynn and Murphy, 1996; Boyne et al., 2002; Lægreid, 2014).

Therefore, it is argued that NPM doctrines resulted in changes in public accountability (Stone, 1995; Boyne, 2002; Christensen and Lægreid, 2015). Under NPM reforms the focus in many parts of the public sector was transferred from political accountability to managerial accountability (Glynn and

Murphy, 1996; Boyne et al., 2002; Ospina et al., 2004; Lægreid, 2014). For example, agencies are required to receive a periodic formal external evaluation through performance indicators. Subsequently, NPM doctrines tends to replace vertical accountability with horizontal accountability such as social accountability to users of public services (Lægreid, 2014). For instance, public agencies and public administrators are supposed to be autonomous and focus on signals from their clients or customers (Lægreid, 2014). In spite of the change in public accountability since the NPM reforms, there is limited research on the effect of performance management systems on managerial and social accountability of public organisations (Jantz et al., 2015). This thesis aims to fill the gap in previous research.

Since the late 1990s, the agenda in public management reform appears to have shifted from NPM to other management techniques which preference networks, partnerships and digital-era governance (Dunleavy et al., 2005). However, it is important to recognize that NPM reforms are still prevalent in a number of countries including Australia, Italy, and New Zealand and remain a strong force for change in most public administrations (Willis et al., 2017; Bracci, Maran and Inglis, 2017; Pollitt and Bouckaert, 2017; Hammerschmid et al., 2019; Lapuente and Van de Walle, 2020). NPM reforms focus on the improvement of accountability through explicit standards of performance and output control, however, the relationship between performance management and accountability is still contested (Pollitt, 2011; Christensen and Lægreid, 2015). The key premise of NPM reforms was that performance management systems would make public organisations more accountable (Van De Walle and Cornelissen, 2014). For example, the 1993 Government Performance and Results Act in the US declared that "new performance-oriented regime would improve federal program effectiveness and public accountability by promoting a new focus on results, service quality and customer satisfaction". In contrast, many scholars argue that performance management has limited

capacity for strengthening accountability (Jantz et al., 2015; Van Dooren et al., 2015; Jones and Bouckaert, 2017). This is because practitioners and public agencies can manipulate the measurement of results and performance management adds complexity, ambiguity and the potential for conflict in public agencies (Bevan and Hood, 2006; Halligan, 2007; Jones and Bouckaert, 2017).

The UK is considered one of the originators and most vigorous proponents of the NPM (Pollitt and Dan, 2013). Therefore, this thesis draws on previous research on cases in the UK to explore the effect of performance management on accountability. Research into the National Health Service (NHS) in the UK highlights the debate about the relationship between performance management and accountability. Hauck and Street (2007), for example, argue that the strong performance management regime operating in English hospitals has contributed to a higher level of accountability than in Welsh hospitals in terms of waiting times for hospital admission. The impact of performance management on accountability in the public sector is also found in the research on the effect of annual publication of the school performance tables in the UK (Bradley, Crouchley, Millington and Taylor, 2000). Using a data set covering all publicly funded secondary schools in England for 1992-1998, Bradley et al. (2000) argue that since the school performance tables were first published in 1992 schools have achieved better exam results by competing with top-performing schools.

In contrast, Bevan (2006) argues that whilst the new UK Labour Government of 1997 introduced a system of NHS regulation, the percentage of significantly poor financial performance rose from three percent of trusts in 2001 to nearly 30 percent in 2005. The system was designed to measure annual performance against targets and give each organisation a single summarised score. Pollitt (2011) suggests that the NHS case gives little support for the argument that performance management

improves accountability. This is because, for the NHS, there are a lack of interest and a general misunderstanding concerning performance measurement and, gaming, together with a lack of public trust in the official figures presented (Pollitt, 2011). In addition, the relationship between performance management and accountability can be explored in Norway and Germany where the reforms of welfare and employment administration have combined elements of NPM since the 2000s (Jantz et al., 2015). The introduction of performance management in Norway and Germany has not contributed to accountability for learning in local offices (Jantz et al., 2015). This is because the focus of the performance management is on "What went wrong and who should be blamed" instead of on learning aspects and goals in the performance management (Jantz et al., 2015). To sum up the discussion so far, considerable research has been undertaken on the relationship between performance management affected by NPM reforms and accountability. However, there is limited research on the relationship between performance management and different types of accountability (Jantz et al., 2015; Christensen and Lægreid, 2015). This thesis contributes to literature on the relationship between performance management and accountability by examining the effect of performance management systems on managerial and social accountability of public organisations.

2.3 Performance-based incentive schemes and performance indicators

2.3.1 Agenda of performance management research

Performance management systems can be defined as formal mechanisms, processes, systems and networks for conveying key objectives set by managers, assisting ongoing management and

facilitating organisational learning and change (Ferreira and Otley, 2009). After the global economic disturbances of the 1970s, the belief that governments had become 'overloaded' and inefficient spread (Propper and Wilson, 2003; Van Dooren et al., 2015; Pollitt and Bouckaert, 2017). Consequently, the desire to make government to increase efficiency arose, and hence many countries including Australia, New Zealand, the UK and the US developed performance management and introduced more competition into the public sector (Pollitt and Bouckaert, 2017). As a result, performance management systems have been prevalent in the public sector (Pollitt, 2013; Jantz, Christensen and Lægreid, 2015; Taylor, 2021). Moreover, the financial crisis for 2007-2008 intensified the need to make best use of reduced resources in the public sector (Arnaboldi, Lapsley and Steccolini, 2015). The global finance crisis has accentuated the effect of performance management in the public sector (Arnaboldi et al., 2015).

Performance management research has covered the design and use of performance management systems (Ferreira and Otley, 2009; Bracci et al., 2017). Research into the design of such systems is concerned with features such as reward systems, key performance indicators, and target-setting when implementing performance management systems (Agostino and Arnaboldi, 2012; Bracci et al., 2017). The research on use of the systems investigates how the information collected by performance management systems is used to influence the behaviour of people within an organisation (Bracci et al., 2017). This thesis explores the effect of reward systems, performance indicators and target-setting on public accountability and performance. Therefore, the thesis sheds light on the agenda of design and use of performance management systems.

Reviewing research on performance management in the period 1980-2017, Schleicher et al. (2018)

suggest seven key agendas: (i) setting performance expectations, (ii) observing employee performance, (iii) integrating performance information, (iv) the rendering of a formal summative performance evaluation, (v) generating and delivering performance feedback, (vi) the formal performance review meeting, and (vii) performance coaching. Among the seven agendas, considerable research is concentrated on the rendering of a formal summative performance evaluation and generating and delivering performance feedback (Schleicher et al., 2018). Regarding the agenda of effect of a formal summative performance evaluation, a handful of empirical research has been undertaken (Schleicher et al., 2018). Whilst some scholars argue that a formal performance evaluation can have a positive effect on the job satisfaction of employees and on product quality (Kaya, Koc and Topcu, 2010; Nayyab et al., 2011), other scholars argue that it has no effect on satisfaction and commitment of employees (Pullin and Haidar, 2003). With respect to the agenda of performance feedback, scholars suggest that organisations implementing feedback intervention see improved performance and job satisfaction of employees (Pampino et al., 2004; Kaya et al., 2010). This thesis examines how the summative performance evaluation and the feedback of financial and non-financial incentives through performance management systems affect accountability and performance of public organisations. Therefore, the thesis is related to the two key agendas of performance management research.

2.3.2 Definition of performance

Before discussing performance-based incentive schemes and performance indicators in performance management systems, it is necessary to define performance. As with the term of accountability, performance can mean many things (Talbot, 2010; Van Dooren et al., 2015). Dubnick (2005, p. 392)

defines performance as a behaviour motivated or guided by some intent or purpose and makes a typology of performance, relying on two aspects of performance; "the quality of the actions being performed and the quality of what has been achieved as a result of those actions" (see Table 2-2).

Table 2-2: Types of performance

		Level of attention on quality of performance achievement	
		Low	High
Level of attention on quality of performance actions	High	Competence	Productivity
	Low	Production	Results

Source: Adapted from Dubnick (2005, p. 392)

While 'production performance' focuses on the process of production actions, 'competence performance' highlights the quality of performance actions (Dubnick, 2005). 'Production performance' breaks any job down to its basic tasks and assesses whether the tasks are being done appropriately. However, 'competence performance' considers the application of higher levels of knowledge and skill when actors conduct tasks (Dubnick, 2005; Hood and Lodge, 2004). For example, 'production performance' is for anyone to cook according to a recipe, however, 'competence performance' is for a professional chef to cook with high level of knowledge (Dubnick, 2005). 'Results performance' focuses attention on what is produced, as opposed to the process of production (Dubnick, 2005). 'Results performance' refers to both outputs and outcomes and is, therefore, the most prevalent type of performance in the literature on the NPM principle which emphasises performance through measurement of outputs (Van Dooren et al., 2015; Dubnick, 2005; Pollitt and Bouckaert, 2017). In contrast to 'results performance', 'productivity performance' concentrates on the quality of action as well as what is achieved (Dubnick, 2005). As explained earlier, managerial accountability focuses on outputs or results which are made by using public resources (Jantz et al.,

2015), and hence managerial accountability is connected to 'results performance' which focuses on output and outcomes. Social accountability is linked to improvement of public service quality by public agencies (Grandvoinnet et al., 2015; Brinkerhoff and Wetterberg, 2016; Hickey and King, 2016), and hence social accountability is related to 'results performance' which highlights outcome as well as output. Considering the link between 'results performance' and managerial and social accountability, performance in the public sector in this thesis refers to 'results performance'.

Multiple dimensions of performance (e.g. efficiency, economy, effectiveness) are discussed in the literature, with assessments based on inputs, outputs and outcomes (Carter et al., 1995; Talbot, 2010; Van Dooren et al., 2015). 'Inputs' generally mean resources such as labour, physical resources and IT services consumed in producing public sector service (Talbot, 2010). While 'outputs' are defined as actual goods or services produced by organizations, 'outcomes' are the actual changes to health, wealth, happiness and other desired social results (Carter et al., 1995). 'Efficiency' refers to the relationship between inputs and outputs, 'economy' means reducing the costs of inputs, and 'effectiveness' is defined as the extent to which outputs achieve desired outcomes (Carter et al., 1995; Talbot, 2010).

Skelcher (2008) conceptualises performance in three ways. Firstly, 'organisational performance' is defined as the substantive outputs and outcomes of a public organisation. The 'organisational performance' is constructed with reference to the metrics of effective implementation, productivity, service outcomes and client satisfaction (Skelcher, 2008). Secondly, 'democratic performance' refers to the extent to which a public organization can demonstrate mechanisms for legitimacy, consent and accountability. Legitimacy means political validation of institutional authority. Consent is relevant to

the capacity of citizens to agree organisational action and accountability refers to the explanation of action linked to a renewal of mandate by its principal (Skelcher, 2008). Thirdly, 'system performance' means the extent to which a system of public organisations is integrated. The 'system performance' is relevant to the overall functioning of a government system and its capacity to coordinate actions (Skelcher, 2008).

As explained earlier, public agencies were established at arm's length from ministerial control and since the 1980s have provided public services in many countries in accordance with the NPM principle of disaggregation (Pollitt and Bouckaert, 2017; Dunleavy et al., 2005). Public agencies lack exposure to political direction in response to public concern and hence it is important to hold public agencies accountable for their public service performance through performance management systems. Therefore, it is necessary to review the meaning of public service performance to explore the effect of performance management systems on accountability of public agencies. Public service performance is judged by multiple stakeholders such as consumers, taxpayers, staff and politicians (Boyne, 2004; Andrews, 2011). Although each stakeholder may use different criteria to judge public service performance (Boyne, 2004), there are three key elements of public service performance that all stakeholders value (Andrews, 2011). The three elements are the speed of public service delivery, efficiency of public service providers, and the satisfaction level of public service users (Boyne, 2004; Andrews, 2011). It is against these three elements of public service performance that the next sections discuss the concepts of performance-based incentive schemes and performance indicators in performance management systems.

2.3.3 Performance-based incentive schemes in performance management systems

Accountability processes consist of three phases: information, discussion and sanction (Brandsma and Schillemans, 2013; Jantz et al., 2015; Han, 2020). The 'information phase' is when agents provide an account of their conduct and behaviour to their principals (Jantz et al., 2015). The 'discussion phase' is when principals assess the performance information provided by their agents and the agents have an opportunity to justify their performance (Brandsma and Schillemans, 2013; Jantz et al., 2015). The 'sanction phase' is when principals pass a judgement on the behaviours of their agents and reward or punish them based on that judgement (Jantz et al., 2015). The 'sanction phase' is connected to performance-based incentive schemes in performance management systems due to the reward and punishment (Ferreira and Otley, 2009; Jantz et al., 2015). Performance-based incentive schemes such as performance-related pay and public league tables have been adopted under performance management systems in many countries in accordance with the NPM principle of incentivisation (Perry et al., 2009; Lapuente and Van de Walle, 2020). As performance-based incentive schemes can significantly influence the behaviours of employees and are an important condition of performance management systems (Burgess and Ratto, 2003; Ferreira and Otley, 2009; Lockwood and Porcelli, 2013; Mayumana et al., 2017), the thesis examines the effect of performance-based incentive schemes on public accountability.

The performance-based incentive schemes in the public sector can be divided into financial and non-financial incentive schemes (Ferreira and Otley, 2009; Jalava, Joensen and Pellas, 2015; Kefay and Kero; 2019; Burgess et al., 2021). Financial incentives involve a transfer of monetary values or equivalents such as salary increase, performance-related pay or housing (Mathauer and Imhoff, 2006;

Ferreira and Otley, 2009). Among them, performance-related pay has been adopted extensively to improve the performance of employees in the public sector (Burgess and Ratto, 2003; Propper and Wilson, 2003; Weibel, Rost and Osterloh, 2010; Belle and Cantarelli, 2015). For example, Brazil instituted a major incentive reform to improve tax enforcement in 1989 (Kahn, Silva and Ziliak, 2001). The essence of the reform was to provide tax collectors with monetary compensation according to their performance in finding and collecting taxes from tax evaders (Kahn et al., 2001). Using a panel data set for 1987-1992, Kahn et al. (2001) suggest that fines per inspection increased by about 75% after the incentive reform and there was substantial heterogeneity in the effect of the reform across tax regions. In addition, Mayumana et al. (2017) argue that performance-related pay introduced in the Tanzanian health system in 2011 contributed to improvement of internal accountability between managers and providers through shared goals. However, many scholars argue that performance-related pay in the public sector is ineffective in improving employees' accountability and performance (Perry et al., 2009; Belle and Cantarelli, 2015). This is because financial incentives are too small to be valued and employees perceive the implementation of performance-related pay to be unfair due to lack of transparency (Perry et al., 2009). Considering the debate about the impact of financial incentive schemes on public accountability and performance, this thesis examines the effectiveness of financial incentive schemes in performance management systems.

Non-financial incentives refer to any means of incentives that do not directly involve money or transfers of monetary values (Mathauer and Imhoff, 2006; Kefay and Kero; 2019). Job promotion, recognition, training, rank and awards are examples of non-financial incentive schemes (Kefay and Kero; 2019; Jalava et al., 2015). For instance, a published school league table functions as a non-financial incentive scheme because a high league position rewards teachers with a short run 'prestige' benefit and a long run 'working conditions' benefit if good students can choose schools in a high

league position and teachers prefer teaching good students (Propper and Wilson, 2003). Non-financial incentives can be the key to improving employees' motivation, job satisfaction and performance (Franco et al., 2004; Kefay and Kero; 2019). However, limited research has been undertaken on the impact of non-financial incentives on the motivation of employees in the public sector and students (Franco et al., 2004; Jalava et al., 2015). This thesis fills the gap in the literature on non-financial incentives by examining the effectiveness of non-financial incentive schemes in performance management systems.

2.3.4 The effect of performance-based incentive schemes on motivation

It is important to consider employees' motivation when designing performance-based incentive schemes in performance management systems (Sansone and Harackiewicz, 2000; Franco et al., 2004; Jalava et al., 2015; Lee, 2021). Motivation is defined as an individual's degree of willingness to exert and maintain an effort towards goals and is largely divided into extrinsic and intrinsic motivation (Franco et al., 2004; Ryan and Deci, 2017). Extrinsic motivation is represented by behaviours that are instrumental for some consequences such as an external reward, social approval and the attainment of a valued outcome (Ryan and Deci, 2017). Many studies have focused on financial incentives as means to enhance extrinsic motivation (Jalava et al., 2015). Intrinsic motivation refers to motivation that is driven by an interest in, or enjoyment of a task itself (Jalava et al., 2015). Non-financial incentives such as verbal rewards, trophies and ranking significantly enhance intrinsic motivation because non-financial incentives give employees non-material benefits such as social recognition and improved self-esteem (Sansone and Harackiewicz, 2000; Mathauer and Imhoff, 2006; Jalava et al., 2015).

Abundant research has been undertaken to explore how financial incentives affect intrinsic motivation (Frey and Oberholzer-Gee, 1997; Weibel et al., 2010; Belle and Cantarelli, 2015; Greener, 2019; Lee, 2021). For example, Lee (2021) argues that financial rewards can motivate public employees even if they have strong intrinsic motivation. This is because the practice of financial incentives increases the possibility of employees to be recognised and respected as high performers, therefore, intrinsic motivation of public employees increases by the financial incentives (Lee, 2021). However, many scholars argue that there is a negative interaction between financial incentives and intrinsic motivation under specific conditions (Frey and Oberholzer-Gee, 1997; Belle and Cantarelli, 2015; Frey and Gallus, 2016). It is argued in the literature, that financial incentives crowd out intrinsic motivation if employees perceive financial incentives to control them and are to a high degree intrinsically motivated (Weibel et al., 2010; Frey and Gallus, 2016).

On the other hand, non-financial incentives play an important role in enhancing intrinsic motivation and hence result in improvement of accountability and performance (Burgess and Ratto, 2003; Mathauer and Imhoff, 2006; Frey and Gallus, 2016; Kefay and Kero; 2019). For example, it is found that non-financial incentives of prestigious awards in academia considerably raise the productivity of academics (Frey and Gallus, 2016). This positive relationship between the non-financial incentive and increased productivity results from the fact that the prestigious award creates honours valuable to the recipients because the award reveals the recipients' talent and commitment (Frey and Gallus, 2016). Therefore, intrinsic motivation provides a rationale for the use of non-financial incentive schemes (Burgess and Ratto, 2003). This thesis will consider the intrinsic motivation of employees in the public sector to examine the effectiveness of non-financial incentive schemes in performance

management systems.

2.3.5 Definition and types of performance indicator

Performance indicators play an important role in improving the public accountability of ministers to parliament and the public for public services (Carter et al., 1995; Bird et al., 2005; Van Dooren et al., 2015). This importance of performance indicators is one of the reasons why reform initiatives such as the Financial Management Initiative and the Next Steps led to performance indicator systems for most public services (Carter et al., 1995; Van Dooren et al., 2015). It is virtually impossible to measure fully any aspect of performance, because the cost of trying to gauge every possible aspect of performance would probably be prohibitive (Talbot, 2010). For example, over 600 possible areas of measurement were established by a team in Careers Scotland which was a relatively small agency (Talbot, 2010). The existence of so many areas of measurement, even in small institutions, invokes the call for performance indicators and selection of performance indicators (Talbot, 2010; Van Dooren et al., 2015).

Many scholars have sought to define performance indicators (Carter et al., 1995; Pidd, 2008; Talbot, 2010). For example, performance indicators are defined as weighted aggregations of different aspects of performance (Pidd, 2008). According to this definition, performance indicators are used to summarize the performance of a public service organization in terms of outputs and outcomes (Pidd, 2008). Talbot (2010, p. 39) suggests that "performance indicators are - as the name suggests - indicators of performance rather than direct measures of absolute performance". Therefore, performance indicators usually rely on a sampling approach or a selective approach rather than a full

approach (Talbot, 2010). As performance indicators summarize different aspects of performance with a sampling approach, the number of performance indicators should be as small as possible, and that performance indicators can explicitly show 'results performance' (Van Dooren et al., 2015). For example, in 2004 the European Commission proposed reducing 42 indicators covering general economic and environmental performance for the next decade down to 14 key indicators (Van Dooren et al., 2015). In order to reduce the number of performance indicators in performance management systems, only performance indicators that are decision-relevant and successfully survive cost benefit analysis should be used (Johnsen, 2005). In other words, if a performance indicator adds value to its users beyond the cost of developing and using the performance indicator, then the performance indicator would deserve to be used.

The use of performance indicators falls into three categories: prescriptive, descriptive and proscriptive (Carter et al., 1995). Firstly, prescriptive performance indicators are used to monitor progress towards an organization's objectives or targets. Carter et al. (1995, p. 49) state that "The ideal prescriptive indicator operates like a dial, providing a precise measure of inputs, outputs and outcomes based on a clear understanding of what good and bad performance entails". Secondly, descriptive performance indicators simply record change, providing a map rather than a prescribed route. According to Carter et al. (1995), performance indicators tend to be used descriptively in practice, because they do not give complete answers but provide an incomplete picture. While implicit assumption of prescriptive performance indicators is that standards of performance are unambiguous, that of descriptive performance indicator is that performance is a contestable notion (Carter et al., 1995). Thirdly, proscriptive performance indicators present, not targets, but things which simply should not happen in a well-run organization. Proscriptive performance indicators operate like an alarm-bell, providing warnings of events which cannot be tolerated in a well-run organization (Carter et al., 1995). Among

these three kinds of performance indicators, prescriptive performance indicators appear to be appropriate for improving political and managerial accountability of public organisations. This is because prescriptive performance indicators provide public organizations with explicit targets and performance against the targets, which can be accounted for and justified to parliament and the public. The link of prescriptive performance indicators with political and managerial accountability is supported by the suggestion that government policy for public service uses prescriptive performance indicators to determine whether policy objectives are being met (Bird et al., 2005).

As mentioned earlier, the number of possible dimensions of performance can increase the amount of measurement demanded, and one effect of such increase may be 'information overload' (Talbot, 2010). For example, in Japan, about 10,000 evaluation and performance reports containing multiple pieces of information have been published annually in accordance with the Government Policy Evaluation Act (Talbot, 2010). 'Information overload' has led to a tendency to organize and simplify performance data into 'composite indicators' (Talbot, 2010; Van Dooren et al., 2015). The 'composite indicators' are used to compare and rank country performance in areas such as industrial competitiveness, sustainable development and innovation (Talbot, 2010). For example, the World Bank's 'Worldwide Governance Indicators (WGI)' are an exemplar of 'composite indicators', which compare 212 countries on a wide range of measures to arrive at a single score rating (Talbot, 2010). The OECD formulated several criteria for constructing 'composite indicators'; (i) exposed methodological choice in weighting and aggregation, (ii) comparable indicators, and (iii) different approaches for imputing missing values exposed (OECD, 2009; Van Dooren et al., 2015). In addition, 'star ratings' have been used as 'composite indicators' to compare and rank hospitals and schools in the UK (Talbot, 2010). The composite indicator of 'star rating' can create competition between public service organisations (Talbot, 2010). As 'composite indicators' have the advantages of easy

interpretation, tracking over time and comparisons for public service organisations (Talbot, 2010), they can be used in performance management systems in the public sector.

According to previous research on performance in the public sector, it is more difficult to measure or assess public sector performance for some characteristics such as multiple goals and lack of competition compared with the private sector (Dixit, 2002; Van Thiel and Leeuw, 2002). Public agencies have several principals or forums such as central or local government, parliament and the public (Dixit, 2002). Therefore, public agencies have multiple, potentially conflicting goals (Carter et al., 1995; Andrews, 2011). Moreover, most public agencies used to be monopolies, hence they lack the background of competition (Dixit, 2002). According to Carter et al. (1995), competition can lead to a variety of performance indicators strategies, and these can vary from concentrating on performance indicators measuring productive efficiency to those measuring effectiveness of customer service. However, lack of competition in public agencies hinders the variety of performance indicators because comparing efficiency and effectiveness of public agencies through performance indicators is difficult (Carter et al., 1995). Considering the limitation on use of performance indicators in the public sector (Dixit, 2002; Van Dooren et al., 2015), it is important to explore the criteria for good performance indicators.

2.3.6 Criteria for good performance indicators

Many scholars have sought to find criteria for good performance indicators in order to make performance management systems effective for improving accountability in the public sector (Carter et al., 1995; Bird et al., 2005; Van Dooren et al., 2015). According to Carter et al. (1995), there are

three criteria for good performance indicators. The first criterion is that a performance indicator should measure performance owned by the organization. In other words, it should not measure performance determined by external actors or environmental factors. The second criterion is that a performance indicator should be relevant to the objectives of organizations. The third criterion is that a performance indicator should not be susceptible to manipulation by employees to be assessed. Bird et al. (2005) suggest criteria of good performance indicators more specifically. Firstly, a performance indicator should be consistent over time. When changes regarding performance indicators are necessary in exceptional circumstances, an impact assessment or a new baseline needs to be made. Secondly, a performance indicator should be straightforward to interpret. In other words, a performance indicator should avoid ambiguity as to whether the performance being monitored has improved or deteriorated. Thirdly, a survey-based indicator such as patient satisfaction should use a shared methodology and common questions across the organizations being assessed to enable a fair comparison between the organisations. Fourthly, a performance indicator needs to conform to international standards if these exist because the international standards present shared targets across the globe. Lastly, a performance indicator should not impose an undue burden in terms of cost on those providing the information.

Van Dooren et al. (2015) propose that the main qualities of good performance indicators are that they are sensitive to change and are precisely defined and actionable. For example, regarding a performance indicator sensitive to change; if the performance indicator for customer satisfaction relies on a "yes" or "no" response, its measurement fails to distinguish between someone being just satisfied and another being very satisfied. In order to avoid ambiguous understanding amongst employees assessed in an organization, detailed indicator descriptions are required. An actionable indicator means that it guides employees in an organization towards actions to improve performance.

This thesis uses the criteria for good performance indicators in order to examine how the evaluation processes of performance management systems affect the accountability of public institutions. Despite the importance of performance indicators in performance management systems, limited research has been undertaken on the effect of performance indicators on accountability of public organisations under performance management systems (Carter et al., 1995; Bird et al., 2005; Van Dooren et al., 2015). This thesis fills the gap in literature on performance indicators in performance management systems.

2.3.7 Guidance on performance management systems in the public sector

This thesis uses a guidance on performance management systems in the public sector based on performance measurement to examine how the processes of performance management systems affect accountability and performance of public organisations (Fryer, Antony and Ogden, 2009; Van Dooren et al., 2015). Performance measurement involves quantifying input, output, or level of activity involved in an event, and is a crucial element of performance management system (Fryer, Antony and Ogden, 2009). Performance measurement consists of five steps: (i) prioritizing measurement efforts, (ii) selection of indicators, (iii) data collection, (iv) analysis, and (v) reporting (Van Dooren et al., 2015).

The guidance for the first step is to understand the essence of an organization and its representation via an organizational chart depicting divisions of tasks and accountability (Van Dooren et al., 2015). With respect to the second step, scholars suggest that a cost-benefit analysis is necessary for each indicator and the indicators that survive the analysis and are decision-relevant should be used

(Johnsen, A., 2005; Fryer, Antony and Ogden, 2009). The guidance for the third step is to use an existing administrative registration system as a default data source and consider other data sources when the system cannot provide necessary data (Van Dooren et al., 2015). For example, a survey of citizens or customers can be used as an alternative data source to obtain outcome information (Van Dooren et al., 2015). The fourth step is to transform data into information that can lead to decisions (Fryer, Antony and Ogden, 2009). The guidance for the fourth step is to analyse data with targets (Van Dooren et al., 2015). Targets can be set based on dimensions of time, other organisations within and outside the sector and other countries (Van Dooren et al., 2015). For example, target can be to do at least as good as last year on dimension of time and can be the average performance of other organisations within the sector (Van Dooren et al., 2015). After analysis of data, public league tables are created and used as information about hospitals, local councils and schools in the public sector (Jacobs and Goddard, 2007; Fryer, Antony and Ogden, 2009). The guidance for the last step is related to target groups of information and right formats for information (Van Dooren et al., 2015). If the target group for the information is interest groups or other government, the reporting has to be specialised and detailed (Van Dooren et al., 2015). In addition, annual reporting is a good instrument for reporting to interest groups (Van Dooren et al., 2015). The guidance for all five steps will be used to explore the effect of performance indicators, target-setting and management performance reports in performance management systems on public accountability and performance.

2.4 Strategic behaviour and gaming

Even with sound performance indicators in place, over time performance management systems tend to become vulnerable to strategic behaviour and gaming of actors under the systems, and hence have a perverse effect on accountability and performance of actors (Hood, 2006; Van Dooren et al., 2015; Greener, 2019; Taylor, 2021). For example, a 'performance target paradox' can occur, which means that a performance indicator fails to discriminate between good and bad performers over time because actors adapt their performance for target of performance indicator, resulting in performance clustering around the target (Van Dooren et al., 2015). Considering the negative relationship between strategic behaviour or gaming and accountability or performance of actors, it is important to explore the strategic behaviour and gaming of public organisations to develop performance management systems that improve accountability and performance (Terman and Yang, 2016; Lu, Yang and Thomas, 2021). This section discusses principal-agent theory which presents a useful framework for analysing strategic behaviour and gaming (Terman and Yang, 2016), and then reviews various strategic behaviour and gaming under performance management systems.

2.4.1 Overview of the principal-agent theory

Principal-agent theory refers to formal models in which an actor identified as an agent undertakes an action on behalf of another actor identified as a principal and the principal makes decisions that affect the incentives of the agent to take any of various possible actions (Gailmard, 2014; Bertelli, 2012). As the agent's incentives constitute a contract, principal-agent theory is a specific area of contract

theory (Gailmard, 2014). Principal-agent theory is much developed in relation to insurance and the regulation of commercial businesses, however, the theory relates to non-marketed service and hence it has been applied to the public sector as well as the private sector (Levacic, 2009). In the public sector, principal-agent theory is widely used when designing an optimal government contract in the presence of moral hazard which is the tendency for agents to engage in undesirable behaviours (Bertelli, 2012; Song, Yu and Sun, 2020; Mankiw and Taylor, 2020). A main issue for principal-agent theory in the public sector is what structure of rewards and sanctions best incentivises public organisations to produce outputs desired by government and the public (Levacic, 2009; Gailmard, 2014). Principal-agent theory also covers issues of accountability because the theory examines the ways in which principals hold agents accountable for their performance (Levacic, 2009; Gailmard, 2014). Therefore, principal-agent theory can provide a useful theoretical framework for analysing how well rewards and sanctions work to incentivise public organisations to improve accountability and performance.

A principal-agent relationship occurs whenever a person or organisation contracts another person or organisation to perform services or supply goods (Levacic, 2009). A principal is someone or some organisation who wants something to be done, but cannot perform the task themselves due to a lack of expertise or time (Bertelli, 2012). An agent is someone or some organisation who performs services or supplies products on behalf of a principal (Bertelli, 2012). While a principal would like an agent to commit a great deal of time and effort in carrying out specific tasks, there is always the risk that the agent may not act in the principal's best interests (Sloman, Garratt and Guest, 2020). For example, governments encourage public agencies to improve the quality of public services, however, the public agencies might prioritise their employees' interests.

The problem for principals known as the 'principal-agent problem' is caused by (i) asymmetric information where agents have more information than the principals and (ii) goal divergence where principals and agents have a conflict of interests (Sloman, et al., 2020; Bertelli, 2012; Song et al., 2020). The principal cannot directly observe specific action choice of the agent under asymmetric information, and hence there is the possibility for the agent to take opportunistic practices of deceit and fraud in goal divergence between the principal and the agent (Song et al., 2020). In other words, the agent can take advantage of asymmetric information to reduce its effort and pursue the maximization of its own interest (Song et al., 2020). The 'principal-agent problem' is related to strategic behaviour and gaming of the agents under performance management systems. For example, the agents may make their targets lower than their abilities and decrease efforts to the level of targets to achieve the targets more easily after a year under performance management systems.

According to principal-agent theory, principals can try to address the 'principal-agent problem' through two elements (Sloman et al., 2020; Bertelli, 2012). The first element is that principals have a way to monitor the performance of agents (Sloman et al., 2020). The second element is that principals design a set of incentives to compel agents to behave for the principals' interests (Bertelli, 2012; Sloman et al., 2020; Levacic, 2009). In relation to the incentive mechanism, whether agents respond to the incentives is ultimately an empirical question (Bertelli, 2012).

2.4.2 Strategic behaviour and gaming under performance management systems

Performance management systems are based on the assumption that they have the capacity to improve accountability and performance in the public sector (Jantz et al., 2015; Gerrish, 2016; Han 2020). However, empirical studies on performance management systems show their vulnerabilities to dysfunctional behavioural effects in the form of strategic behaviour and gaming surrounding performance management systems (Hood, 2006; Heinrich, 2007; Mizrahi and Minchuk, 2019b; Taylor, 2021). According to Hood (2006), at least three major types of strategic behaviour and gaming surrounding performance management systems exist. The first type is ratchet effects that refer to the expected tendency of target setters to fix next year's targets as an incremental advance over last year's results (Hood, 2006). Ratchet effects occur when the target setters or those having to achieve targets intentionally set lower targets than their abilities to achieve the targets easily (Taylor, 2021). The second type is threshold effects where a uniform target applying to all units in a performance management system provides no incentive to excellence (Hood, 2006). Threshold effects provide those who are performing above the uniform target with a perverse incentive to allow their performance to deteriorate to a uniform target level (Taylor, 2021). The third type is output distortion or data manipulation (Hood, 2006; Taylor, 2021). For example, according to Bohte and Meier (2000), school districts cheat by liberally exempting certain students from exams to raise the overall pass rate of the district.

Strategic behaviour and gaming occur during performance monitoring, collecting and reporting of performance data (Taylor, 2021). Firstly, when performance data is being monitored, organisations devote more effort to a performance dimension that is being measured at the expense of an unmeasured dimension (effort substitution) (Bevan and Hood, 2006). Secondly, organisations are likely to select only samples that lead to a positive evaluation when performance data is being collected (biasing sample) (Bohte and Meier, 2000). Thirdly, organisations can manipulate reporting

timing to secure a favourable outcome when they report performance data (Taylor, 2021). For example, the US Job Training Partnership Act training centres delayed reporting unfavourable outcomes and reported these outcomes when they faced the minimum risk of losing their awards (Courty and Marschke, 2007). To sum up, strategic behaviour and gaming can occur across all stages of the performance management process from target-setting to performance data reporting (Taylor, 2021).

Principal-agent theory can explain the strategic behaviour and gaming via two concepts: (i) moral hazard or hidden action and (ii) adverse selection or hidden information problem (Dionne and Harrington, 1992; Bertelli, 2012; Dunn, 2013; Pindyck and Rubinfeld, 2018). Moral hazard refers to the tendency of an agent who is imperfectly monitored to engage in dishonest or undesirable behaviour (Mankiw and Taylor, 2020). Adverse selection means that an agent has information that a principal cannot have, and so the information can be used for the agent's private benefit (Bertelli, 2012). This thesis contributes to literature on strategic behaviour and gaming by analysing actions of agents under performance management systems using the two concepts of moral hazard and adverse selection.

2.4.3 Screening and signalling

In order to address the problems of moral hazard and adverse selection, principal-agent theory presents the two related theoretical processes of screening and signalling. While screening means an action taken by an uninformed principal to determine the information possessed by an informed agent, signalling signifies an action taken by an informed agent to send information to a less informed

principal (Perloff, 2018, p. 656). As Boyne et al. (2002, p. 696) suggest, "Information is essential for the effective operation of both internal and external accountability", using adequate information through screening and signalling is important to improve public accountability and performance through performance management systems. This thesis explores how screening and signalling work in practice to mitigate the strategic behaviour and gaming of public organisations under performance management systems.

2.5 Conclusion

This chapter has reviewed literature on public accountability, agenda of performance management research, performance-based incentive schemes, performance indicators and strategic behaviour or gaming. The focus in many types of public accountability was transferred from political and vertical accountability to managerial and social accountability under NPM reforms (Boyne et al., 2002; Ospina et al., 2004; Lægreid, 2014). Managerial accountability emphasises outputs or results created through public resources and social accountability is linked to the improvement of public service quality (Jantz et al., 2015; Grandvoinnet et al., 2015). Therefore, managerial and social accountability are connected to 'results performance' which focuses on output and outcome. This thesis will examine the effect of performance management systems on managerial and social accountability of public organisations in order to explore the impact of performance management systems on 'result performance'.

The key agendas of performance management research are the design and use of performance

management systems (Bracci et al., 2017; Schleicher et al., 2018). Performance-based incentive schemes such as financial and non-financial incentives and performance indicators are design features of performance management systems (Agostino and Arnaboldi, 2012). The effectiveness of financial and non-financial incentive schemes is related to how performance management systems influence the behaviours of employees in public organisations (Bracci et al., 2017). This thesis will examine the effect of financial and non-financial incentive schemes and performance indicators on accountability and performance of public organisations. The examination sheds light on how conditions and processes of performance management systems affect public accountability.

Principal-agent theory is a useful theoretical framework for analysing how rewards and sanctions under performance management systems incentivise public organisations to improve accountability (Bertelli, 2012). Many strategic behaviour and gaming under performance management systems can be explored via principal-agent theory's concepts of moral hazard and adverse selection. Principal-agent theory provides screening and signalling as methods to mitigate the problems of moral hazard and adverse selection. This thesis will explore strategic behaviour and gaming of public organisations under performance management systems through the theoretical framework of principal-agent theory. The exploration contributes to literature on the influence of performance management systems on strategic behaviour and gaming.

Figure 2-1 shows the connection of public accountability, performance management systems and strategic behaviour or gaming. Many countries have tried to improve accountability and performance in the public sector through performance management systems including performance-based incentive schemes, performance indicators and targets. However, public organisations may take

strategic behaviour and gaming under the performance management systems due to asymmetric information and goal divergence.

Figure 2-1: The connection between main concepts

< Purpose >

Public accountability ↑

(Managerial & social accountability)

< Method >

Performance management systems

(Performance-based incentive schemes, performance indicators, targets)

< Dysfunction >

Strategic behaviours & gaming

(Ratchet effects, threshold effects, output distortion, etc.)
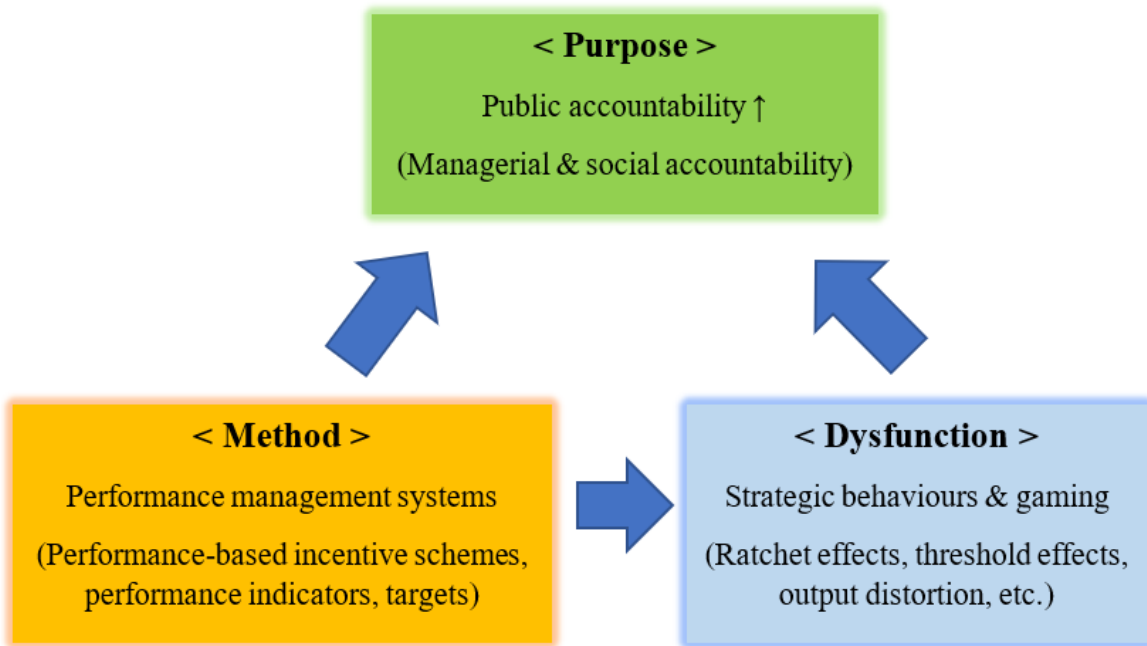
Source: Author's construction

Chapter 3 Performance management system and accountability in South Korea

3.1 Introduction

This chapter examines the key international guidelines of performance management systems in the public sector, which have been presented by the OECD since 2005. Some best practices in accordance with the key international guidelines are summarised. Subsequently, the chapter discusses different countries' performance management systems in the public sector in terms of these guidelines. Based on this discussion, this chapter presents the differences between the South Korean Public Institution Management Evaluation System (PIMES) and other countries' systems. It discusses how the PIMES follows the international guidelines and the study of PIMES as a case fills a gap in literature. The South Korean political context including the history of its democracy, political system and governance is discussed to enhance the understanding of PIMES. Moreover, this chapter explains how the principles of NPM affected the South Korean performance management systems and the PIMES.

An overview of PIMES is provided across five sub-sections. The first sub-section presents the definition and functions of PIMES. Although there are many functions in the PIMES according to the Ministry of Economy and Finance (MOEF) (2011), it is necessary to study how these functions work in practice. The second sub-section explains the subjects of PIMES evaluation and the scope of public

services covered in the PIMES evaluation. The third sub-section reviews the history of PIMES in South Korea since 1984 when it was first implemented. In the fourth sub-section, the evaluation process in the PIMES is introduced. The evaluation process consists of 3 steps: preparing for the PIMES evaluation, implementing businesses and conducting evaluation on a 3 year basis. The last sub-section discusses the performance indicators used in the PIMES. The different target-setting methods across two main categories of performance indicators are emphasised because the methods are related to the strategic behaviour and gaming of public institutions.

3.2 International guidelines on performance management systems

As explained earlier, performance management systems have been prevalent in the public sector across the globe. Reviewing international guidelines on performance management systems is necessary to enhance understanding how performance management systems work in practice. The OECD has provided international guidelines on how governments operate state-owned enterprises efficiently, transparently and accountably through 'OECD guidelines on corporate governance of state-owned enterprises' since 2005 (OECD, 2005). The guidelines are internationally agreed standards for governments' state ownership to avoid the pitfalls of both passive ownership and excessive state intervention (OECD, 2015). A state-owned enterprise is defined as a corporate entity in which state exercises ownership and state-owned enterprises are the main providers of key public services such as energy, telecommunications and infrastructure in many countries (OECD, 2015). Therefore, the international guidelines on how governments operate state-owned enterprises can be applicable to the performance management systems for public service organisations.

Although there are many guidelines in the 'OECD guidelines on corporate governance of state-owned enterprises' (OECD, 2015), this thesis focuses on four key guidelines on performance management systems for state-owned enterprises. The first guideline is that the ultimate purpose of state ownership of state-owned enterprises should be to maximise value for society through an efficient allocation of resources. This guideline is linked to managerial and social accountability of state-owned enterprises because managerial accountability includes efficient use of resources and social accountability impacts on improved provision of public services (Jantz et al., 2015; Grandvoinnet et al., 2015).

The second guideline is that governments should define broad objectives for state-owned enterprises and communicate specific financial and non-financial performance objectives with them, and regularly monitor their implementation (OECD, 2015). The objectives include avoiding market distortion and pursuing profitability such as the organisation's rate-of-return (OECD, 2015). A best practice example regarding the second guideline is the Ownership Guidelines in Lithuania (OECD, 2020). The Ownership Guidelines outline the rights and responsibilities of all state-owned enterprises in Lithuania and define how they should be managed. In accordance with the Ownership Guidelines, the Lithuanian government introduced a target-setting tool called 'Letter of Expectations' in 2017 (OECD, 2020). Each ownership ministry drafts Letter of Expectations for each controlled state-owned enterprise where the ministries set financial and non-financial expectation for state-owned enterprises (OECD, 2020).

The third guideline is that governments should set up a reporting system to monitor and assess state-owned enterprises' performance and publish aggregate report on state-owned enterprises annually. The reporting system enables governments and the public to compare productivity, the use of labour

and capital of state-owned enterprises with those of private or public sector entities both domestically and abroad (OECD, 2015). A best practice example regarding the reporting system is the 'Standardised Monitoring System' that has been implemented by the German Federal Ministry of Finance since 2016 (OECD, 2020). The German Federal Ministry of Finance conducts its 'Standardised Monitoring System' based on state-owned enterprises' financial statements for the previous fiscal year at the beginning of each year (OECD, 2020). The uniform calculation of financial performance of state-owned enterprises and the description of each state-owned enterprise's business situation are provided in one standardised data sheet every year (OECD, 2020). The data sheet increases transparency and comparability within the state-owned enterprises (OECD, 2020).

The fourth guideline is that governments should establish a clear remuneration policy for a state-owned enterprise's boards, which fosters the long and medium term interests of the state-owned enterprise (OECD, 2015). The remuneration level for board members of state-owned enterprises should reflect market conditions to attract and retain highly qualified board members (OECD, 2015). This remuneration policy is related to performance-based incentive schemes such as financial incentives in performance management systems. A best practice example regarding this guideline is Poland's Law of 9 June 2016 that states the principles of remuneration for managers of companies including state-owned enterprises (OECD, 2016). The remuneration consists of a basic monthly salary and a variable amount that depends on the level of achievement of management objectives (OECD, 2020). The management objectives include an increase in net profit, increase in production and sales figures, reducing management or operating costs and achieving specific indicators (OECD, 2020).

These four key guidelines on performance management systems for state-owned enterprises are connected to the key agendas of performance management research, which were explained in Chapter 2: (i) rendering of a formal summative performance evaluation and (ii) generating and delivering performance feedback. It is against the four key guidelines that the next section discusses different performance management systems in different countries and compares them with the Korean PIMES.

3.3 Performance management systems in different countries

This section focuses on the performance management systems for public corporations and quasi-governmental institutions in Canada, France, Sweden and the UK. This thesis selected the performance management systems in Canada, France, Sweden and the UK because the systems have complied with at least one of the four key guidelines on performance management systems for state-owned enterprises. According to the OECD (2018), a public corporation can be defined as an organisation whose main decision-making is affected by government and which provides public services and supplies a market with commodities or services. A quasi-governmental institution is defined as an organisation whose main decision-making is affected by government and which provides public services (OECD, 2018).

The Canadian performance management system for public corporations complies with the third guideline of reporting system. Public corporations in Canada should submit a management plan, budget and quarterly financial statements to the Parliament of Canada (Park et al., 2020). A sponsoring department controls the management of public corporations and the Department of

Finance reviews and approves the budget of public corporations (Park et al., 2020). The French performance management system for public institutions closely follows the fourth guideline of remuneration policy (OECD, 2020). The French Law of 9 December 2016 on transparency and anti-corruption modernised the approval system for the remuneration of the management of listed public institutions by putting in place a binding vote of shareholders (OECD, 2020). In France, the Agence des Pariticipations de l'État (hereafter APE)[1] under the Ministry of Economy and Finance supervises public corporations, and the responsible ministry manages public institutions except public corporations (Park, 2013). A 3-4 year management contract is signed between the responsible ministry and the public corporation, and the management contract evaluation on public corporation is performed by APE (Park, 2013).

The Swedish performance management system for public corporations is a best practice example regarding the guidelines of target-setting, reporting system and remuneration policy. In Sweden, an annual report on public corporations has been published since 2000, and outcomes of public corporations have been evaluated based on financial targets and public policy targets (Park et al., 2020). Executives of public corporations in Sweden can receive salary and bonus according to their performance against the targets (Park et al., 2020). The UK's performance management system for public corporations and Non-Departmental Public Bodies (NDPB) complies with the third guideline of reporting system. In the UK, the evaluation of public corporations has been performed by the Shareholder Executive[2] and the evaluation of NDPB has been conducted by sponsoring departments since 2002 (Park, 2013). The Tailored Review for 2015-2020 states that reviews of public bodies

---

[1] Agence des pariticipations de l'État (APE): It was established to supervise public corporations, however, does not intervene in the management of public corporations but functions as a consultant.
[2] Shareholder Executive was replaced by UK Government Investments (hereafter UKGI) in 2016. UKGI is wholly owned by the HM Treasury, and aims to be the UK Government centre of excellence in corporate finance and corporate governance (UKGI, 2018).

should be undertaken at least once in the lifetime of the Parliament by review teams of sponsoring department and UKGI[2] advises the sponsoring department on major corporate finance matters and corporate governance (Cabinet Office, 2019; UKGI, 2019). The Minister for Cabinet Office appoints chair and all board members of advisory NDPB and can dismiss individuals whose performance is unsatisfactory (Cabinet Office, 2019).

The Korean PIMES is very different from the performance management systems for public corporations and quasi-governmental institutions in Canada, France, Sweden and the UK. The first difference is that the South Korean Ministry of Economy and Finance (MOEF) evaluates both 36 public corporations and 93 quasi-governmental institutions as of 2020 by constituting an independent management evaluation team (MOEF, 2020). The management evaluation team consists of over 100 evaluators who are professors, public accountants, lawyers and people working for a government-funded research institute with a doctor's degree. Second, public corporations and quasi-governmental institutions are regularly evaluated every year in accordance with the Act on the Management of Public Institutions in South Korea. In the UK, NDPBs are evaluated according to the plan of sponsoring department within 5 years, and, in France, public corporations are evaluated based on the management contract every three or four years (Park, 2013). The third difference is that the PIMES evaluation results are not only announced to the public but also linked to financial incentives of senior executives and all employees compulsorily.

The PIMES complies with the four key guidelines on performance management systems for state-owned enterprise, therefore, the study of PIMES as a case can fill the gap in the literature on performance management systems, which was identified in Chapter 2. In accordance with the

guideline of reporting system, the PIMES has a reporting system to provide the public with certain information such as profit, assets and debt of South Korean public institutions and the level of customer satisfaction with public services. The profit, assets and debt of public institutions are related to the efficient use of public resources, which managerial accountability emphasises. The level of customer satisfaction with public services may increase by the improvement of public service quality, which social accountability impacts on. Therefore, the reporting system of PIMES enables this study to examine the effect of performance management systems on managerial and social accountability of public organisations. Subsequently, the PIMES is operating its remuneration scheme to reward additional pay (bonus) to employees of public institutions according to the evaluation results in line with the guideline of remuneration policy. The study of the remuneration scheme in the PIMES sheds light on the effect of performance-based incentive schemes on public accountability and performance. Lastly, the PIMES sets targets and evaluates the public institutions based on the targets in accordance with the guideline of target-setting. The public institutions that are evaluated in the PIMES may engage in strategic behaviour in relation to the targets or game the target-setting to receive better evaluation results. The study of public institutions' strategic behaviour and gaming in relation to the targets contributes to research on the on the strategic behaviour of public organisations under performance management systems.

3.4 The South Korean political context

3.4.1 South Korean politics

Politics can be defined as the exercise of power, the science of government and the allocation of scare resources (Heywood, 2013). The political system of the Choson dynasty which lasted from 1392 to 1910 was a monarchy and bureaucracy (Kihl, 2015). After the state of Korea as a colony of Japan for 1910-1945, Korea was divided into Northern half occupied by the Soviet Union and Southern half by the US (Kihl, 2015; Moon and Moon, 2020). This division was institutionalised as the South established the Republic of Korea (South Korea) and the North established the Democratic People's Republic of Korea (North Korea) in 1948 (Moon and Moon, 2020). South Korea was established with democratic institutions in 1948 (Moon and Moon, 2020). The Korean War started when North Korea invaded South Korea in 1950, and lasted more than three years (Moon and Moon, 2020). In July 1953, the Korean War ended and a division between the countries became fixed after the signing of the armistice agreement (Moon and Moon, 2020). From the authoritarian rule of President Syng-man Rhee during the first Republic (1948-1960) and onwards, authoritarianism was dominant in South Korea until 1987 (Kihl, 2015). For example, Major General Chung-hee Park toppled the Second Republic in 1961, and ruled the country for 17 years (Moon and Moon, 2020).

In 1987, however, as a result of massive public protests, the ruling coalition promised the public that the government would make a constitutional amendment for direct presidential elections, greater local autonomy and freedom of press (Kim, 2003; Moon and Moon, 2020). Since then, South Korea

(hereafter Korea) has been classified as a "free country" or "full democracy" in the Freedom House Index and Economist Intelligence Unit democracy index (Lim and Roh, 2020). After the upheaval in 1987 made the first meaningful democratic opening and transition, Korea has moved toward democratic consolidation (Kihl, 2015; Moon and Moon, 2020). In 1992, Korea elected its first civilian leader in three decades, and in 1997, Dae-jung Kim won the presidential election and made the first peaceful transfer of political power (Kim, 2003; Moon and Moon, 2020). Since then, frequent shifts of political power between the conservative and progressive parties have occurred (Moon and Moon, 2020). For example, Moo-hyun Roh (2003-2007) and Jae-in Moon (2017-2022), progressive leaders, served as presidents and Myung-bak Lee (2008-2012) and Geun-hye Park (2013-2016), conservative leaders, served as presidents.

The constitution is the most important foundation of a political institution because it determines overall norms, principles, rules and procedures related to governance (Lim and Roh, 2020). The Korean Constitution has been amended nine times since it was promulgated in 1948. In accordance with the Ninth amendment in 1987, the Sixth Republic was established in Korea and still exists. This thesis focuses on the Ninth amendment because the Korean PIMES since 2014 is the case in this thesis. The Ninth Amendment was institutionalised as the '1987 constitution' which stipulates the authority of the President, the National Assembly and the Judiciary. First, the President is elected via direct election every five years and is prohibited from re-election, and represents the country as the head of state in both domestic and foreign affairs (Lim and Roh, 2020). The President appoints the Prime Minister who serves as the principal executive assistant to the President (Lim and Roh, 2020). Heads of executive ministries (ministers) are appointed by the President on the recommendation of the Prime Minister. Second, the National Assembly has the legislative power and consists of 300 members as of 2023, which are elected by direct election. The term of members of the National

Assembly is four years and the incumbents have the possibility of re-election (Lim and Roh, 2020). Lastly, the Judiciary consists of the judges of Supreme Court, High courts, District Courts, Family Courts, Administrative Courts, Patent Court and the Constitutional Courts (Lim and Roh, 2020). To sum up, Korea adopts the presidential system where the three bodies of Executive, Legislative and Judiciary are separated on the principle of the separation of power (Lim and Roh, 2020).

3.4.2 South Korean governance

Fukuyama (2013, p. 350) defines governance as a government's ability to make and enforce rules, and to deliver services, regardless of whether that government is democratic or not. Governance is essentially composed of two distinctive dimensions (Kim, 2013). The first dimension means the institutional arrangements where socio-economic resources are allocated to address issues in society (Kim, 2013). The second dimension is the activity of governing which is embodied in a policy-making processes (Deacon, 2007; Kim, 2013). The activity of governing consists of managing, coordinating and controlling resources among actors in policy initiation, formulation and implementation (Kim, 2013). Considering the definition and dimensions of governance, it is useful for understanding the Korean PIMES to review the South Korean various governance.

South Korean governance is based on strong presidency because the president appoints the prime minister and the cabinet, and directs executive and advisory agencies (Ringen et al., 2011). Central government consists of 18 ministries pursuant to the Government Organization Act as of March 2023, and ministers play a significant role in making policy decisions and implementing policies in government (Sung, 2014). For example, the Minister of Economy and Finance administers

coordination of economic and financial policies, execution and performance management of budgets, international finance and management of public institutions. Therefore, the Ministry of Economy and Finance (MOEF) is responsible for the operation of PIMES pursuant to the Act on the Management of Public Institutions. Responsible ministries receive management performance reports from the public institutions under their jurisdiction, however, the ministries do not involve the PIMES evaluation process.

Local self-governance had been a constitutional aspiration already in the First Republic, however, the elections of local government heads were first held in 1995 (Ringen et al., 2011). Local governments administer the collection of local taxes, regional development and promotion of welfare for local residents, however, work under the pressure of financial and administrative controls from the central government (Ringen et al., 2011). Local governments are responsible for the evaluation of local public institutions apart from the PIMES. The South Korean National Assembly has the authority of legislative consent on the executive policies (Ringen et al., 2011). Therefore, the Minister of Economy and Finance should report the PIMES evaluation results to the South Korean National Assembly by 20th June every year. The Economy and Finance committee of National Assembly can ask the Minister of Economy and Finance for an account of the PIMES evaluation results pursuant to National Assembly Act. Moreover, the South Korean public institutions should submit the contents of fixed or changed budget to the South Korean National Assembly.

3.4.3 The influence of NPM principles on the PIMES

Following the NPM principles of explicit standards and measures of performance and emphasis on

output controls, South Korea has increasingly developed performance management systems including performance budgeting and related performance management strategies (Yang and Torneo, 2015; Roh, 2018; Park, Kim and Kim, 2019). In the 2000s, the South Korean government attempted systematic efforts to manage performance in the public sector as a strategy of public sector reform in response to the economic crisis of the late 1990s (Yang and Torneo, 2015; Song, 2015). For example, the South Korean National Assembly enacted a new Framework Act on Government Performance Evaluation in 2006 to improve government efficiency, effectiveness and accountability (Yang and Torneo, 2015). The Act covers both performance management and evaluation of all government organisations, and was inspired by the Government Performance and Results Act in the US in 1993 (Yang and Torneo, 2015). Moreover, the Act introduced performance evaluation of public corporations, quasi-governmental institutions and local governments (Yang and Torneo, 2015).

Presidents Dae-jung Kim (1998-2002) and Moo-hyun Roh (2003-2007) strengthened the performance evaluation of public institutions by establishing a team which inspects and evaluates the management innovation of public institutions and expanding the subjects of the PIMES evaluation in 2004 (MOEF, 2011; Song, 2015). The efforts of Presidents Dae-jung Kim and Moo-hyun Roh were influenced by the NPM doctrine of improving performance in the public sector through disaggregation, competition and incentivisation (Song, 2015; Dunleavy et al., 2005; Lapuente and Van de Walle, 2020). For example, the subjects of PIMES evaluation were expanded from 24 government-invested institutions to 89 public institutions in 2004 to promote competition and improve performance in the public sector (Song, 2015). In the PIMES, performance-based remuneration (financial incentive scheme) has been operated to improve accountability and performance of public institutions from 1984 (MOEF, 2011). The financial incentive scheme in the PIMES is based on incentivisation of the NPM principle (Yang and Torneo, 2015; Lee and Ra, 2015;

Park, Kim and Kim, 2019). In addition, the results of customer satisfaction survey should be reflected on the PIMES evaluation pursuant to the Act on Management of Public Institutions. The emphasis of customer satisfaction in the PIMES aligns with the principle of NPM that the public sectors can maximise customer satisfaction with public services (Funck and Karlsson, 2019; Singh and Slack, 2022).

3.5 What is the South Korean Public Institution Management Evaluation System (PIMES)?

3.5.1 The definition and functions of PIMES

The South Korean government operates a legal system that encourages the public institutions to rationalise management and enhance transparency in operation. The legal system has been based on the Act on Management of Public Institutions after 2007. The Act determines basic things concerning the establishment of self-controlling and accountable management system as well as the operation of public institutions in order to improve the quality of public service. Therefore, the PIMES is based on the Act on Management of Public Institutions. According to the MOEF (2011, p. 31), the PIMES is defined as 'a cyclical process where an objective evaluation is made and differentiated incentives are decided based on it, the evaluation results are reflected in the management plan of public institutions in the following year'.

According to the MOEF (2011, p. 31), the functions of PIMES are as follows. Firstly, the PIMES

clearly defines the roles of the South Korean government and public institutions through management contract containing specific goals and performance-based compensation. The intention of the contract is to minimise intervention by the South Korean government and establish self-controlling of public institutions and an accountable management system. A similar management contract is signed between the state-owned enterprises and the governments in Canada, France, Sweden and the UK. Secondly, the PIMES enhances the management efficiency of public institutions by promoting a sense of accountability and providing the motivation to achieve the goals (MOEF, 2011). The MOEF is operating the financial and non-financial incentive schemes in the PIMES to motivate the public institutions to achieve their goals. Thirdly, the PIMES increases the management transparency of public institutions through the disclosure of the actual state of management such as financial statement on the South Korean government official website (MOEF, 2011). If a public institution discloses appointed information late or incorrectly, the MOEF deducts points from the PIMES evaluation result. The function is based on the assumption that disclosure of the performance of public institutions increases the management transparency through the attention of the public and media. Fourth, the PIMES leads the management innovation of public institutions by introducing competition factors such as relative evaluation (MOEF, 2011). Lastly, the PIMES addresses principal-agent problem by specifying the management targets (MOEF, 2011). This thesis examines how these functions of PIMES work in practice through the three research questions.

3.5.2 The subjects of PIMES evaluation

The subjects of PIMES evaluation are the South Korean public institutions that provide the public with public service. The subjects consist of 36 public corporations whose self-generating revenue

reaches or exceeds 50% of the amount of total revenue and 93 quasi-governmental institutions which are other than public corporations as of 2020 (See Table 3-1: Subjects of evaluation in the PIMES). The MOEF classifies public corporations into public corporation I which is responsible for infrastructure and has relatively large assets and public corporation II which constitutes all not included in public corporation I. Quasi-governmental institutions are divided into fund-management-type institutions to which the management of a fund is assigned or commissioned pursuant to the National Finance Act and commissioned-service-type institutions which are other than fund-management-type institutions. As the commissioned-service-type institutions are 80 as of 2020, the MOEF separates relatively small-sized institutions from the commissioned-service-type institutions for evaluation. Therefore, the public institutions evaluated in the PIMES fall into one of five categories in the PIMES evaluation. (i) public corporation I, (ii) public corporation II, (iii) fund-management-type institution, (iv) commissioned-service-type institution, and (v) strong and small institution (See Table 3-1: Subjects of evaluation in the PIMES).

The scope of public services covered in the PIMES evaluation is wide because the South Korean public institutions provide various types of public services. The first type is to construct and manage infrastructure such as highway, railroad and port. The second type is to create energy such as electricity and gas or dig up mineral resources and supply them to companies and the public. These types of public services are mainly provided by the public corporations. The third type is to manage the South Korean government pension system that is divided into National Pension, government employee pension, teachers pension and the military pension. Fund-management-type institutions are responsible for this type of public service. The fourth type is to provide the public with health and welfare service such as national health insurance and childcare services. The fifth type is to present schools with education information and students with training services and national scholarship. The

sixth type is to support small and medium-sized companies, start-ups and self-employed. The last type is to manage public facilities such as national parks. Commissioned-service-type institutions and strong and small institutions provide the fourth to the last type of public services.

Table 3-1: Subjects of evaluation in the PIMES

Category		Name of public institution	Weight of main performance indicators
Public corporation (36)	Public corporation I (10)	Korea Gas Corporation, Korea Electric Power Corporation, Korea Expressway Corporation, Korea Water Resources Corporation Korea Railroad Corporation, Korea Land & Housing Corporation et al.	Job creation (6) Work Efficiency (5) Finance operation and Performance (5) Main businesses (45)
	Public corporation II (26)	Korea Resources Corporation, Korea Hydro & Nuclear Power Company, Korea Minting, Grand Korea Leisure Company, Korea Racing Authority, Korea Gas Technology Corporation, Korea Marine Environment Management Corporation, Korea Coal Corporation et al.	
Quasi-governmental Institution (93)	Fund-management-type (13)	Teacher's pension, Korea Sports Promotion Foundation, National Pension Service, Korea Workers' Compensation & Welfare Service, Korea SMEs and startups Agency et al.	Job creation (6) Finance operation and Performance (2) Fund management and performance (5) Main businesses (50)
	Commissioned-service-type (37)	Korea Student Aid Foundation, Korea Trade-Investment Promotion Agency, Korea Gas Safety Corporation, Korea Energy Agency, Korea Infrastructure Safety Corporation, Social Security Information Service et al.	Job creation (6) Safety and environment (5) Main businesses (55)
	Strong and small institution (43)	Korea Public Finance Information Service, Korea Education and Research Information Service, Korea Creative Content Agency, National Institute of Ecology, Korea Human Resource Development Institute for Health & Welfare et al.	Job creation (6) Safety and environment (3) Main businesses (55)

Sources: *Announcement of public institutions management performance evaluation result* (MOEF, 2020)

3.5.3 The history of PIMES

The PIMES was first implemented in 1984 for 24 government-invested institutions with the enactment of the Framework Act on Management of Government-invested Institutions. Before the Act, government-invested institutions were operated by the principle of ex-ante management by responsible ministries (MOEF, 2011). For example, the responsible ministries exercised the right to appoint and dismiss the head of a government-invested institution (MOEF, 2011). Therefore, the management of government-invested institutions was carried out by the intervention of responsible ministries, and hence the government-invested institutions could not manage autonomously (Kim, 2014). In order to establish an autonomous management of government-invested institutions, the Framework Act on Management of Government-invested Institutions was enacted in 1984. According to the Act, the way of managing the institutions was changed into ex-post management, and hence the institutions could thus manage autonomously (MOEF, 2011). For example, the government-invested institutions could make business plans and budget by themselves after 1984 (Kim, 2014), and were evaluated in the PIMES by the MOEF not responsible ministries in return for autonomous management (MOEF, 2011).

The PIMES had evaluated between 20 and 25 government-invented institutions every year until the early 2000s, however, it was impossible to evaluate the government-affiliated institutions which provided the public with crucial public services. In order to systematically manage and evaluate the government-affiliated institutions, the Framework Act on Management of Government-affiliated Institutions was enacted in 2003. As a result, the subjects of PIMES evaluation were expanded to 89 institutions including government-invested institutions and government-affiliated institutions in 2004

(MOEF, 2011). However, some public institutions which provided crucial public services were still excluded from the subjects of PIMES evaluation. Furthermore, the government-invested institutions and government-affiliated institutions were evaluated based on the different Acts and hence it was difficult to evaluate the government-invested and government-affiliated institutions consistently (MOEF, 2011). Therefore, the MOEF merged the two Acts into the Act on Management of Public Institutions in 2007, and hence the government-invested and government-affiliated institutions were evaluated by the same method based on the Act on Management of Public Institutions.

The PIMES was significantly reformed according to the Act on Management of Public Institutions in 2007 (MPB, 2007). Firstly, the evaluation results of score and ranking were transformed into overall grade which is composed of S, A, B, C, D, E in diminishing order of excellence. The reform about the evaluation result format aimed to prevent the dysfunction of ranking. In the former system, public institutions ranked according to their score, and so even the score gap of 0.1 could generate different financial incentives between the public institutions. However, if the two public institutions with similar ranking are at the same overall grade in the reformed system, they would receive the same financial incentives. It may enhance the rationality of incentive scheme in the PIMES. Secondly, the weighting range of management performance rose from 30~40% to 40~45% to lead to the improvement of management performance. Lastly, the two separate evaluation systems for government-invested institutions and government-affiliated institutions were unified as the PIMES to define the evaluated public institutions clearly. As a result, government-invested institutions were named public corporations and government-affiliated institutions were named quasi-governmental institutions.

In 2013, the PIMES was significantly reformed again to normalise the management of public institutions. The South Korean public institutions had been criticised by the public because the debt of institutions increased by about 70% during over the four years and the institutions provided excessive benefits to employees (MOEF, 2013). The South Korean government focused on the decrease of the debt and excessive benefit of public institutions via the reform of PIMES and other tools. In the PIMES, the weight of the indicators that related to controlling excessive pay and benefits increased from 8% to 12% to control the excessive pay and benefit of public institutions (MOEF, 2013). Moreover, the South Korean government expanded its disclosure of information concerning the benefits of public institutions on the South Korean government official website (MOEF, 2013). The aim of reform was to reduce the debt of public institutions dramatically and establish an autonomous management innovation and inspection system (MOEF, 2013). As explained in Chapter 2, managerial accountability includes responsibility for an efficient use of resources (Ospina, Grau and Zaltsman, 2004). As the reduction of debt of public institutions can be achieved through the efficient use of resources, the aim of reform in 2013 relates to the improvement of managerial accountability. Therefore, this thesis examines the effect of PIMES on the managerial accountability of public institutions from 2014 to 2019.

3.5.4 The PIMES evaluation process

In South Korea, the management evaluation on the public institutions is operated every year based on the Act on the Management of the Public Institutions. The PIMES evaluation process is composed of three steps on a three year basis as follows.

① Preparation (previous year)

Criteria and methods for the management evaluation such as performance indicators and evaluation weighting are finalized and provided to the public institutions by the MOEF by the end of December based on the Enforcement Decree of the Act on the Management of Public Institutions.

② Business performing (current year)

Public corporations and quasi-governmental institutions conduct business to achieve their management goals from January to December.

③ Evaluation (next year)

Public corporations and quasi-governmental institutions should submit management performance reports to the MOEF and the responsible ministries by 20th March. The management performance reports are evaluated and due diligence is carried out by over 100 evaluators who are separated from the MOEF. As the evaluators are experts on the public institutions or evaluation, an efficient evaluation of management performance and consultation concerning management performance are provided by the evaluators. The evaluation results of overall grades and follow-up measures are finalized by the MOEF by 20th June and the results are reported to the President and the South Korean National Assembly by the end of August.

A customer satisfaction survey is conducted for most public institutions which are evaluated in the PIMES every year and the survey results are used in the PIMES evaluation process. The MOEF is responsible for the customer satisfaction survey and selects an agency to design survey questions and the sample and selects several agencies to conduct survey. The agencies survey the customer satisfaction level of sample extracted from the customer list via email or phone. After the survey, the MOEF analyses the survey results and calculates the customer satisfaction score.

According to the evaluation results of overall grades, a different set of financial incentives (bonus) is provided to the employees of public corporations and quasi-governmental institutions (see

Table 3-2: Financial incentives according to the evaluation results of overall grades). In accordance with the evaluation results of overall grades, the employees of public corporations would receive 0~250% of monthly pay additionally as a bonus, whilst the employees of quasi-governmental institutions would receive 0~100% of that as a bonus every year. However, if a public institution receives the overall grade of D or E in the PIMES evaluation, it would not receive any bonus. Moreover, the institution would not only submit a managerial improvement plan to the MOEF and to the responsible ministries but also would be managed to implement the plan. Furthermore, if the overall grade of a public institution is 'D', disciplinary measures such as a warning would be issued to the public institution head who has worked for the institution for over six months. Even recommendations for dismissal would be made to the public institution head who has worked for the institution for over six months if the overall grade of public institution is 'E'. In addition, all the public institutions are required to disclose their management information such as current status of business, workforce, finance to the public via the South Korean government official website of ALIO addressed

as www.alio.go.kr.

Table 3-2: Financial incentives according to the evaluation results of overall grades

Overall grade	S	A	B	C	D and E
Public corporation	250% of monthly pay	200% of monthly pay	150 of monthly pay	100% of monthly pay	0%
Quasi-governmental institution	100% of monthly pay	80% of monthly pay	60% of monthly pay	40% of monthly pay	0%

Sources: News release 'Confirmation of evaluation result on public institution management performance in 2015' (MOEF, 2016)

3.5.5 The performance indicators of PIMES

The PIMES evaluation of public institutions uses between 22 and 30 performance indicators as of 2019. According to the guidance on the management evaluation of public institutions (MOEF, 2019a), the performance indicators of PIMES are composed of two categories: business administration and main businesses to systematically and comprehensively evaluate the public corporations and quasi-governmental institutions. The category of business administration in the public corporations consists of six sub-categories: (i) management strategies and leadership, (ii) social value realisation, (iii) work efficiency, (iv) management of organization, personnel, finance, (v) management of pay and welfare benefits, and (vi) innovation and communication (see Table 3-3: Performance indicators and weight criteria for the public corporations). The category of business administration in the quasi-governmental institutions is made up of five sub-categories except work efficiency (see Table 3-4: Performance indicators and weight criteria for the commissioned-service-type quasi-governmental institutions).

Targets of performance indicators are set differently across the two categories. In the first category of business administration, the MOEF set the targets to induce the public institutions to follow government policies. For example, the MOEF gives all the public institutions the target that disabled employees must be above a certain percentage of all employees. The target represents 'equal opportunity and social integration' performance indicator in the sub-category of 'social value realisation'. However, the public institutions can propose their own new targets in the second category of main businesses if exiting targets become inappropriate in a changing environment or are difficult to achieve. The public institutions discuss the new targets with evaluators and new targets are confirmed as proposed or with a variation after the discussion. The difference of target-setting methods between the first and second categories of performance indicators is related to different strategic behaviour. This thesis will explore the different strategic behaviour of public institutions concerning the target-setting methods in Chapter 6.

The weighting of qualitative and quantitative performance indicators is 56% and 44% respectively regardless of the category of public institution as of 2019. The grade of qualitative indicators is determined by a nine-point scale, and the score of quantitative indicators is calculated by comparing management performance and predetermined target. The overall grade of a public institution is determined according to the sum of the score of quantitative indicators and the grade of qualitative indicators.

Table 3-3: Performance indicators and weight criteria for the public corporations

Category	Performance indicator	Sum	Qualitative	Quantitative
Business Administration	1. Management strategies and Leadership	6	6	
	- Strategy planning		2	
	- Management enhancement		2	
	- Leadership		2	
	2. Social value realization	24	17	7
	- Job creation	7	4	3
	- Equal opportunity and social integration	4	3	1
	- Safety and environment	5	5	
	- Co-prosperity, cooperation and regional development	5	2	3
	- Ethical management	3	3	
	3. Work efficiency	5		5
	4. Management of organization, personnel, finance	7	4	3
	- Organization, personnel generalities	2	2	
	- finance operation and Performance	5	2	3
	5. Management of pay and welfare benefits	8	5	3
	- pay and welfare	3	3	
	- Management of total labour cost	3		3
	- Labour-management relations	2	2	
	6. Innovation and communication	5	3	2
	- Innovation effort and performance	3	3	
	- People's feedback	2		2
	Subtotal	55	35	20
Main businesses	Comprehensive evaluation on plan, operation and performance of main businesses	45	21	24
	Total sum	100	56	44

Source: Guidance on management evaluation of public institutions (MOEF, 2019a)

Table 3-4: Performance indicators and weight criteria for the commissioned-service-type quasi-governmental institutions

Category	Performance indicator	Sum	Qualitative	Quantitative
Business Administration	1. Management strategies and Leadership	6	6	
	- Strategy planning - Management enhancement - Leadership		2 2 2	
	2. Social value realization	22	15	7
	- Job creation - Equal opportunity and social integration - Safety and environment - Co-prosperity, cooperation and regional development - Ethical management	6 3 5 5 3	3 2 5 2 3	3 1 3
	3. Management of organization, personnel, finance	4	3	1
	- Organization, personnel generalities - finance operation and Performance	2 2	2 1	 1
	4. Management of pay and welfare benefits	8	5	3
	- pay and welfare - Management of total labor cost - Labor-management relations	3 3 2	3 2	 3
	5. Innovation and communication	5	3	2
	- Innovation effort and performance - People's feedback	3 2	3 	 2
	Subtotal	45	32	13
Main businesses	Comprehensive evaluation on plan, operation and performance of main businesses	55	24	31
	Total sum	100	56	44

Source: Guidance on management evaluation of public institutions (MOEF, 2019a)

3.6 Conclusion

This chapter provided four key international guidelines on performance management systems and best practices based on the 'OECD guidelines on corporate governance of state-owned enterprises' (OECD, 2015; 2020). In accordance with the four international guidelines, different performance management systems in Canada, France, Sweden and the UK were explained. The chapter compared the Korean PIMES with the performance management systems in Canada, France, Sweden and the UK, and discussed how the PIMES is situated in terms of the four international guidelines. The discussion concluded that the PIMES is designed complying with the four international guidelines, therefore, the study of PIMES fills the gap in literature on design and effect of performance management systems in the public sector.

South Korean democracy follows the governance in which the Executive, the Legislative and the Judiciary are separated on the principle of a separation of power (Lim and Roh, 2020). In accordance with the principle, the PIMES is being operated by the Ministry of Economy and Finance and the South Korean National Assembly has authority to ask the Ministers of Economy and Finance for an account of the PIMES evaluation results. The PIMES was influenced by NPM principles of disaggregation, competition and incentivisation (Song, 2015; Dunleavy et al., 2005) and hence subjects of PIMES evaluation expanded and financial incentive scheme was adopted in the PIMES (Song, 2015; Singh and Slack, 2022). Moreover, the PIMES reflects the level of customer satisfaction with public services on the PIMES evaluation.

The PIMES evaluates over 100 public institutions every year and the public services evaluated in the PIMES include infrastructure, energy, national pensions, health and welfare service, education, support for start-ups and the management of national facilities. This wide scope of PIMES evaluation enables this study to examine how the design of performance management systems regarding reward affects accountability of various public service organisations. The PIMES uses the target-setting method that the South Korean government sets targets or the public institutions propose new targets and discuss them with evaluators. This target-setting method may cause various types of strategic behaviour and gaming, therefore, the study of target-setting method in the PIMES fills the gap in literature on strategic behaviour or gaming under performance management systems.

Chapter 4 Research Design and Methodology

4.1 Introduction

As mentioned in Chapter 2, NPM reforms are still a strong force for change in most public administrations (Willis et al., 2017; Pollitt and Bouckaert, 2017; Hammerschmid et al., 2019). A major objective of NPM reforms is to obtain good performance in the public sector and many see the path to this as lying in the improvement of accountability (Pollitt and Bouckaert, 2017; Lapuente and Van de Walle, 2020; Christensen and Lægreid, 2015). This view is not universally held – for example, Dubnick (2005) argues an 'accountability paradox' in which more accountability actually diminishes organizational performance. This is because giving account such as speech and reporting demands energy and time of account giver, which would otherwise be devoted to improvement of performance (Dubnick, 2005). However, the mainstream view, as argued, for example, by Bovens (2010) argues that public accountability can serve as a tool to induce governments and agencies to improve their performance by providing external incentives such as rewards and sanctions related to the effects of their policies. This thesis focuses on public accountability as a tool to improve performance in the public sector based on the NPM reforms' assumption that improved accountability will increase performance.

This chapter sets out an appropriate research design and methodology for addressing the three research questions. The chapter starts with discussion of the three research questions. It summarises how the three research questions are developed from the literature review. The first section also

reviews earlier arguments on why the three research questions are important and shows in detail how the chosen research method is appropriate for addressing each research question. The second section discusses the research philosophy that constitutes the foundation of the whole research design. The third section outlines the case study design of this research and the methodological framework, which is a mixed methodology incorporating quantitative and qualitative methods. The last two sections present the specific research design for each research question. The quantitative approach consists of a longitudinal study to analyse data from 2014 to 2019. A qualitative approach based on interviews is used to examine the effect of operational methods of PIMES and the strategic behaviour and gaming of public institutions from the perspective of principal-agent theory.

4.2 Research questions

As noted, the objective of this thesis is to examine the impact of performance management systems on accountability in the public sector, which is a contested area of theory as seen in Chapter 2. Specifically, the aim is to explore this through a case study of the effectiveness of the South Korean PIMES with respect to accountability and performance. In line with the key factors identified in the literature review in relation to the accountability-performance relationship, the key focal points in the case study of PIMES are the PIMES evaluation results, the PIMES incentive schemes, the performance indicators and processes of PIMES and the strategic behaviour of South Korean public institutions.

The research objective will be addressed via the three research questions below.

(1) Do performance management systems improve the accountability of public organisations?

(2) How do the conditions and processes of performance management systems affect the accountability of public organisations?

(3) What is the impact of performance management systems on the strategic behaviour of public organisations?

The first research question aims to examine the relationship between performance management systems and accountability in the public sector, which is still contested. A mainstream view from NPM principles is that sophisticated performance management systems can considerably improve public accountability which leads governments and agencies to improve their performance (Pollitt, 2011; Bovens, 2010). As explained in Chapter 3, the South Korean government reformed the management of South Korean public institutions based on the principles of NPM in the 2000s and hence the PIMES is to some extent designed according to the principles of the NPM reforms (Park, Kim and Kim, 2019; Lee and Ra, 2015). Therefore, research on the PIMES can shed light on the academic debate about the effect of NPM reforms and the relationship between performance management systems and accountability in the public sector. On the basis of the link between the PIMES and the NPM reforms, this study investigates how the PIMES has affected the accountability of South Korean public institutions to address the first research question.

In quantitative approaches to an issue, researchers use theory deductively and test the theory rather than develop it (Creswell and Creswell, 2018). While the NPM theories argue the positive effect of

performance management systems on public accountability, some literature suggests that evidence on the NPM theory that performance management systems lead to better performance via improved accountability is scant (Pollitt, 2011; Van Dooren et al., 2015). Therefore, it is reasonable to examine the hypotheses derived from the NPM theory through a quantitative approach to answer the first research question. Testing of the effect of new system or reformed system can be conducted by using panel data on the operation of the system (Gayle and Lambert, 2018). This study's panel data from 2014 to 2019 enables an examination of the effect of PIMES on the accountability of South Korean public institutions over time. The panel data is collected from the South Korean government official website www.alio.go.kr and the South Korean government's evaluation reports which are released on the website. The evaluation results and values of performance indicators of PIMES are released annually and so longitudinal study with panel data has been feasible. Considering these reasons, this research adopts a quantitative approach in order to address the first research question.

The second research question seeks to investigate how the operational methods of performance management systems affect the accountability in the public sector, using South Korean data. The operational methods of PIMES involve the incentive schemes, performance indicators and target-setting methods as described in Chapter 3. The investigation of the operational method of PIMES is differentiated from the examination of the first research question in that the current operational methods of PIMES are the subject of investigation rather than the evaluation results of PIMES over time. As there is a lack of research that analyses how operational methods of performance management system affect public accountability (Van Dooren et al., 2015), this investigation contributes to studies on the effect of performance management systems as a working system on public accountability.

The investigation of the operational methods of PIMES requires a large picture to mirror the ways that events operate in the real world rather than a simple linear model of cause and effect (Creswell and Creswell, 2018). The qualitative approach seeks to develop a complex picture of the problem or issue under study (Creswell and Creswell, 2018), and hence such a qualitative approach is the most appropriate method to address the second research question. In order to understand the operational methods of PIMES, this research uses semi-structured interviews and document analysis as methods of data collection.

The third research question focuses on exploring the strategic behaviour and gaming of public organisations under performance management systems from the perspective of principal-agent theory. According to principal-agent theory, as described in Chapter 2, public institutions are agents and the MOEF and the public are principals regarding the operation of PIMES and the provision of public services. In the PIMES as the case in this research, the South Korean government as a principal seeks to mitigate the strategic behaviour of the agents. The principals suffer from asymmetric information but the PIMES is used to mitigate this problem. With respect to principal-agent theory, exploring the strategic behaviour of public institutions can contribute to examining whether agency problems such as moral hazard and gaming can be mitigated using performance management systems. In addition, this exploration may generate information that is likely to have useful policy implications for refining the current PIMES.

In order to explore strategic behaviour with respect to principal-agent theory, qualitative research is useful, because it is concerned with how people think and act in their everyday lives and focuses on exploring individuals' experiences with a phenomenon of interest (Taylor, Bogdan and DeVault,

2016; Plano Clark and Ivankova, 2016). In addition, it is useful to examine the quantitative data that objectively shows strategic behaviour, because it is likely that participants will be explicitly reluctant to provide information that shows them in an unfavourable light. Therefore, the third research question will be answered using a mixed methods research that involves integration of quantitative and qualitative research and data (Creswell and Creswell, 2018).

4.3 Research philosophy

The philosophical perspective of the researcher forms the basis of their reasoning as well as the research method and methodology employed in the practice of social research (Williams and May, 1996; Grix, 2010; Agassi, 2011; Risjord, 2014). Moreover, the philosophical perspective is important for understanding the kinds of support that the researcher can claim for conclusions (Perri and Bellamy, 2012). Therefore, it is necessary to explain the philosophical perspective which is adopted in this research.

According to Grix (2010), there are three broad paradigms in the philosophy of social science: positivism, interpretivism and critical realism. Positivism views that the world exists independently of our knowledge and uses scientific methods to analyse the social world (Grix, 2010). Positivists believe that regular relationships between social phenomena can be established by using theory to generate hypotheses, which can then be tested by direct observation (Grix, 2010). Positivists believe in the possibility of making causal statements and suggest that causal explanation is relevant to understanding how observation relates to deeper causes (Grix, 2010; Perri and Bellamy, 2012). On

the other hand, interpretivism takes the view that the world does not exist independently of our knowledge and the separation of fact and value is not as clear-cut as positivism claims (Grix, 2010). In contrast to positivists, interpretivists believe that there is a clear distinction between the social and the natural world and hence the social world needs to be studied with methods different from those used in the natural sciences (Grix, 2010). Interpretivists focus on understanding the meaning given to the world and hold that knowledge achieved in secondary interpretation is valuable (Grix, 2010; Perri and Bellamy, 2012). Perri and Bellamy (2012) explain that a secondary interpretation is the researchers' account of actors' primary interpretation and the more accurately a secondary interpretation captures the essential structure of primary interpretations, the greater is its validity.

Since the 1970s critical realism has grown in importance as a powerful alternative to both positivism and interpretivism and has proved very fruitful in stimulating new research agenda (Grix, 2010; Benton and Craib, 2011; Sayer, 2000). Critical realism attempts to combine 'explanation' linked to positivism and 'understanding' linked to interpretivism by bridging the gap between the two extremes (Grix, 2010). In other words, critical realism shares with interpretivism the view that social phenomena are concept-dependent and intrinsically meaningful, however, it does not rule out causal explanation because material change should be explained too (Sayer, 2000). Critical realism proposes a stratified ontology divided into three domains: the real, the actual and the empirical (Sayer, 2000; Zachariadis, Scott and Barrett, 2013).

'The real' includes objects and structures with inherent causal powers and liabilities that result in mechanisms. 'The actual' refers to the events generated from mechanisms and is a subset of 'the real'. 'The empirical' is a subclass of observable and experienced events (Sayer, 2000; Downward and

Mearman, 2007; Zachariadis et al., 2013). Zachariadis et al. (2013) argue that critical realism views causality not as a relationship between distinct events but as a realization of the processes and conditions under which one event causes another event. This critical realism position on causality means that a different mode of inference from deduction or induction needs to be adopted in order to explain events by postulating mechanisms that are capable of producing the events (Zachariadis et al., 2013). This logic of different inference allows an approach to research in critical realism to embrace a variety of methods where qualitative and quantitative approaches can be integrated in order to identify the mechanisms (Sayer, 2000; Downward and Mearman, 2007; Zachariadis et al., 2013).

This research is based on the philosophical perspective of critical realism. Firstly, this research examines a causality regarding the conditions and processes of PIMES that affect the accountability of South Korean public institutions. In other words, the research not only examines whether the PIMES evaluation affects the managerial and social accountability of public institutions but also explores which conditions and processes result in the effect of PIMES evaluation. Therefore, the analysis of causality in this research matches the critical realism's view that causality should be about realizing the conditions and processes under which one event causes another event (Zachariadis et al., 2013). Secondly, critical realism is compatible with a relatively wide range of research methods and particular choices among them should depend on the nature of the object of study and the research questions under consideration (Sayer, 2000; Zachariadis et al., 2013). As explained earlier, this research adopts a mixed methods approach that can employ a quantitative approach or qualitative approach or some combination in order to address the three main research questions (Creswell and Creswell, 2018). Therefore, the research method of this study corresponds with the epistemological assumptions of critical realism. Considering the causality and research method of this study, the philosophical perspective adopted is in line with critical realism.

4.4 Methodological framework

4.4.1 Case study design

According to Yin (2018), there are four methodological modes: experimental design, survey research, random assignment and case study design. Among these, this research adopts a case study design as its methodological mode, reviewing the three important conditions that Yin (2018) suggests to distinguish the different methodological modes. The first condition is the form of research question. For example, 'how' and 'why' questions are likely to favour an experimental design and case study design, whilst 'what', 'who' and 'where' questions are likely to lead to the use of survey research. In this study, the overarching research question is how performance management systems affect the accountability of public organisations. Therefore, an experimental design or a case study design is more appropriate for the 'how' research question of this study. The second condition is the control that a researcher has over actual behavioural events. An experimental design requires a researcher to manipulate behaviour in a laboratory or a field setting, however, case study design or survey research is appropriate when the relevant behaviours cannot be manipulated. The evaluation results and the operational methods of PIMES and the strategic behaviour of public institutions being studied in this research cannot be manipulated by a researcher. Hence, case study design or survey research is appropriate. The third condition is the degree of focus on contemporary events which include the recent past and the present. The four methodological modes deal with contemporary events, hence this research investigating the impact of the PIMES on the accountability of South Korean public institutions can adopt all four methodological modes regarding the third condition. As a result, case study design is the most appropriate methodological mode reviewing the three conditions.

Yin (2018, p.15) defines case study design as, "an empirical method that investigates a contemporary phenomenon (the "case") in depth and within its real-world context". This definition distinguishes case study design from an experimental design and a survey research in that an experimental design deliberately separates a phenomenon from its context by controlling the context in a laboratory environment and a survey research's ability to investigate a context is limited (Yin, 2018). Yin (2018) argues that case study research designs are sorted into four types: single-case holistic design, single-case embedded design, multi-case holistic design and multi-case embedded design. Among these four types, the case study design of this research is a single-case embedded design that involves units of analysis at more than one level (Yin, 2018). According to Miles, Huberman and Saldaña (2013), a case is a phenomenon of some sort occurring in a bounded context. In this case study design, the single-case is the PIMES in Korea from 2014 to 2019.

There are five rationales for single-case study designs, which are related to the theory or theoretical propositions of a research: critical, unusual, common, revelatory and longitudinal case (Yin, 2018). Among these five rationales, the rationale of this single-case study design is of being a critical case with regard to theoretical propositions about the effect of performance management systems on public accountability. As explained in Chapter 3, the PIMES is designed in accordance with the four key international guidelines on performance management systems. Therefore, the single-case of this study is critical to how the design of performance management systems affects the accountability and performance of public organisations. This single-case study design involves units of analysis at more than one level. In other words, this research analyses many South Korean public institutions and their employees as the second level, which are embedded in the PIMES as the first level. Thus, this research has adopted a single-case embedded design (Yin, 2018).

4.4.2 Mixed methods research

Figure 4-1: Conceptual framework for convergent mixed methods design
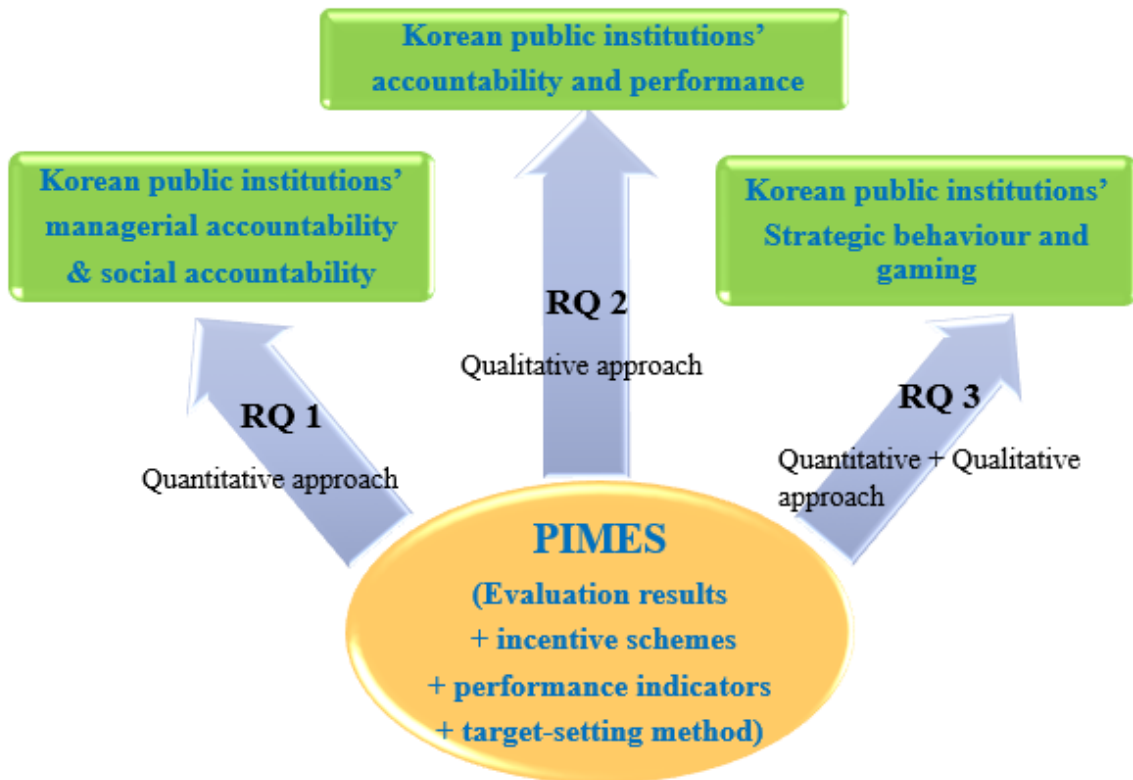
Source: Author's construction

Punch (2014, p.302) defines mixed methods research as, "empirical research that involves the collection and analysis of both qualitative and quantitative data". Creswell and Creswell (2018, p.4) also define mixed methods research as, "an approach to inquiry involving collecting both quantitative and qualitative data, integrating the two forms of data, and using distinct designs that may involve philosophical assumptions and theoretical frameworks". As case study design can include quantitative evidence as well as qualitative evidence (Yin, 2018), case study design is compatible

with a mixed methods research design.

Researchers began to develop the groundwork for mixed methods research in the 1990s and there has been a growth of interest in mixed methods research since the 2000s (Punch, 2014). For example, there are journals specifically devoted to mixed methods research such as the *Journal of Mixed Methods Research* and the *International Journal of Multiple Research Approaches* (Creswell, 2015). Moreover, the number of Google Scholar search results related to mixed methods research and OLS regression had increased between 2000 and 2013 and the number of search results for mixed methods research especially grew from 8% to 47% of the number of search results for OLS regression during the period (Seawright, 2016).

Creswell and Creswell (2018) suggest reasons for choosing mixed methods research at general, practical and procedural level. At a general level, researchers choose mixed methods research because it has the strength of drawing on quantitative and qualitative research. It also minimises the limitation to ethical applicability in quantitative research and the lack of data control in qualitative research (Punch, 2014; Blaikie, 2010). At a practical level, mixed methods research provides a sophisticated approach to research that attracts those at the forefront of new research procedures, and it can be an ideal approach when researchers have access to both quantitative and qualitative data. At a procedural level, mixed methods research can be a useful strategy to have a more complete understanding of research questions such as comparing different perspectives drawn from quantitative and qualitative data and evaluating both the processes and outcomes of a policy decision. In addition, Plano Clark and Ivankova (2016) suggest that researchers can produce more credible studies that present both complementary and corroborating evidence about research problems by using a mixed methods

95

research. Considering these advantages of mixed methods research, the rationale for using mixed methods research is to gain deeper insights into research problems and then ensure that the inferences made are valid (Plano Clark and Ivankova, 2016).

In mixed methods research, the practical considerations are timing, integration and priority, which produce meaningful inferences in response to research questions (Plano Clark and Ivankova, 2016). Timing is divided into concurrent timing which means that researchers collect and analyse both quantitative and qualitative data at the same time and sequential timing which implies that researchers collect and analyse quantitative and qualitative data in sequence (Plano Clark and Ivankova, 2016). Integration refers to an explicit interrelating of quantitative and qualitative methods in a mixed methods research (Creswell and Plano Clark, 2011). Approaches to integration can be sorted into either combining quantitative and qualitative results during their joint interpretation at the completion of both data collection and analysis or connecting quantitative and qualitative methods when the results from the first quantitative or qualitative phase are used to inform the data collection of the sequent qualitative or quantitative phase (Plano Clark and Ivankova, 2016). Priority is the relative importance of quantitative and qualitative methods for addressing research questions (Creswell and Plano Clark, 2011). Mixed methods research has either a quantitative priority or a qualitative priority or equal priority depending on the role that both types of data play in answering research questions (Plano Clark and Ivankova, 2016).

Basic designs at the heart of every mixed methods research consist of three types: convergent design, explanatory sequential design and exploratory sequential design (Creswell, 2015). Convergent design is carried out by collecting both quantitative and qualitative data concurrently, analysing both sets

separately and comparing results. The purpose of the convergent design is to compare two different perspectives on a topic and explain or interpret the resultant convergence or divergence. The procedure of explanatory sequential design is followed by collecting and analysing quantitative data, scrutinising the results, and then following up on the results with qualitative data collection and analysis. Explanatory sequential design aims to help explain the quantitative results with the qualitative data. Exploratory sequential design begins with collecting and analysing qualitative data and researchers use the results from the analysis to develop something quantitative such as an instrument or an intervention and then applies or tests the quantitative instrument or intervention. The intent of the design is to explore instrument or intervention activities to study a population. Most mixed methods research falls into one of three basic designs or some variation of them (Creswell, 2015).

As explained earlier, this study is based on the philosophical perspective of critical realism and adopts a mixed methods approach within critical realism. One of the purposes of mixed methods research within critical realism is to ensure as complete a picture as possible (as detailed as possible) of the phenomenon under research (Zachariadis et al., 2013). This purpose can be achieved when a specific object in research has a quantifiable property or a researcher wants to measure the number of objects that belong to a class while the researcher is endeavouring to improve qualitative understanding at the same time (Zachariadis et al., 2013). In this study, the evaluation results and values of performance indicators of PIMES are quantifiable. In addition, this study aims to improve qualitative understanding about the effect of design of performance management systems on public accountability and performance through a case study of the PIMES. Therefore, it is argued that a mixed methods approach within critical realism is appropriate for seeking to provide as complete a picture of PIMES as possible as the specific object and the case of this research.

When quantitative and qualitative methods are used alongside each other, establishing an appropriate alignment between research questions and research methods is very important especially in today's research environment (Punch, 2014). In order to achieve such alignment, it is helpful to arrange for the research methods used to follow the specific nature of research questions (Punch, 2014; O'Leary, 2017; Downward and Mearman, 2007). As mentioned earlier, the overarching research question in this study is how performance management systems affect the accountability of public organisations and this study investigates the PIMES as a single case to answer the overarching research question. The investigation of the PIMES requires collecting both quantitative and qualitative data regarding the evaluation results, values of performance indicators and operational methods of PIMES and strategic behaviour of public institutions, analysing both sets separately and integrating the results of each analysis. Therefore, it is argued that mixed methods research following the above procedure is the most appropriate methodology for this study to address the overarching research question.

Specifically, the mixed methods research of this study falls into a convergent design that is one of the three basic designs (Creswell, 2015). Regarding timing, both quantitative and qualitative data are collected concurrently. The main source for the quantitative data is the evaluation results and values of the performance indicators of PIMES from 2014 to 2019, while the source of qualitative data is semi-structured interviews and documents. With respect to integration, this research analyses both quantitative and qualitative data separately and combines quantitative and qualitative findings during their joint interpretation. In terms of priority, this research emphasises both databases equally, which is considered the best approach for convergent design (Creswell and Creswell, 2018). If the reason for choosing a mixed methods research is to compare different perspectives drawn from quantitative

and qualitative data, a convergent design is appropriate (Creswell and Creswell, 2018). This research compares findings produced through quantitative and qualitative methods and hence a convergent mixed methods design is the most appropriate. The convergent mixed methods design of this study is summarised in Figure 4-1. The next two sections discuss the quantitative and qualitative approaches of this study in more depth.

4.4.3 Retroductive process

Downward and Mearman (2007) present 'retroduction' as the logic of inference espoused by critical realism. 'Retroduction' can be contrasted to other inferences such as deduction and induction, as not simply developing specific claims from general premises nor general claims from specific premises (Downward and Mearman, 2007). Sayer (1992) defines 'retroduction', as the mode of inference in which events are explained by postulating mechanisms that are capable of producing events.

Four main phases are identified to achieve a mixed methods research's objective that is to link the structures and the causal powers of objects under study to the events through the notion of causal mechanism (Zachariadis et al., 2013). The first phase is to describe the research context and identify the composite events of phenomena under study. As it is not feasible to examine all possible constituents of a phenomenon, a researcher should select which constituents to examine. This phase enables a researcher not only to identify data sources such as interviews, historical documents and surveys but also to conduct a deeper examination into realistic insights afforded by each data type and analysis method. The second phase is the actual retroductive analysis of the data. This phase involves hypothesising about the mechanisms or structures capable of generating the phenomenon

observed and analysing objects in terms of their constitutive structures and causal powers. In this phase, a researcher engages in a process of abstraction where propositions are developed for use in subsequent phases of investigation. In addition, this analytical dynamic at work in the process of abstraction involves iterative cycles of reflection between academic literature, data and propositions in order to achieve analytical stability about the mechanism. The third phase involves the critical assessment and elimination of the alternative explanations. This phase includes comparison between the findings or inferences produced by the combination of method. The fourth phase is about action to circulate research findings to see if the findings are satisfactory to an intended audience with expertise. This retroductive process can be usefully used to guide activities as a study progress (Zachariadis et al., 2013).

This research methodology follows the retroductive process. The process of this research methodology based on the three phases of retroduction except circulation of research findings is as follows.

① Description of research context and identification of constituents of phenomenon under research

In Chapter 2, the issues of accountability deficits, the relationship between performance management systems and public accountability, effectiveness of incentive schemes and principal-agent problems were discussed. The elements of performance management systems and public accountability that are focused on in this study are the evaluation results and values of performance indicators of the PIMES and the managerial and social accountability emphasized under NPM reforms in the quantitative approach. In the qualitative approach, the incentive schemes, the performance indicators and the

target-setting method of PIMES and the strategic behaviour of South Korean public institutions are selected because they are most relevant to public accountability.

② Retroductive analysis of data

This research hypothesises about the relationship between the PIMES and the accountability of South Korean public institutions with reflection on the literature review and the overview of PIMES in Chapters 2 and 3. In the quantitative approach, the numerical data regarding the managerial and social accountability of South Korean public institutions and the strategic behaviour of public institutions are analysed. Regression models are employed to test the hypotheses about the impact of PIMES on the accountability of South Korean public institutions. In the qualitative approach, the data collected from semi-structured interviews and documents is analysed. The analysis aims to further understand the different views about the impact of PIMES on the accountability of South Korean public institutions and explore the conditions and processes in the PIMES to improve the accountability of South Korean public institutions.

③ Critical assessment of alternative explanations

This research examines the findings produced by both quantitative and qualitative approaches in order to reach an overall synthesis of findings. This phase enables this research to establish the validity and robustness of findings based on triangulation from the quantitative and qualitative approaches.

4.5 The quantitative approach

4.5.1 Introduction

A quantitative approach is used in this thesis to analyse the impact of PIMES on the managerial and social accountability and strategic behaviour of public institutions. A quantitative approach focuses on measuring a set of variables to answer research questions or to test hypotheses through surveys or experiments (Creswell and Creswell, 2018). It also enables researchers to make inferences about relationships between variables (Creswell and Creswell, 2018). The quantitative approach of this research employs secondary data analysis which is a study of problems through analysis of existing data collected for another purpose (Smith, 2008; Punch, 2014). The first research question is addressed through a quantitative approach as follows. Firstly, the researcher uses the theoretical analysis in Chapter 2 to develop hypotheses and then operationalises the concepts with a number of appropriate variables that the researcher has collected. Secondly, the researcher measures the variables and then analyses these data to test the hypotheses. Thirdly, the researcher interprets findings from the quantitative analysis. In addition, this quantitative approach is used to answer the third research question by inferring the relationship between the variables in relation to the PIMES evaluation results and the strategic behaviour of South Korean public institutions.

According to Punch (2014), a continuum of quantitative research designs consists of experiment, quasi-experiment and non-experiment. A researcher manipulates independent variables in an experiment and has naturally occurring treatment groups in a quasi-experiment. On the other hand, a

researcher has naturally occurring variation in independent variables in a non-experiment. A non-experiment studies a world after things have happened with correlation and regression as their main features (Punch, 2014). The quantitative approach of this research follows the case of non-experiment because the researcher has already occurring variation in the independent variable regarding the evaluation results of PIMES and adopts regression models using the occurring variation in the independent variable. This quantitative approach applies bivariate and multiple linear regressions to examine the effect of the evaluation results of PIMES on each of the three dependent variables in relation to the managerial accountability, social accountability, and strategic behaviour of South Korean public institutions. Each of the three dependent variables is regressed first on the PIMES evaluation results, and secondly on the PIMES evaluation results and a set of relevant control variables.

4.5.2 Hypotheses

Creswell and Creswell (2018, p.136) define quantitative hypotheses as, "predictions that the researcher makes about the expected outcomes of relationships between variables". In order to test hypotheses, researchers employ statistical procedures in which researchers draw inferences about a population from a sample (Creswell and Creswell, 2018). Hypotheses are divided into null (H0) and alternative hypotheses (H1): a null hypothesis makes a prediction that no relationship or significant difference exists between variables, while an alternative hypothesis makes a prediction about the expected outcome based on prior literature or studies on the topic, which argue for a potential outcome (Creswell and Creswell, 2018). In this quantitative approach, the first and second null and alternative hypotheses below are set out as follows in order to address the first research question and the third

null and alternative hypotheses below are for the third research question.

① Hypotheses about the relationship between the PIMES and the managerial accountability of South Korean public institutions

Null hypothesis H0$_1$: The evaluation results of PIMES have no effect on managerial accountability as measured by the ratio of profit to assets for South Korean public institutions.

Alternative hypothesis H1$_1$: The evaluation results of PIMES have a positive effect on managerial accountability as measured by the ratio of profit to assets for South Korean public institutions.

② Hypotheses about the relationship between the PIMES and the social accountability of South Korean public institutions

Null hypothesis H0$_2$: The evaluation results of PIMES have no effect on social accountability as measured by the customer satisfaction score for South Korean public institutions.

Alternative hypothesis H1$_2$: The evaluation results of PIMES have a positive effect on social accountability as measured by the customer satisfaction score for South Korean public institutions.

③ Hypotheses about the relationship between the PIMES and the strategic behaviour of South Korean public institution from the perspective of principal-agent theory

Null hypothesis H0$_3$: The evaluation results of PIMES have no effect on strategic behaviour of South Korean public institutions as measured by the average welfare benefit of regular employees of South Korean public institutions.

Alternative hypothesis H1$_3$: The evaluation results of PIMES have a negative effect on the strategic behaviour of South Korean public institution as measured by the average welfare benefit of regular

employees of South Korean public institutions.

4.5.3 Definitions of variables

This study now sets out the definitions of dependent variables, independent variable and control variables. It explains the theoretical basis for the independent and control variables and their expected signs of effect (positive or negative). For convenience this information is also set out in Table 4-1 below.

Table 4-1: Variables for quantitative analysis

Category		Name of variable	Expected sign	
Dependent variables		*PROFIT* (profit/assets of a public institution)	-	
		SATIS (customer satisfaction score of a public institution)		
		WELFARE (average welfare benefit of regular employees)		
Independent variable		*RANK* (PIMES standardised rank of a public institution)	*PROFIT*	Positive
			SATIS	Positive
			WELFARE	Negative
Control variables	Organisation character	*TYPE* (public corporation 1 vs quasi-governmental institution 0)	*PROFIT*	Positive
			SATIS	Positive
			WELFARE	Negative
		LISTED (listed on the South Korean Stock Exchange 1 vs not 0)	*PROFIT*	Positive
			SATIS	Positive
			WELFARE	Negative
	Personnel character	*SIZE* (number of regular employees)	*PROFIT*	Positive
			SATIS	Positive
			WELFARE	Positive
		PAST (past career of head is a government official 1 vs not 0)	*PROFIT*	Positive
			SATIS	Positive
			WELFARE	Negative
	Finance character	*WAGE* (average wage of regular employees)	*PROFIT*	Negative
			SATIS	Positive
			WELFARE	Positive

Dependent variables

The first research question is divided into two sub questions. First, do better PIMES evaluation results have a positive effect on managerial accountability and second, do better PIMES evaluation results have a positive effect on social accountability. According to Day and Klein (1987), managerial accountability can be sub-divided into regularity, efficiency and effectiveness accountability. Considering the NPM reforms' aim to improve efficiency of the public sector, this research focuses on efficiency accountability. Power (1997) defines efficiency accountability as accountability for ensuring that maximum output is achieved from the given resources or that minimum resources are used to accomplish a given level of output.

Public institutions providing public services are generally constrained on whether, and by how much they can raise the price of their services by central or local government and constrained in increasing the extent of the services they provide due to limitation of area where they are providing the services. However, the public institutions can increase their profit by lowering overall cost, for example by increasing productivity or reducing welfare benefit which consists of scholarship, medical expense and severance pay provided to employees by public institutions. Therefore, the greater efficiency accountability a public institution providing public services has, the less resources are used to accomplish a given level of output and then profit of the public institution increases. Hence, profit of a public institution can be used in order to operationalise efficiency accountability. However, the absolute profit of a public institution is largely determined by the size of the public institution, and so the 'ratio of profit to assets' is employed to measure the managerial accountability related to efficiency. This thesis builds on Choi and Hong's (2017) 'ratio of profit to assets' to examine the

effect of PIMES evaluation on the managerial accountability of public institutions. The 'ratio of profit to assets' is the first dependent variable in this study. This dependent variable (Y_{1it}) is a continuous variable and the scale of it is 'ratio'. This study names this dependent variable *PROFIT*.

Social accountability refers to more direct and explicit accountability relations between public agencies and citizens (Bovens, 2010). As explained in Chapter 2, the key outcome of social accountability is improved provision of public services (Grandvoinnet et al., 2015; Hickey and King, 2016). NPM reforms view citizens as clients who reward good public service by loyalty or punish poor public service by choosing to exit and hence citizens provide information to promote better social accountability by assessing public service delivery through loyalty and exit (Grandvoinnet et al., 2015). The South Korean government announces a customer satisfaction level for most of the public institutions that are evaluated in the PIMES annually, which is to some extent designed according to the principle of NPM (Singh and Slack, 2022). Hence, the customer satisfaction level of a public institution can be seen as providing information on its social accountability. Therefore, the regression estimated adopts the 'score of customer satisfaction level' as the second dependent variable, using the variable as a measure of social accountability. This study names the second dependent variable *SATIS*.

The effect of performance management systems on the 'score of customer satisfaction level' has not been estimated in previous literature. Therefore, the estimation in this thesis fills a gap in literature on the effect of performance management systems on social accountability of public organisations. The second dependent variable (Y_{2it}) is a continuous variable and the scale of it is 'ratio' because the score of customer satisfaction level is announced by the MOEF as a number between 0 and 100. The

higher *SATIS* is, the better is the level of customer satisfaction.

The third dependent variable is designed to operationalise the strategic behaviour of South Korean public institutions from the perspective of principal-agent theory. Choi and Hong (2017, p. 75) suggest that if the risk of bankruptcy disappears because of government's safety-net support, public institutions would have an incentive to increase their expenditure including labour costs and welfare benefits above an optimal level. Based on the suggestion, Choi and Hong (2017) estimated the effect of expenditure on welfare benefits on the PIMES results with panel date for 2008-2013 to examine whether the PIMES is designed in such a way that it can penalise the moral hazard of public institutions by delivering a lower PIMES ranking to institutions that give their employees higher welfare benefits. Another way of conceptualising this problem which is adopted in this thesisis, as argued by Kornai, Maskin and Roland (2003) that public institutions can be seen as facing a soft budget constraint (SBC). A SBC occurs if one or more supporting organizations are ready to cover the deficit of the SBC organization. Kornai et al. (2003) note that the idea of a SBC was initially developed by the study of socialism and applied to problems of post-socialist transition. However, the SBC effect also occurs in the public sectors of non-socialist countries according to many empirical studies (Kornai et al., 2003). For example, various non-profit organizations such as hospitals, schools and universities have not been permitted to go bankrupt because their deficits have been covered out of the state budget (Kornai et al., 2003). Kornai et al. (2003) suggest that BC organizations anticipate being rescued from bankruptcy. This anticipation usually results in distortion of their behaviour such as an attenuation of effort to reduce costs and overinvestment in risky ventures.

Therefore, there is always the possibility for a public institution to commit moral hazard such as

increasing the welfare benefits of employees paid by the public institutions excessively and conducting risky investment. The excessively increased welfare benefits and risky investment especially lead to exploiting public funding and so the South Korean government has endeavoured through the PIMES to incentivise institutions to keep the welfare benefits under control and also to lower the ratio of debt to equity to combat the intrinsic moral hazard of public institutions (MOEF, 2013). In contrast with the research of Choi and Hong (2017), but instead applying the theory of the soft budget constraint, this study adopts the average welfare benefit of regular employees as the third dependent variable regarding the strategic behaviour of public institutions from the perspective of principal-agent theory. The dependent variable (Y_{3it}) is continuous variable and the scale of it is 'ratio'. This study names the third dependent variable *WELFARE*.

Independent variable and control variables

The principal independent variable of this quantitative approach is the PIMES evaluation results. This study also includes a number of control variables. The PIMES evaluation results are announced as overall grades which consist of S, A, B, C, D, E in diminishing order of excellence. The order of overall grades means that a public institution with 'A' receives a higher rank than a public institution with 'B' and the overall grade of 'S' refers to super grade. The majority of overall grades are located in B or C (see Table 4-2: Distribution of overall grade in evaluation of performance in 2019). Therefore, it is reasonable that the overall grade could be transformed into a standardised rank, which is done in this study, in order to subdivide the same overall grade. This transformation of overall grade into the standardised rank follows a rank transformation in which each ordinal outcome score is transformed into the rank of the score (Grace-Martin, 2020). The overall quantitative approach of

this research drew heavily on the approach of Choi and Hong (2017). In particular, the linear regression of this research uses Choi and Hong (2017)'s definition of the standardised rank of a public institution as the independent variable.

As explained in Chapter 3, in the PIMES evaluation, the public institutions are placed in five categories: (i) public corporation I, (ii) public corporation II, (iii) fund-management-type institution, (iv) commissioned-service-type institution, and (v) strong and small institution. Each institution is ranked in one of the five categories. Choi and Hong (2017) define the standardised rank of a public institution as an institution's rank in its category divided by the number of institutions in its category multiplied by 100.

i.e. for a particular year

Choi and Hong's standardised rank for institution i in category j = $(rank_{i,j} / N_j) \times 100$

Where N_j is the number of institutions in category j and $rank_{i,j}$ is the institution's position in order of merit in the N_j set of institutions.

Thus, if there are 50 institutions in a category, Choi and Hong's standardised rank for the lowest ranking institution in its category would be $50/50 \times 100 = 100$. The top ranking institution would have a standardised rank of $1/50 \times 100 = 2$. Therefore, the closer to 100 is Choi and Hong's standardised rank, the worse performing is the institution.

In the study carried out by this thesis, independent variable (*RANK*) indicates the reverse rank of a public institution in its category contrary to the Choi and Hong (2017)'s above definition. Why is the ranking reversed? This is done following Choi and Hong (2017) who also used the reverse standardised ranking in their regression estimations, so that the higher the numerical value of *RANK* (with an upper limit of 100), the better performing is the institution, as indicated by the PIMES. Consequently, the independent variable (X_{it}) can be expressed as follows.

X_{it} = Standardised rank of institution i at time t

$$\frac{number\ of\ public\ institutions\ in\ i's\ category\ at\ t - rank\ of\ i\ in\ i's\ category\ at\ t}{number\ of\ public\ institutions\ in\ i's\ category\ at\ t} = \quad \times 100$$

(i: a public institution, t: fiscal year)

Thus, if there are 50 institutions in a category, the standardised rank for the lowest ranking institution in its category would be (50-50)/50 ⨯ 100 = 0 in this study. The top ranking institution would have a standardised rank of (50-1)/50 ⨯ 100 = 98. Thus, the higher the standardised rank (*RANK*) calculated, the better performing is the institution.

The researcher calculated a public institution's ranking in its category by using the evaluation reports that are released on the South Korean government official website of www.alio.go.kr. As explained in Chapter 3, the PIMES evaluates the public institutions through between 22 and 30 performance indicators as of 2019, which consist of quantitative and qualitative indicators (see Table 3-4). The evaluation reports show the scores of quantitative indicators and grades of qualitative indicators of

all the public institutions. However, the reports do not show the total score of each public institution, which is necessary for calculating each public institution's ranking in its category. Therefore, the researcher calculated the total score of each public institution by transforming its grades for qualitative indicators into scores and adding them to its scores of quantitative indicators.

The grades of qualitative indicators consist of A^+, A^0, B^+, B^0, C, D^+, D^0, E^+, E^0 in diminishing order of excellence, and are transformed into scores of 100, 90, 80, 70, 60, 50, 40, 30, 20 respectively in accordance with the guidance on management evaluation of public institutions (MOEF, 2018). Each of the qualitative indicators is given a *weight* by the MOEF and these weights range from 1 to 12. The final score of a qualitative indicator is obtained by multiplying the above transformed score by the appropriate weight and dividing by 100. For example, if the weight and grade of a qualitative indicator of a public institution are '2' and 'B^0', its final score would be $2 \times 70/100 = 1.4$. A public institution's total score in the PIMES evaluation is obtained by adding together all the scores of its qualitative indicators and adding this to the scores of the quantitative indicators taken from the PIMES evaluation reports. The researcher calculated all the public institutions' total scores and ranked them in their categories in order of their scores.

Taking the public corporations of 'G' and 'H' in the category of public corporation I as examples, both public corporations received 'B' as overall grade in the PIMES evaluation in 2019. On the other hand, their rankings were '5th' and '7th' respectively in the category of public corporation I according to the above calculation. There were 10 public corporations in the category of public corporation I in 2019. Consequently, the standardised ranks of 'G' and 'H' are '$(10-5)/10 \times 100 = 50$' and '$(10-7)/10$

$\times$ 100 = 30' respectively according to the formula. So, if 10 public institutions are ranked first, second,

third… to tenth, their standardised ranks would be 90, 80, 70… to zero. In addition, the public

corporations of 'I', 'J', 'K' in the category of public corporation II received 'A', 'B', 'C' as overall

grade in the PIMES evaluation in 2019 respectively. On the other hand, their rankings were '2nd',

'7th', '16th' respectively in the category according to the above calculation where qualitative and

quantitative factors were transformed into a total score. There were 26 public corporations in the

category of public corporation II in 2019. Consequently, their standardised ranks are '(26-2)/26 $\times$ 100

= 92.31', '(26-7)/26 $\times$ 100 = 73.08', '(26-16)/26 $\times$ 100 = 38.46' respectively according to the formula.

Thus, overall grades of S, A, B, C, D, E were not used when the researcher calculated the standardised

ranks of public institutions.

The distance of standardised ranks of public institutions is determined by the number of public

institutions in their categories. Applying this construction on the scores allows an independent

variable (X_{it}) to become a continuous variable having a measured continuum. The scale of the

independent variable is 'ratio' because the score of standardised rank is from 0 to just less than 100.

This study names this independent variable *RANK*. The larger *RANK* is, the better is the PIMES

evaluation result. If the evaluation result of a public institution improves for the current year, the

difference between *RANKs* in the current year and the previous year ($X_{it} - X_{it-1}$) is positive (+).

Based on the functions of PIMES (see Chapter 3), this study makes a prediction about the independent

variable's hypothesised effect (positive or negative) on the three dependent variables. Firstly, it is

hypothesised that *RANK* would have a positive effect on *PROFIT* because the PIMES aims to improve

management efficiency of South Korean public institutions (MOEF, 2014). Secondly, it is hypothesised that *RANK* would have a positive effect on *SATIS* because one of the purposes of PIMES is to improve the quality of public services (MOEF, 2014). Thirdly, it is hypothesised that *RANK* would have a negative effect on *WELFARE* because the MOEF has endeavoured to restrain the welfare benefits of public institutions through the PIMES (MOEF, 2013). With respect to excessive welfare benefit, the MOEF (2013) released some examples such as a public institution's support of medical expense up to 5 million won (3,100 pound) per employee in a year.

Table 4-2: Distribution of overall grade in evaluation of performance in 2019

Overall grade	S	A	B	C	D	E
Ratio of overall grade	0%	16.3%	39.5%	31.0%	12.4%	0.8%

Sources: News release 'Announcement of evaluation result on public institution management performance in 2019' (MOEF, 2020)

Other factors in addition to *RANK* can affect *PROFIT*, *SATIS* and *WELFARE*. If *PROFIT*, *SATIS* and *WELFARE* are influenced by these other factors than *RANK*, then it is important to control for them when it comes to interpreting the relationship between independent and dependent variables, because omitting these variables may cause specification bias (Punch, 2014). Researchers can control for a variable by holding its value constant and the variable controlled is called a control variable (Agresti and Finlay, 2009). In this research, control variables are placed into three categories: organisation character-related factors, personnel character-related factors and finance character-related factor. Control variables that have a ratio scale are expressed X_{2it}, X_{3it} and control variables that have binary scale are indicated by dummy variables (D_{1it}, D_{2it}, D_{3it}), which take the value one for some

observations and zero for the remaining observations (Greene, 2012, p. 189).

Firstly, the organisation character-related factors include type and status in the South Korean stock exchange of a public institution. As mentioned in Chapter 3, the South Korean public institutions are divided into public corporations and quasi-governmental institutions and the performance indicators for public corporations are different from those for quasi-governmental institutions. The performance indicators may affect *PROFIT*, *SATIS* and *WELFARE*, hence whether a public institution is a public corporation or a quasi-governmental institution should be controlled in this study's regressions. The control variable about type of a public institution is set up as a dummy variable (D_1) which is coded as 1 (public corporation) or 0 (quasi-governmental institution). This control variable is named *TYPE* in this study. As explained in Chapter 3, the self-generating revenue of public corporation reaches or exceeds 50% of the amount of its total revenue and hence the public corporations make more effort to increase self-generation revenue than the quasi-governmental institutions. As a result, the public corporations are expected to have more *PROFIT* than the quasi-governmental institutions. Moreover, the public corporations are better known to the public and are monitored more by the mass media than the quasi-governmental institutions. Consequently, the public corporations are expected to make more effort to increase *SATIS* and reduce *WELFARE* than the quasi-governmental institutions. Therefore, it is hypothesised that *TYPE* would have a positive relationship with *PROFIT* and *SATIS*, and, a negative relationship with *WELFARE*.

Some public institutions are listed on the South Korean Stock Exchange. For example, 13 public institutions among the 129 public institutions evaluated in the PIMES were listed on the South Korean Stock Exchange as of 2019. If a public institution is listed on a stock exchange, the public institution

is hypothesised to be more likely to endeavour to manage more efficiently since it is faced with shareholder pressure. As a result, whether a public institution is listed on the South Korean Stock Exchange can affect *PROFIT*, *SATIS* and *WELFARE*. Therefore, it should be controlled in the regressions. The control variable about the status in the South Korean Stock Exchange is set up as a dummy variable, so it (D_2) is coded as 1 (listed) or 0 (not listed). The control variable is named *LISTED* in this study. Considering shareholder pressure in the stock exchange, *LISTED* is expected to have a positive relationship with *PROFIT* and *SATIS*, but a negative relationship with *WELFARE*.

Secondly, the personnel character-related factors include the number of regular employees and the past career of a public institution's head. An organisation's size can be conceptualised as the number of employees of the organisation and is identified as an important control variable in public service performance studies (Walker, Andrews, George and Tu, 2023). Choi and Hong (2017) also argue that a public institution's size measured by the number of its employees may have an effect on the PIMES evaluation result, although in principle the result should be unaffected by its size because it is irrelevant to the efforts of the institution. In this study, the number of regular employees is used as a control variable (X_{2it}) and named *SIZE*. *SIZE* is a continuous variable and the scale of it is 'ratio'. According to the perspective of economies of scale, larger organisations can be more efficient than smaller organisations because fixed costs of production are spread across more units of output in a bigger organisation than a smaller one (Walker et al., 2023). Considering the perspective of economies of scale, *SIZE* is expected to have a positive relationship with *PROFIT*. *SIZE* also can affect *SATIS* and *WELFARE* because the more employees, the more advantageous it is to improve the quality of their services and negotiate for better welfare benefit. Therefore, it is hypothesised that *SIZE* would have a positive relationship with *SATIS* and *WELFARE*.

It can be argued that the past career of an institution's head could have an effect on *PROFIT*, *SATIS* and *WELFARE* of the institution. According to Choi and Hong (2017), if an institution's head was previously a government official, that would enable better management ability. Here is argue that that could lead the institution to increase *PROFIT* and *SATIS* and reduce *WELFARE*. Alternatively, a former official's better networking with government could produce that effect. Following the research of Choi and Hong (2017), this study therefore adopts the past career of an institution's head as a control variable (D_3) which is coded as 1 (has experience as a government official) or 0 (otherwise). This control variable is named *PAST* in this study. It is hypothesised that *PAST* is expected in this present study to have a positive relationship with *PROFIT* and *SATIS*, and a negative relationship with *WELFARE*.

Thirdly, a finance character-related factor is the average wage of regular employees in a public institution. This study employs the average wage of regular employees as a control variable ($X3_{it}$) and names the control variable *WAGE*. *WAGE* is a continuous variable and the scale of it is 'ratio'. A higher *WAGE* of an institution results in higher cost and lower profit of the institution. Therefore, it is hypothesised that WAGE would have a negative relationship with PROFIT. Choi and Hong (2017) suggest that employees' ability is reflected on their average wage. In other words, higher WAGE of an institution indicates, better ability of its employees and hence can result in better quality of public service which the institution provides. Consequently, *WAGE* is expected to have a positive relationship with *SATIS*. Employees with better ability are more likely to demand a higher *WELFARE* suitable for their ability, therefore, *WAGE* is expected to have a positive relationship with *WELFARE*.

4.5.4 Constructing a model for analysis: the linear regression model

In order to test the three hypotheses regarding the first and third research questions, a linear regression model that can handle both continuous and categorical variables is designed as follows, which is based on the dependent, independent and control variables.

Table 4-3: Analytical model for verifying the factors which affect the accountability and strategic behaviour of public institutions

$$Y_{it} = \alpha + \beta_1 X_{1it} + (\beta_2 D_{1it} + \beta_3 D_{2it} + \beta_4 X_{2it} + \beta_5 D_{3it} + \beta_6 X_{3it}) + \varepsilon_{it}$$

<Dependent variables>
· Y_{1it}: Profit/Assets of i at t
· Y_{2it}: Customer satisfaction score of i at t
· Y_{3it}: Average welfare benefit of regular employees of i at t

<Independent variable>
· X_{1it}: PIMES standardised rank of institution i at time t

$$* X_{1it} = \text{Standardised rank of i at t}$$
$$= \frac{number\ of\ public\ institutions\ in\ i's\ category\ at\ t - rank\ of\ i\ in\ i's\ category\ at\ t}{number\ of\ public\ institutions\ in\ i's\ category\ at\ t} \times 100$$
(i: a public institution, t: fiscal year)

<Control variables>

Organisation character-related factors
· D_{1it}: Type (public corporation/quasi-governmental institution: public corporation 1, quasi-governmental institution 0)
· D_{2it}: Stock (listed on the South Korean Stock Exchange or not: listed 1, not 0)

Personnel character-related factors
· X_{2it}: Number of regular employees of public institution of i at t
· D_{3it}: Past career of public institution's head (government official or not: government official 1, not 0)

Finance character-related factor
· X_{3it}: Average wage of regular employees of i at t

<Others>
· α and β are the constant and unstandardised coefficients in the linear regression equation
· ε_{it}: denotes an error term

4.5.5 Data collection

The data employed in this quantitative approach is secondary data that includes data generated through systematic reviews, documentary analysis as well as results from large-scale datasets (Smith, 2008). Among the secondary data, this quantitative research focuses on government surveys that can often represent the highest quality data because the high-level funding and expertise go into the government surveys (Smith, 2008). As the government surveys are often produced at regular time periods (Smith, 2008), the evaluation results and the values of performance indicators of PIMES are announced on the website of www.alio.go.kr annually, which is the South Korean government official website for public institutions. The data on *PROFIT*, *SATIS*, *WELFARE*, *RANK, TYPE, SIZE* and *WAGE* were collected from the website. In addition, this quantitative approach uses such data for a number of years which enables the study to look at continuity and change in behaviour over time (Smith, 2008).

The population for study in this research is all the public institutions which had been evaluated in the PIMES. The PIMES was significantly reformed in 2013, and it is difficult to compare *RANK* after 2014 with *RANK* before 2013. Therefore, in this research the target population which means the actual focus or aim of the research (Saunders, Thornhill and Lewis, 2019) is limited to the public institutions which have been evaluated in the PIMES after 2014. It is possible to collect data from this entire target population as the evaluation results and the values of performance indicators of PIMES are published on the government official website, therefore, the sampling technique is not required in this quantitative approach. The number of public institutions evaluated in the PIMES for 2014-2019 is shown in Table 4-4: Number of subjects for the quantitative approach. Adding these numbers, the

number of subjects for this quantitative approach is 731.

Table 4-4: Number of subjects for the quantitative approach

Year	Number of evaluated public institutions		
	Public corporation	Quasi-governmental institution	Sum
2014	30	86	116
2015	30	86	116
2016	30	89	119
2017	35	88	123
2018	35	93	128
2019	36	93	129
Total Sum	196	535	731

4.5.6 Data analysis

The study of the relationships between variables and accounting for variance are two sides of the same coin because knowing how much variance a researcher can account for gives the researcher a strong indication of how important the relationship is (Punch, 2014). In order to test the hypotheses regarding the relationships between variables, this research employs linear regression models and uses Stata program to carry out estimations. The value of coefficients of β_1, β_2, β_3, β_4, β_5, β_6 and statistical significance level have been estimated through the Stata program. In order to examine the statistical significance of the regression models, the F-values to test the overall impact of the independent and control variables and t-values testing the impact of each explanatory variable have

been computed.

4.6 The qualitative approach

4.6.1 Introduction

The purpose of the qualitative approach in this research is to examine the effect of design of performance management systems on public accountability and performance and explore strategic behaviour and gaming of public organisations. A qualitative approach is the most appropriate for the purpose because of the following important features of qualitative research. The first feature is that qualitative research enables a researcher to study people, things and events in their natural settings (Punch, 2014; O'Leary, 2017; Creswell and Creswell, 2018). In order to examine the effect of design of performance management systems and explore strategic behaviour and gaming of public organisations, it is necessary to collect data through events and people that are involved in the natural setting of performance management systems. As the second research question (*How do the conditions and processes of performance management systems affect the accountability of public organisations?*) is about the 'events' in the PIMES evaluation process, a qualitative research design is appropriate for addressing this question. The third research question (*What is the impact of performance management systems on the strategic behaviour of public organisations?*) is about the 'people' who are stakeholders in performance management systems, therefore, a qualitative research design can give answers to this question.

The second feature is that the researcher's role in qualitative research is to gain a holistic picture of the context (Punch, 2014; Creswell and Creswell, 2018). Performance management systems consist of several operational methods such as incentive schemes, performance indicators and target-setting methods, therefore, answering the second research question requires a holistic picture of the effect of the operation methods on public accountability and performance. In this research, a qualitative approach presents various experiences and opinions about the operational methods of PIMES through semi-structured interviews with different stakeholders of PIMES. The experiences and opinions enable the researcher to gain a holistic picture of the operational methods of PIMES and its effect.

The third feature is that qualitative researchers typically gather multiple forms of data such as interviews, observation and documents rather than rely on a single data source (Creswell and Creswell, 2018). In order to investigate strategic behaviour of public organisations, it is necessary to collect multiple forms of data regarding strategic behaviour of public organisations. Documents as well as semi-structured interviews are used in this qualitative approach in order to triangulate the data. This is particularly important here due to the possibility that the employees of public institutions are unlikely to present much information about immoral behaviour of public institutions. This section describes sources of data, sampling, data collection and analysis and ethics of this qualitative approach.

4.6.2 Sources of data

Documents and interviews are the qualitative data sources used in this research. This qualitative

approach collected and analysed public documents such as the South Korean government's evaluation reports and guidance on management evaluation and newspapers regarding the South Korean public institutions. This is because documents such as theses have the potential to inform (and indicate) the decisions which people make on a daily and long-term basis (May, 2011). However, documents are not produced to answer particular research questions and may have bias inherent in the classification and coding systems that they employ (O'Leary, 2017). For these reasons, this study uses documents in order to corroborate information and help form interview protocol.

As the interview is a very good way of accessing people's perceptions of situations and constructions of reality (Punch, 2014), the main source of data in this research is interviews. May (2011) suggests that there are four types of interviews used in social research: structured, semi-structured, unstructured or focused, and group interviews. The structured interview is associated with survey research in that each participant is asked the same question in the same way and comparability between responses is permitted and there is little room for variation in response (May, 2011; Punch, 2014). The semi-structured interview represents an opening up of interview method to understand how interviewees generate meanings in social life (May, 2011). This type of interview enables an interviewer to have more latitude to probe beyond the answers and compare the information that is generated in each interview (May, 2011). The unstructured interview has open-ended character and provides qualitative depth and a greater understanding of the subject without imposing a priori categorization that might limit the inquiry (May, 2011; Punch, 2014). Among the four types of interviews, the semi-structured interview is the most appropriate for this qualitative approach because it enables the researcher to elicit different stakeholders' opinions about the effect of PIMES, which is difficult to obtain by the structured interview (May, 2011). Furthermore, a semi-structured interview allows the researcher to examine the individual effect of operational methods of PIMES on

the accountability and performance of public institutions through comparison of the opinions, which is almost impossible in the unstructured interview (May, 2011; Punch, 2014).

4.6.3 Sampling

O'Leary (2017, p. 203) defines sampling as "the process of selecting elements of a population for inclusion in a research study". Sampling is as important in qualitative research as it is in quantitative research (Punch, 2014). In this qualitative approach, sampling is required for the semi-structured interviews. Target interviewees for the semi-structured interview included the evaluators of PIMES, the government officials of MOEF and the employees of public institutions, in order to access the various opinions of different stakeholders of PIMES.

Firstly, the evaluators refine some performance indicators submitted by public institutions as well as evaluate the public institutions in the PIMES. Therefore, the evaluators can present expertise and opinions about both performance indicators and operational methods of PIMES. Secondly, the governmental officials design the operational methods of PIMES and are responsible for operating the PIMES evaluation and hence, interviews with them enable this qualitative research to access the opinions about the impact of operational methods of PIMES on accountability and performance of public institutions. Thirdly, the employees are evaluated in the PIMES and significantly influenced by the PIMES evaluation results. Thus, they are the most sensitive stakeholders to the performance indicators and the operational methods of PIMES and might behave strategically in order to achieve a good overall grade. Therefore, interviews with them help the researcher understand the effect of PIMES on the strategic behaviour of public institutions.

According to Blaikie (2010), sampling methods are divided into single-stage probability, multi-stage and single-stage non-probability. Among them this qualitative approach uses single-stage non-probability sampling because there is no need to give every population element a chance to being selected in this semi-structured interview (Blaikie, 2010). This is because every employee or government official does not have enough expertise and experience to provide opinions about the individual effect of operational methods of PIMES. This research's qualitative approach uses purposive sampling where participants are selected on the basis of the researcher's judgement of the relevance of the participant to answering the research question and also uses snowball sampling in which one member of a network is chosen and then the person can be asked to identify other members (Saunders and Townsend, 2018; Blaikie, 2010). Snowball sampling is an effective method to access potential relevant participants the researcher may not have been aware of prior to data collection (Cohen and Arieli, 2011). However, snowball sampling can lead to biased sampling because it is not random and selects participants on the basis of a participant's networks and connections (Browne, 2005). But, for this research the qualitative sample was purposive and so the risk of bias was not as acute.

The purposive sampling criterion used was that the interviewee has professional experience and knowledge of working in the PIMES. As the researcher had worked in the MOEF over eight years, the researcher has a professional network of government officials of the MOEF and the employees of South Korean public institutions. Using the professional network, the researcher could select the samples of the government officials of MOEF and the employees of Korean public institutions who have professional experience and knowledge of working in the PIMES. When selecting the sample of the evaluators of PIMES, the researcher asked the government officials of MOEF to recommend the appropriate evaluators who have much experience of evaluating in the PIMES. After selecting

some evaluators as interviewees, the researcher identified other evaluators by being recommended from the first evaluators.

Saunders and Townsend (2016) suggest that the number of interviewees depends on the balance between representativeness and quality of responses in obtaining sufficient information. With respect to how many participants are likely to be needed within qualitative research, scholars suggest that data collection should continue until saturation is reached (Guest, Bunce and Johnson, 2006; Saunders and Townsend, 2018). The criterion of saturation has been used as an important standard when researchers determine purposive sampling size (Guest et al. 2006). Theoretical saturation occurs when all the main variations of the phenomenon under study have been identified and incorporated into an emerging theory (Guest et al., 2006). This research determined the number of interviewees based on the standard of saturation.

As mentioned earlier, there are five categories in the PIEMS evaluation: (i) public corporation I, (ii) public corporation II, (iii) fund-management-type institution, (iv) commissioned-service-type institution, and (v) strong and small institution. In order to identify main effects of operational methods of PIMES on the accountability and performance of public institutions in each category, the researcher selected five interviewees from each category. Guest et al. (2006) argue that data saturation occurred for the most their research parts when they had analysed 12 interviews. Considering the argument of Guest et al. (2006), this research selected 12 evaluators with professional experience and knowledge of working in the PIMES evaluation. As one or two government officials of MOEF are responsible for one of the five categories respectively, this research selected five government officials as interviewees considering the five categories. Consequently, this research selected 42 interview

participants (see Table 4-5: Classification of interviewees). This study indicates the employees, evaluators, and government officials as 'E', 'V', and 'G' respectively when quotations from the interviews with them are used in this thesis.

Table 4-5: Classification of interviewees

Group	Employees of public institutions (E)	Evaluators of PIMES (V)	Government officials of MOEF (G)	Total
Public corporations	10	12	5	42
Quasi-governmental institutions	15			
Total	25	12	5	

4.6.4 Data collection

Before starting interviews, pilot interviews were undertaken from the UK through video software of Zoom during 19th and 20th January 2021 to increase the validity of interview questions. Pilot interviews can indicate questions that need expanding or strengthening in some way (Gillham, 2007). As the purpose of this semi-structured interview was to obtain a rich understanding, it was necessary to expand or strengthen interview questions through these pilot interviews. The three groups of the interviewees represent different stakeholders in the PIMES, hence the pilot interviews consisted of a South Korean government official, an employee of South Korean public institution and a South Korean professor with professional experience and knowledge of working in the PIMES. The pilot interviews took between one and two hours. The pilot interview resulted in many suggestions that

were useful for improving interview questions and selecting interviewees.

The questions of the semi-structured interview are sorted into themes relating to the performance indicators, operational methods of PIMES and strategic behaviour of public institutions. The purpose of the interview questions about performance indicators is to understand how the structure of performance indicators in the PIMES evaluation affects the accountability and performance of South Korean public institutions. The interview questions about operational methods of PIMES aim to explore how the financial and non-financial incentive schemes and target-setting method in the PIMES affect the accountability and performance of South Korean public institutions. Whilst the interview questions about both the performance indicators and operational methods are designed to address the second research question, the purpose of the interview questions about the strategic behaviour of public institutions is to examine how far the strategic behaviour of South Korean public institutions is influenced by the PIMES.

Regarding the interview questions, the three pilot interviewees agreed that most interview questions were generally clear and easily understood. However, they suggested that some interview questions should be revised to elicit relevant experiences and opinions of the evaluators and employees of public institutions for the qualitative approach. Consequently, some interview questions were revised based on the suggestions of pilot interviewees. Firstly, the interview question regarding distorted behaviour of public institutions was revised by adding an example to show the distorted behaviour explicitly. As the employees of public institutions are reluctant to provide unfavourable information, the example might make the interviewees convenient when answering the interview question. Secondly, the interview question regarding the target-setting method was improved by adding an introductory

part that asks interviewees if they know how the targets are set out. Lastly, several follow-up questions were added to garner the specific opinions of interviewees.

After the pilot interviews, potential interviewees were selected by using personal professional network and searching public documents. As the number of employees of each public institution is over 100, the researcher asked colleagues in the MOEF to recommend the employees who had worked or were working in the PIMES. Selecting potential interviewees from evaluators group, the researcher reviewed press releases announced by the MOEF to identify the evaluators who had wide professional knowledge about the PIMES. After selecting potential interviewees, the researcher sent them a participant information sheet including the purpose of this research and the confidentiality of information collected during the interview and sent a consent form to them by email to obtain their consent to interviews.

Positionality is determined by the place where one stands in relation to others (Bourke, 2014). As the positionality of the researcher has the potential to impact research process (Bourke, 2014), it is important to reflect the positionality of researcher. The researcher's positionality of having worked for the MOEF before this research was beneficial for understanding performance of public institutions and having a good rapport with the interviewees. Moreover, the positionality was helpful for communicating with potential interviewees and obtaining the consent to interviews because most potential interviewees have trust in South Korean government officials. However, given the researcher is a South Korean government official, some employees may have been concerned about the possibility that if they discuss some problems of their institutions, that might possibly lead to their institutions being subject to unfavourable attention by the South Korean government. Moreover, some

employees may have tried to focus only on advantages of the PIMES evaluation and hidden drawbacks of the evaluation which they might know of. In order to overcome the potential bias, the researcher convinced interviewees of confidentiality before and during the interviews, and persuaded them to be open about possible drawbacks of the PIMES evaluation by comforting them. After obtaining the consent to interview, the researcher reserved the interview date and time and received a consent form from every interviewee by email. The interviews with 42 participants were conducted through the video software of Zoom from 1st February to 5th March 2021 considering the COVID 19 pandemic. Each interview lasted for about one hour and every interview was recorded through the video software of Zoom as every interviewee signed the consent form stating that interviews would be recorded. The recoding through the video software of Zoom was used to produce transcriptions of the 42 interviews. The video recording enables researchers to identify participants' nonverbal cues such as facial expressions and gestures (Archibald, Ambagtsheer, Casey and Lawless, 2019). When the researcher used the video recording to transcribe, it was sometimes possible to understand the implicit intention of interviewees through facial expressions. However, the video recording can cause security issue or privacy breaches (Archibald et al., 2019). One of my interviewees refused video recording and the researcher replaced it with voice recording after getting permission from the interviewee. All the interviewees were very cooperative and provided their experiences and opinions regarding the PIMES.

4.6.5 Approach to analysis

All the interview recordings were reviewed and 42 full interview transcripts were made for qualitative data analysis. In order to find patterns for addressing the second and third research questions, the

qualitative data analysis of this research uses thematic analysis that is a method identifying and analysing patterns within data (Braun and Clarke, 2006). The thematic analysis consists of six phases: (i) transcribing data, (ii) coding interesting features of data in a systematic fashion across the entire data set, (iii) collating codes into potential themes, (iv) checking if the themes work in relation to the coded extracts and the entire data set, (v) generating clear definitions for each theme, and (vi) producing a report (Braun and Clarke, 2006).

Coding as a starting point in qualitative analysis is essential to discover regularities in the data (Punch, 2014). This qualitative approach adopts concept codes that assign macrolevels of meaning to data or data analytic work in progress (Miles, Huberman and Saldaña, 2020). Concept codes are appropriate for studies focused on theory and transcending the local and particular of the study to more abstract or generalizable context (Saldaña, 2013). This research examines whether the effect of operational methods of PIMES and the strategic behaviour of public institutions link to theories hence concept codes are appropriate for the research. This qualitative approach conducted coding following the eight steps suggested by Tesch (1990): (i) reading all the transcriptions, (ii) picking the interviews that contain the experiences and opinions of interviewees, which are related to the research questions, (iii) making a list of topics that were identified in the literature review, (iv) abbreviating the topics as codes, (v) making total list of categories by grouping topics, (vi) deciding the abbreviation for each category and alphabetizing these codes, (vii) assembling the data belonging each category in one place, and (viii) performing a preliminary analysis, recoding your exiting data if necessary.

Memos that help shape the development of codes (Creswell and Creswell, 2018) were done simultaneously when the researcher conducted the coding, in order to move the qualitative data

analysis towards conceptualising and developing propositions (Punch, 2014). In the process of developing propositions, this qualitative approach followed the iterative nature of explanation building (Yin, 2018). In other words, the propositions of this research result from a series of iterations: making an initial but tentative explanatory proposition, comparing the qualitative data with the proposition, revising the proposition and repeating this process with the other qualitative data (Yin, 2018). NVivo 12 which might be the most widely used software in social science assisted the researcher in conducting the qualitative data analysis (Creswell and Creswell, 2018). The researcher uploaded quotes from the interview transcriptions to NVivo 12, which are relevant to topics as codes for each category. As the researcher uploaded quotes instead of full transcripts, it was feasible in this research to assemble the data belonging each category regarding relevant theory or concepts into NVivo 12, which is seventh step suggested by Tesch (1990). Based on the quotes, the researcher processed a series of iterations such as comparison of the quotes with the tentative proposition and revision of the proposition.

As explained in Chapter 2, this research analyses behaviour of public institutions by using principal-agent theory because actors' strategic behaviour and gaming can be discerned through principal-agent theory. Moreover, principal-agent theory presents how to design rewards and sanctions to mitigate strategic behaviour and gaming (Levacic, 2009; Gailmard, 2014). Principal-agent theory focuses on studying people and events, to which interview data is an appropriate method for understanding. This research ruled out action research which is sometimes done by other researchers and produces practical knowledge (Punch, 2014).

4.6.6 Ethical consideration

O'Leary (2017) suggests that ethical guidelines for the conduct of research generally cover informed consent from respondents, no harm to respondents, and ensuring confidentiality. The researcher received a signed consent form stating the purpose of interview, the future use of interview result, participant withdrawal, and confidentiality from all the interviewees before the interview commenced. When starting the interviews, the researcher explained the interview purpose, participant withdrawal, and confidentiality once more. In addition, the researcher received ethical approval from the University of Birmingham on 8th January 2021 before starting the interviews.

4.7 Conclusion

This chapter has presented the research design and methodology used in this thesis. The chapter has developed the three research questions based on the literature review and explained which research method is appropriate for each research question. It discussed the philosophical perspective of critical realism that forms the basis of case study design, mixed methods research and retroductive process of this research. This research has adopted the PIMES as a single case in order to examine the effect of design of performance management systems on public accountability. A convergent mixed methods research has been adopted in this research to answer the overarching research question of how performance management systems affect the accountability of public organisations.

The quantitative and qualitative approaches in this research have been explained in depth. For the quantitative approach, this chapter has defined and explained the dependent variable, independent variables and control variables, and constructed the multiple regression model used in this research. For the qualitative approach, this chapter has discussed why this research selected documents and interviews as data sources and 42 interviewees based on the standard of saturation, and has explained how the thematic analysis, coding and memos are conducted. The following chapters present the results of both quantitative and qualitative analysis based on the empirical data.

Chapter 5 The impact of PIMES *RANK* on the accountability and strategic behaviour of public institutions

5.1 Introduction

The purpose of this chapter is to examine the relationship between the PIMES and the accountability and strategic behaviour of South Korean public institutions. This accountability relationship is examined with the aid of the model set out in Chapter 4 and a set of data collected for the years from 2014 to 2019. The three hypotheses set out in Chapter 4 on accountability are tested and the results are analysed. Furthermore, the interviewee quotes from key participants in the PIMES evaluation are used in order to explore the quantitative analysis results. The interviewee quotes provide empirical evidence on why the PIMES evaluation has not significantly affected the managerial accountability and strategic behaviours of public institutions However, the PIMES has affected the social accountability of public institutions significantly.

As explained in Chapter 2, an accountability deficit can be expected in public service agencies because they are less exposed to political control than ministries (Mulgan, 2014; Day and Klein, 1987; Watt, 1998). The PIMES aims to improve the managerial and social accountability of South Korean public institutions based on the principles of NPM. It can be seen as a mechanism to address a likely accountability deficit (Ospina, Grau and Zaltsman, 2004; Lægreid, 2014). It is therefore useful to examine empirically to what extent the PIMES addresses the accountability deficit. The quantitative

analysis in this chapter sheds light on-both the first research question (*Do performance management systems improve the accountability of public organisations?*) and the general question of how well performance management systems based on the principles of NPM address an accountability deficit.

Among the types of strategic behaviour of public institutions set out in Chapter 2, this chapter focuses most on the strategic behaviour of public institutions in terms of moral hazard. As explained in Chapter 4, strategic behaviour in the form of moral hazard can lead to agents exploiting public funding through excessively increased welfare benefit and risky investment. Welfare benefit is used in this thesis to denote scholarship, medical expense and severance pay provided to employees by public institutions. the South Korean government has tried to control the strategic behaviour of public institutions in terms of moral hazard by structuring incentives to reduce excessive welfare benefits and debt of public institutions through the operation of PIMES. However, the current research aims to discern empirically to what extent the PIMES succeeds in controlling such strategic behaviour of public institutions regarding moral hazard. Using panel data, this chapter examines how the strategic behaviour of public institutions regarding moral hazard has played out over time.

The chapter begins with discussion of the data set. It deals with how the data set has been collected and what the variables in the model measure, including discussion of missing data for customer satisfaction scores. That lack of data is an important issue as such data would help an examination of the social accountability of public institutions. After addressing the missing data problem, a preliminary statistical analysis including correlation analysis and scattergrams between variables is presented to map out the broad relationships between independent, dependent and control variables. Based on the preliminary statistical analysis, this chapter then uses regression analysis in seeking to

clarify relationships between the PIMES and the managerial and social accountability and strategic behaviour of public institutions.

The regression analysis was used to examine the effect of PIMES evaluation results, measured by the variable *RANK* on the managerial and social accountability of South Korean public institutions, measured by the variables *PROFIT* and *SATIS*. The results of the regression analysis suggest that *RANK* has a significant positive effect on *SATIS*, however, *RANK* has only a weak positive effect on *PROFIT*. Lastly, this chapter examines the relationship between the PIMES and the strategic behaviour of public institutions in relation to moral hazard as measured by *WELFARE*. *RANK* was hypothesised to have a negative effect on *WELFARE* and such an effect was found, albeit weak.

5.2 Discussion of the data set

In order to provide a thorough analysis of the working of PIMES a large amount of data was collected. As explained in Chapter 3, the PIMES evaluation of management performance of public institutions is conducted annually and numerical data for public institutions are published on the South Korean government official website www.alio.go.kr. To analyse how the PIMES has worked, this study examined data on the evaluation results and other variables for 731 public institutions for 2014-2019. After considerable efforts it proved possible to assemble a large panel of key data for 2014-2019. This section discusses the variables collected in turn.

Firstly, data on the independent variable (*RANK*) was obtained by calculating ranking of the public

institutions in their categories and computing their reverse standardised ranking as explained in Chapter 4. As the *RANK* which arises from the PIMES evaluation each year is based on a public institution's assessed performance in a number of dimensions, as set out in Chapter 2, it is hypothesised that the PIMES incentivises the institution to try to improve its position (*RANK*) by controlling attributes of the institution such as *PROFIT* and *WELFARE*. Striving for *RANK* is hypothesised to provide financial and non-financial incentives to the public institutions.

Secondly, the first dependent variable (*PROFIT*) was generated by extracting the ratio of profit to assets for each public institution from the official website www.alio.go.kr. Some public institutions have several income statements and statements of financial position because they operate several funds. In those public institutions, the researcher summed the profit and summed assets from such institutions' several funds. For such institutions the ratio of profit to assets of the public institutions was calculated by dividing the institution's total profit by its total assets. This variable (*PROFIT*) was taken as a measure the managerial accountability of the public institution taking a view that managerial accountability relates to efficiency accountability and efficiency depends on determinants of profit such as higher output or lower cost (Day and Klein, 1987; Power, 1997). As the PIMES aims to improve management efficiency of public institutions, it is hypothesised that *RANK* would have a positive effect on *PROFIT*.

Thirdly, the customer satisfaction score (*SATIS*) was collected by looking up the institution's satisfaction score in the South Korean government's evaluation reports. This variable (*SATIS*) was taken to be a measure of the social accountability of public institutions. As explained in Chapter 4, social accountability means more direct accountability relations between public agencies and citizens

139

and higher level of social accountability relates to better quality of public services (Bovens, 2010; Grandvoinnet, et al., 2015). The South Korean government surveys citizens' level of satisfaction with public services provided by the majority of public institutions annually. Thus, the survey results consist of customer satisfaction scores that reflect citizens' evaluation of the quality of public services. Therefore, a higher customer satisfaction score relates to a better quality of service and thus can be seen as a measure of the social accountability of public institutions. As a main objective of PIMES is to improve the quality of public services, it is hypothesised that *RANK* would have a positive effect on *SATIS*.

Fourth, data on the average welfare benefit of regular employees *(WELFARE)* was extracted from the official website www.alio.go.kr. Welfare benefit in this thesis denotes scholarships, medical expenses and severance pay provided to employees by public institutions. This variable (*WELFARE*) was taken to measure the (undesirable) moral hazard of public institutions. As explained in Chapter 4, public institutions, in practice, face a soft budget constraint (SBC) because central or local government supports the public institutions to prevent bankruptcy (Kornai et al., 2003). Thus, protected from bankruptcy, public institutions may make excessive increases to the welfare benefits which employees press for. As a result, the average welfare benefit of regular employees is argued to relate positively to the strategic behaviour of public institutions constituting moral hazard. As the MOEF aims to restrain the welfare benefits of public institutions at an appropriate level through the PIMES, it is hypothesised that *RANK* would have a negative effect on *WELFARE*.

Lastly, data on the organisation, personnel and finance character-related factors of public institutions regarding control variables were collected as control variables from the South Korean government

and the public institutions' official websites. The type of public institution (*TYPE*) was taken to measure the effect of differing performance indicators between public corporations and quasi-governmental institutions on *PROFIT*, *SATIS* and *WELFARE*. Whether an institution is listed on the South Korean Stock Exchange (*LISTED*) was taken to measure the effect of listing on efficiency of an institution. As the public institutions listed on a stock exchange are more likely to increase efficiency due to shareholder pressure, *LISTED* is expected to have a positive relationship with *PROFIT* and *SATIS*, however, a negative relationship with *WELFARE*. Regarding personnel character-related factors, number of employees (*SIZE*) and past career of head (*PAST*) were taken to measure the effect of an institution's size and the effect of a head's past career as a government official. A larger organisation is more efficient than a smaller organisation if economies of scale operate, hence *SIZE* is expected to have a positive relationship with *PROFIT*. The past career of an institution's head being a government official is more likely to enable better management ability (Choi and Hong, 2017), or possibly better networking, therefore, *PAST* is expected to have a positive relationship with *PROFIT* and *SATIS*, however, a negative relationship with *WELFARE* is expected. The wage of employees (*WAGE*) was taken to measure the effect of average wage of an institution. As higher *WAGE* results in higher cost and lower profit, *WAGE* is expected to have a negative relationship with *PROFIT*. Employees' ability is likely to be reflected by their average wage and hence, the public institutions paying higher wage are more likely to have more qualified employees (Choi and Hong, 2017). Therefore, *WAGE* is expected to have a positive relationship with *SATIS*. The data set for 2014-2019 was collected over one and half years from various South Korean government's evaluation reports and extraction of data from the official websites.

Among the variables in the regression model, *SATIS* suffered from missing observations because about ten public institutions were exempted from the customer satisfaction survey annually. The

141

missing data pattern is an example of univariate missing data where 'missingness' is confined to a single variable (Little and Rubin, 2002). *SATIS* is the dependent variable for the regression model to examine the effect of PIMES evaluation on the social accountability of public institutions, therefore, it is very important to find an appropriate method to address the missing data problem regarding *SATIS*.

The possible methods to deal with the missing data problem can be grouped into four categories: procedures based on completely recorded units, weighting procedures, imputation-based and model-based procedures (Little and Rubin, 2002). The procedures based on completely recorded units are to discard incompletely recorded units and analyse only the units with complete date (Little and Rubin, 2002). The method is easy to conduct and may be satisfactory when there are small amounts of missing data (Little and Rubin, 2002). In contrast imputation-based procedures are methods to calculate missing values which are then filled in and the resultant completed data are analysed by standard methods (Little and Rubin, 2002). One of the procedures is mean imputation which substitutes the mean of the recorded values for missing values. It is crucial to distinguish missing data mechanisms which cause missing data because the properties of the methods to deal with missing data problem depend on the mechanisms very strongly (Little and Rubin, 2002).

The mechanisms are categorized into missing completely at random (MCAR), missing at random (MAR) and not missing at random (NMAR) (Wooldridge, 2020; Little and Rubin, 2002). MCAR means that missingness does not depend on the data values and the reason the data are missing is independent of observed and unobserved factors affecting the dependent variable (Little and Rubin, 2002; Wooldridge, 2020). If the missing data mechanism is MCAR, missing data cause no statistical

problems (Wooldridge, 2020). MAR means that missingness depends only on the observed components and does not depend on the components that are missing (Little and Rubin, 2002). In other words, the probability of missing data depends on independent variables in MAR (Wooldridge, 2020). NMAR means that missingness depends on the components that are missing (Little and Rubin, 2002). If the missing data mechanism is not MCAR, discarding incomplete cases in procedures based on completely recorded units can cause bias (Little and Rubin, 2002). Therefore, whether the missing data mechanism in *SATIS* is MCAR or not is crucial to determine the application of procedures based on completely recorded units.

Institutionally, missingness of *SATIS* occurs because a few public institutions are exempted from the customer satisfaction survey every year. According to interviews with the government officials of MOEF, the exemption occurs because for these particular institutions the vast majority of people in South Korea do not know of their public services. Consequently, the missingness of *SATIS* does not depend on the data values in the regression and hence the missing data mechanism in the regression is MCAR. In addition, the ratio of the exempted public institutions to all public institutions in the PIMES is under 0.12 from 2014 to 2019, thus there is only a small amount of missing data. Considering the mechanism of MCAR and small amounts of missing data, procedures based on completely recorded units are the most appropriate method to address the missing data problem. Discarding incompletely recorded observations and analysing only completely recorded observations can be seen as appropriate as it is unlikely to cause bias and this method was therefore adopted in this case.

This research therefore discarded 62 incompletely recorded units which did not have the customer

satisfaction score. As a result, the number of subjects for the regression of this research fall to 669 (see Table 5-1). Table 5-2presents descriptive statistics for the dependent, independent and control variables of the regression.

Table 5-1: Number of subjects for the regression

Year	Number of public institutions for the regression		
	Public corporation	Quasi-governmental institution	Sum
2014	24	79	103
2015	24	80	104
2016	24	84	108
2017	29	86	115
2018	29	90	119
2019	30	90	120
Total Sum	160	509	669

Table 5-2: Descriptive statistics for the dependent, independent and control variables

Variable	Mean	Standard deviation	Minimum	Maximum	VIF
PROFIT	0.01297	0.11019	-0.82235	0.95052	-
SATIS	91.84486	7.00075	0	100	-
WELFARE	2161.862	1176.498	207	6461	-
RANK	47.83528	28.81297	0	98.2	1.08
TYPE	0.23916	0.42689	0	1	1.35
LISTED	0.04335	0.20379	0	1	1.25
SIZE	1750.487	3856.258	61	32430	1.14
PAST	0.45441	0.49829	0	1	1.03
WAGE	68174.54	12387.31	36577	99557	1.16

5.3 Preliminary statistical analysis of variables

This section presents a preliminary statistical analysis of relationships between dependent variables (*PROFIT*, *SATIS*, *WELFARE*), independent variable (*RANK*) and control variables (*TYPE*, *LISTED*, *SIZE*, *PAST*, *WAGE*). The analysis then moves on to conducting the linear regression analysis. Subsequently, this section deals with multicollinearity which means a lot of overlap among independent and control variables (Agresti and Finlay, 2009). This is an important issue because if the multicollinearity exists in the regression model, it reduces the clarity of interpretation of regression coefficients in the subsequent discussion (Agresti and Finlay, 2009).

5.3.1 Correlation analysis of variables

The Pearson correlation coefficient (r) indicates the direction and strength of relationships between variables (Agresti and Finlay, 2009; Punch, 2014; Acock, 2018). The value of 'r' does not depend on units of the variables and the variables are positively and negatively related when $1 \geq r > 0$ and $-1 \leq r < 0$ respectively (Agresti and Finlay, 2009). The closer $\mid r \mid$ is to 1, the stronger the relationship is, however, 'r' does not suggest how steep the relationship is (Punch, 2014; Acock, 2018). In other words, 'r' measures the strength of linear association between variables and the regression coefficient indicates how steep the relationship between variables is (Agresti and Finlay, 2009).

Table 5-3 presents the Pearson bivariate correlation between the dependent variables (*PROFIT*,

146

SATIS, WELFARE), independent variable (*RANK*) and control variables (*TYPE*, *LISTED*, *SIZE*, *PAST*, *WAGE*). *RANK* is positively related to *PROFIT*, *SATIS, WELFARE*. However, the correlation between *RANK* and *SATIS* is much stronger than the other correlations. In addition, only the correlation between *SATIS* and *RANK* is statistically significant at the 0.05 level.

Table 5-3: Pearson correlation coefficient (r) between all the variables

	PROFIT	SATIS	WELFARE	RANK	TYPE	LISTED	SIZE	PAST	WAGE
PROFIT	1.0000								
SATIS	0.0176	1.0000							
WELFARE	0.0597	0.1446**	1.0000						
RANK	0.0266	0.1269**	0.0180	1.0000					
TYPE	-0.0059	0.0959**	0.5009**	-0.0767**	1.0000				
LISTED	0.0431	0.0014	0.2236**	-0.1532**	0.3797**	1.0000			
SIZE	0.0057	-0.0989**	0.1228**	0.0061	0.2853**	0.2915**	1.0000		
PAST	0.0534	-0.0114	0.1070**	0.1708**	-0.0120	-0.0026	0.0305	1.0000	
WAGE	0.0448	0.0278	0.5499**	0.0826**	0.3508**	0.1797**	0.0759**	0.0323	1.0000

Note: ** denotes statistically significant at the 0.05 level for a two-tailed test

Table 5-3 also presents the bivariate correlation between the three dependent variables and the five control variables respectively through 'r'. Firstly, the institution being a public corporation rather than a quasi-governmental institution (*TYPE*) is negatively correlated with *PROFIT*, however, it is positively correlated with *SATIS* and *WELFARE*. The correlation between *TYPE* and *PROFIT* is not statistically significant at the 0.05 level, however, the correlations between *TYPE* and *SATIS*, and *TYPE* and *WELFARE* are statistically significant at the 0.05 level.

Secondly, the institution being listed on the South Korean Stock Exchange (*LISTED*) is positively

related with all the dependent variables (*PROFIT, SATIS* and *WELFARE*). While the correlations between *LISTED* and *PROFIT,* and *LISTED* and *SATIS* are not statistically significant at the 0.05 level, the correlation between *LISTED* and *WELFARE* is statistically significant at the 0.05 level.

 Thirdly, the number of regular employees (*SIZE*) is positively related to *PROFIT* and *WELFARE*, however, negatively related to *SATIS*. The correlation between *SIZE* and *PROFIT* is not statistically significant at the 0.05 level, however, the correlations between *SIZE* and *SATIS*, and *SIZE* and *WELFARE* are statistically significant at the 0.05 level.

Fourth, the head being a government official in the past (*PAST*) is positively related to *PROFIT* and *WELFARE*, however, it is negatively related to *SATIS*. While the correlations between *PAST* and *PROFIT*, and *PAST* and *SATIS* are not statistically significant at the 0.05 level, the correlation between *PAST* and *WELFARE* is statistically significant at the 0.05 level.

Lastly, the average wage of regular employees (*WAGE*) is positively related with all the dependent variables (*PROFIT, SATIS* and *WELFARE*). While the correlations between *WAGE* and *PROFIT,* and *WAGE* and *SATIS* are not statistically significant at the 0.05 level, the correlation between *WAGE* and *WELFARE* is statistically significant at the 0.05 level.

5.3.2 Scattergram of variables

A scattergram is a graphic representation of the relationship between two variables (Acock, 2018). An explanatory (independent) variable is generally placed on the horizontal axis and an explained (dependent) variable is placed on the vertical axis in a scattergram (Acock, 2018). A pair of observations on two variables is shown as a point in a graph (Iversen and Gergen, 1997). A scattergram is a very common type of visualisation and has been used in a variety of exploratory contexts (Sarikaya and Gleicher, 2018).

Scattergrams are employed in this preliminary statistical analysis of variables for several reasons. The first is that it shows patterns in two variables (Iversen and Gergen, 1997). For example, if the path of points in a scattergram is from the lower left corner to the upper right corner, the scattergram indicates a positive relationship between the two variables (Iversen and Gergen, 1997). However, it is difficult to discern the relationship between the two variables when there are 500 or more observations (Acock, 2018). Despite that limitation, a scattergram can present a good sense of the relationship between the two variables in the large sample that consists of more than 500 observations through a fitted values line that is in fact a linear regression line (Acock, 2018). This is because a positively sloped fitted values line shows how the higher the observation in the explanatory variable is, the higher the observation in the explained variable is on average and vice versa for a negatively sloped fitted values line (Acock, 2018). The second reason is that no numerical information is lost and simplification of data is gained in a scattergram (Iversen and Gergen, 1997). All observations in a sample are shown as points and the points show whether the observations are distributed with a specific range.

Figure 5-1 shows the relationships between the three dependent variables *PROFIT*, *SATIS*, *WELFARE* and the independent variable *RANK* and the two control variables including *SIZE* and

WAGE through scattergrams. The other control variables such as *TYPE*, *LISTED* and *PAST* are dummy variables, and hence they are not suitable for scattergram. In the scattergrams *RANK*, *SIZE*, *WAGE* to *SATIS*, an influential observation (standardised rank is '0' and customer satisfaction score is '0') was removed because it is outlier which affects fitted values line. While all the nine fitted values lines in Figure 5-1 slope upward, the Pearson correlation coefficient (r) between *SIZE* and *SATIS* is negative and the other eight Pearson correlation coefficients (r) are positive (Agresti and Finlay, 2009). This difference occurs due to the outlier. If the outlier is deleted, the Pearson correlation coefficient (r) between *SIZE* and *SATIS* is positive (0.0702). Furthermore, the four fitted values lines in the scattergrams relating *RANK, SIZE* to *SATIS* and *SIZE, WAGE* to *WELFARE* are much steeper than the five fitted values lines in the scattergrams relating *RANK, SIZE, WAGE* to *PORFIT* and *WAGE* to *SATIS* and *RANK* to *WELFARE*.

Figure 5-1: Scattergrams between the three dependent variables and the independent variable and the two control variables
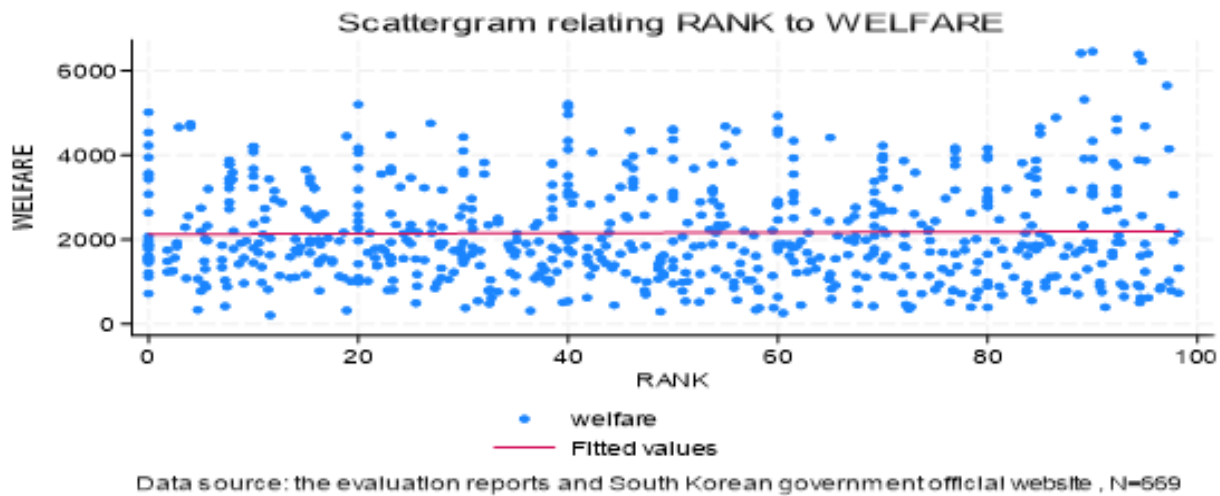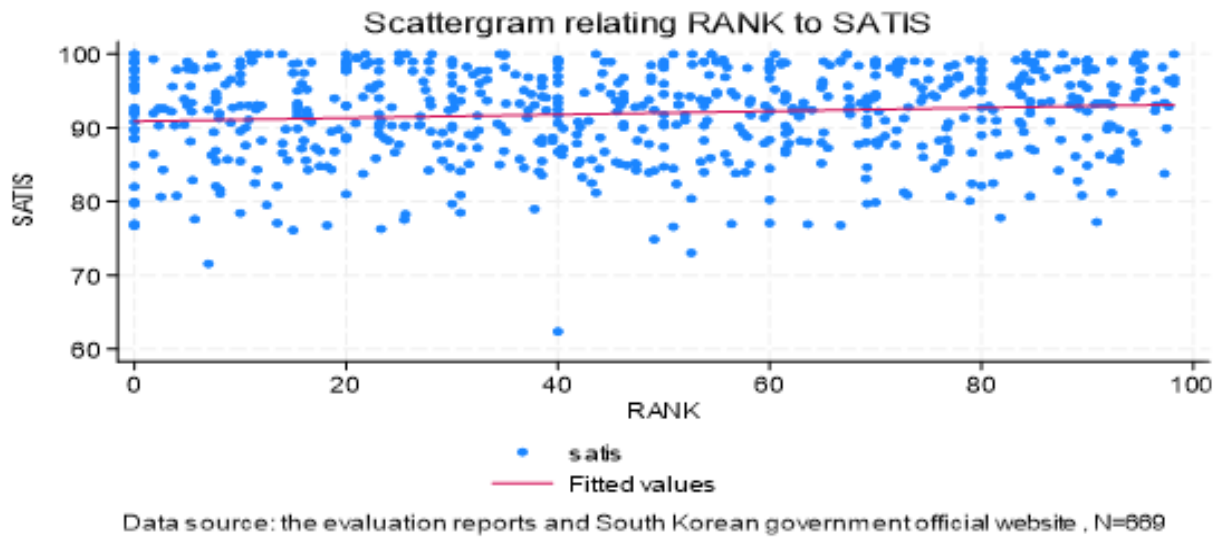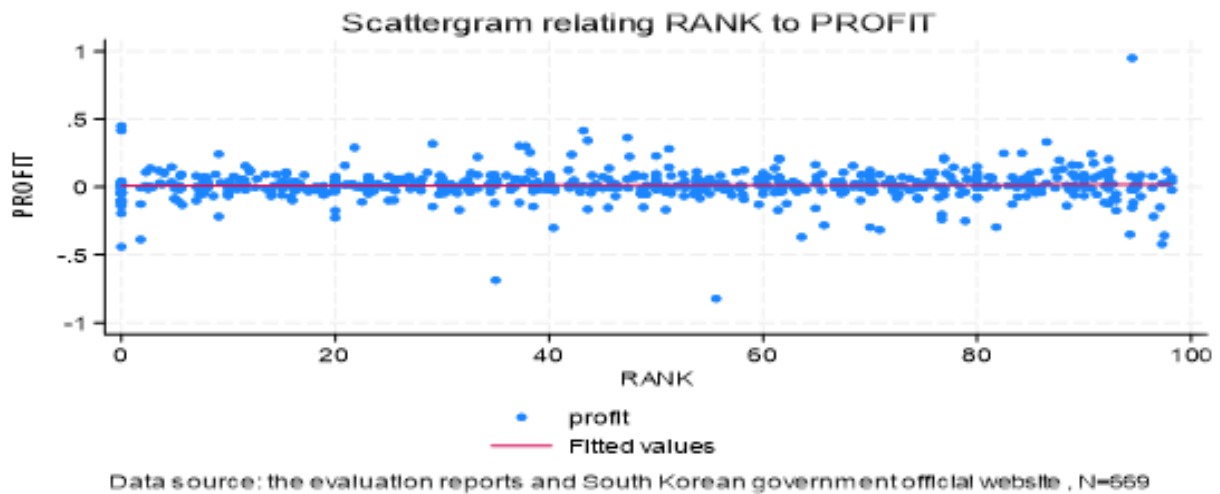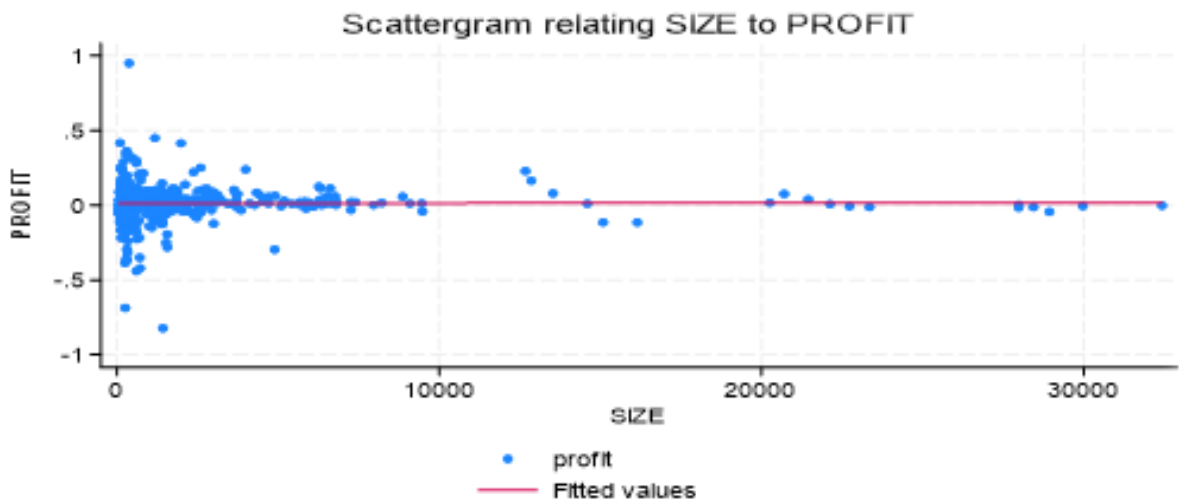
Scattergram relating RANK to PROFIT

profit
Fitted values

Data source: the evaluation reports and South Korean government official website , N=559

Scattergram relating RANK to SATIS

satis
Fitted values

Data source: the evaluation reports and South Korean government official website , N=669

Scattergram relating RANK to WELFARE

welfare
Fitted values

Data source: the evaluation reports and South Korean government official website , N=559

Scattergram relating SIZE to PROFIT

Data source: the evaluation reports and South Korean government official website , N=669

Scattergram relating SIZE to SATIS

Data source: the evaluation reports and South Korean government official website , N=669

Scattergram relating SIZE to WELFARE

Data source: the evaluation reports and South Korean government official website , N=669

152

Scattergram relating WAGE to PROFIT

Data source: the evaluation reports and South Korean government official website , N=669

Scattergram relating WAGE to SATIS

Data source: the evaluation reports and South Korean government official website , N=669

Scattergram relating WAGE to WELFARE
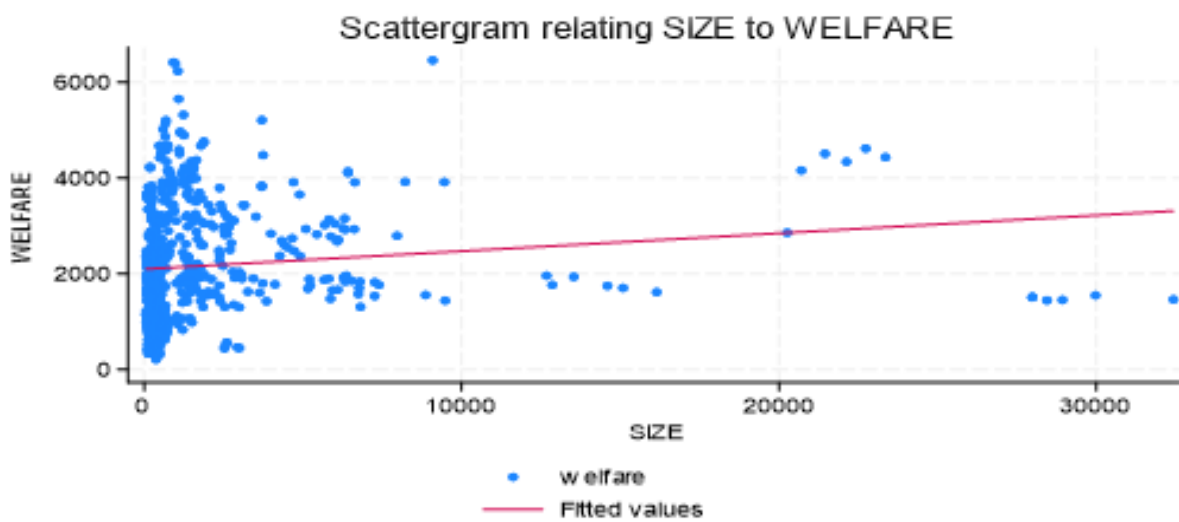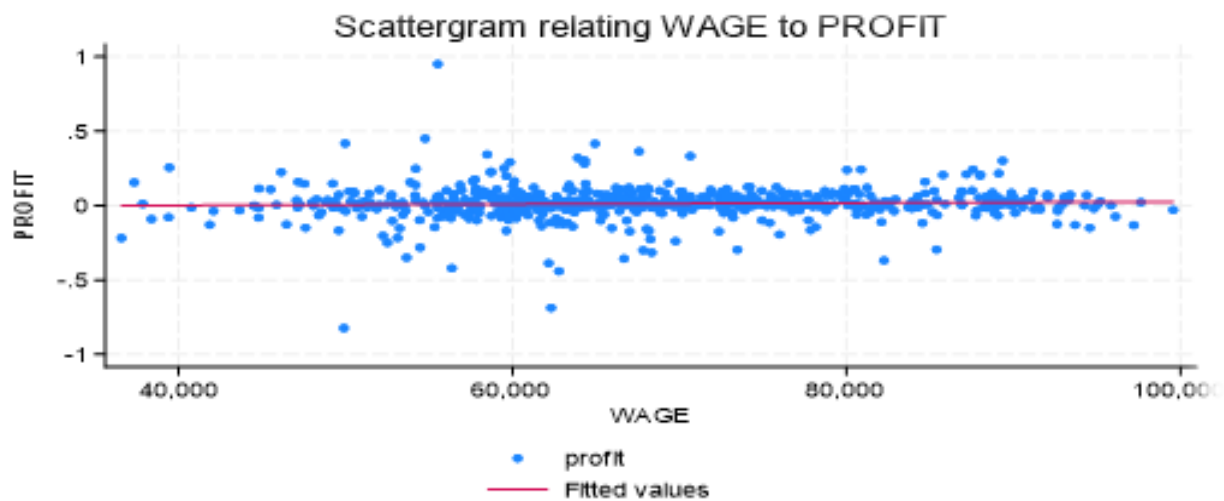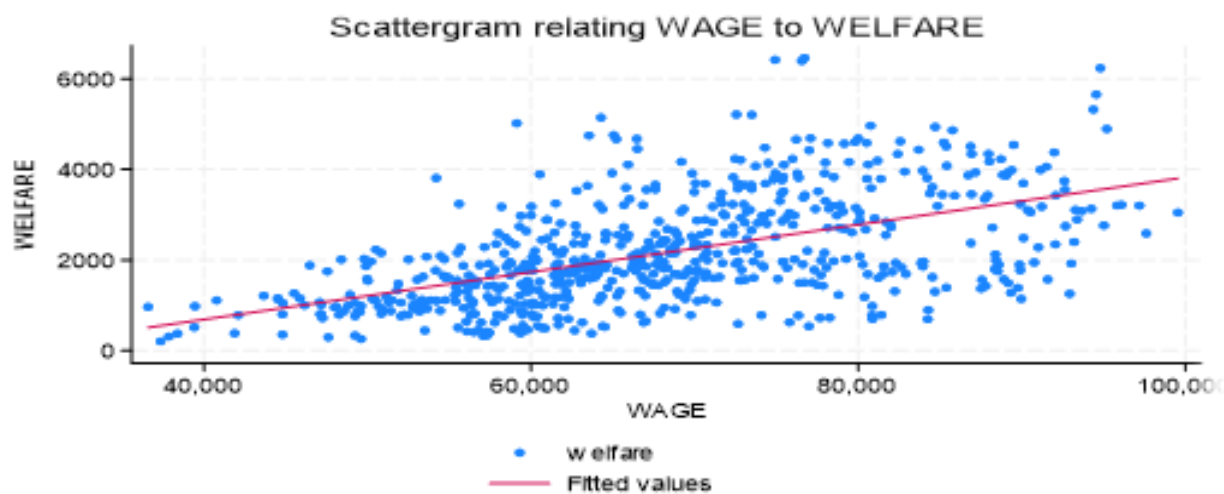
Data source: the evaluation reports and South Korean government official website , N=669

5.3.3 Diagnosis of multicollinearity

Multicollinearity can cause difficulty when interpreting regression coefficients because when multicollinearity exists, the independent variables are highly correlated and no single one has much explanatory power in the data (Agresti and Finlay, 2009). Therefore, it is important to examine whether multicollinearity exists in the multiple linear regression model (Agresti and Finlay, 2009; Acock, 2018). According to Wooldridge (2020), multicollinearity means high correlation between two or more independent variables. If there is high correlation between independent variables or independent and control variables, a regression coefficient in a multiple regression model has less meaning because a regression coefficient represents the effect of an independent or control variable when other variables are held constant (Agresti and Finlay, 2009). The Variance Inflation Factor (VIF) is generally used for diagnosis of multicollinearity (Agresti and Finlay, 2009; Acock, 2018; Wooldridge, 2020). If any independent or control variable's VIF is more than 10, a multicollinearity problem may exist (Agresti and Finlay, 2009; Acock, 2018). Acock (2018) suggests that it is helpful to drop the independent or control variable having a high VIF value.

Table 5-4 presents VIF values of the independent variable and all the control variables in the multiple linear regression model. VIF values of *RANK, TYPE, LISTED, SIZE, PAST, WAGE* are between 1 and 2, and hence much less than 10. Therefore, it can be argued that the multiple linear regression model in this study does not suffer from significant multicollinearity problems.

Table 5-4: VIF values of the independent variable and control variables in the multiple linear regression model

	RANK	TYPE	LISTED	SIZE	PAST	WAGE
VIF	1.08	1.35	1.25	1.14	1.03	1.16

5.4 Analysis of the relationship between the PIMES and the accountability of public institutions

The purpose of this section is to analyse how the PIMES has worked in relation to the managerial and social accountability of public institutions, addressing the first research question: *Do performance management systems improve the accountability of public organisations?* The approach used here for analysing the relationship between the PIMES and the accountability of public institutions is to examine how the PIMES evaluation result (*RANK*) affects *PROFIT* and *SATIS*. This is based on the following rationale. Public institutions are hypothesised to have incentives to try to obtain better PIMES *RANK* because better *RANK* gives the institutions more financial and non-financial benefits. Moreover, it is expected that the public institutions which invest more effort and resources to increase performance receive better *RANK* in the PIMES evaluation. Based on these expectations, it is argued that institutions with a higher *RANK* will have worked harder to achieve greater *PROFIT* and *SATIS*, and hence *RANK* can be expected to have positively affected managerial and social accountability.

The following framework is used to examine whether *RANK* has affected *PROFIT* and *SATIS* significantly. Firstly, regression models based on the model explained in Chapter 4 are estimated

using the data set for 2014-2019. Secondly, the two hypotheses explained in Chapter 4 that *RANK* has a positive effect on *PROFIT* and *SATIS* are examined by inspecting the regression coefficients on *RANK* for regressions with *PROFIT* and *SATIS* as dependent variables Theregression coefficients with associated test statistics and P-values are calculated by using the Stata program. The hypotheses are tested at a specific significance level by examining the P-values. In this study, the specific significance level is 0.05 considering the number of subjects of this regression is 669 as explained earlier. Many researchers use significance level of 0.05 when the number of sample is a few hundred (Wooldridge, 2020). In addition, the magnitude of each coefficient is interpreted in practical terms (Wooldridge, 2020). Subsequently, the test results are interpreted alongside the qualitative findings. Furthermore, this section analyses the regression results in relation to control variables that are related to organisation, personnel and finance characters of the public institutions.

5.4.1 Testing the effect of *RANK* on *PROFIT*

Regression models

This study uses a bivariate linear regression model employing only the independent variable *RANK* and a multiple linear regression model employing *RANK* and the control variables *TYPE*, *LISTED*, *SIZE*, *PAST*, *WAGE* to test the effect of PIMES evaluation results on the managerial accountability of public institutions. The bivariate and multiple regression models are as follows.

(i) Bivariate regression: $Y_{it} = \alpha + \beta_1 X_{it} + \varepsilon_{it}$

<Dependent variable>
· Y_{it}: Profit/Assets of institution i at time t

<Independent variable>
· X_{it}: PIMES standardised rank of institution i at time t

$$* X_{it} = \text{Standardised rank of i at t} =$$
$$\frac{number\ of\ public\ institutions\ in\ i's\ category\ at\ t - rank\ of\ i\ in\ i's\ category\ at\ t}{number\ of\ public\ institutions\ in\ i's\ category\ at\ t} \times 100$$
(i: a public institution, t: fiscal year)

· α and β are the constant and unstandardised coefficients respectively in the bivariate linear regression equation
· ε_{it}: denotes an error term

(ii) Multivariate regression: $Y_{it} = \alpha + \beta_1 X_{1it} + (\beta_2 D_{1it} + \beta_3 D_{2it} + \beta_4 X_{2it} + \beta_5 D_{3it} + \beta_6 X_{3it}) + \varepsilon_{it}$

<Dependent variable>
· Y_{it}: Profit/Assets of i at t

<Independent variable>
· X_{1it}: PIMES standardised rank of institution i at time t

$$* X_{1it} = \text{Standardised rank of i at t} =$$
$$\frac{number\ of\ public\ institutions\ in\ i's\ category\ at\ t - rank\ of\ i\ in\ i's\ category\ at\ t}{number\ of\ public\ institutions\ in\ i's\ category\ at\ t} \times 100$$
(i: a public institution, t: fiscal year)

<Control variables>

Organisation character-related factors
· D_{1it}: Type (public corporation/quasi-governmental institution: public corporation 1, quasi-governmental institution 0)
· D_{2it}: Stock (listed on the South Korean Stock Exchange or not: listed 1, not 0)

Personnel character-related factors
· X_{2it}: Number of regular employees of public institution of i at t
· D_{3it}: Past career of public institution's head (government official or not: government official 1, not 0)

Finance character-related factor
· X_{3it}: Average wage of regular employees of i at t

· α and β_j, j=1, 2, 3, 4, 5, 6 are the constant and unstandardised coefficients respectively in the multiple linear regression equation
· ε_{it}: denotes an error term

The hypothesis about the effect of RANK on PROFIT

The first hypothesis is about the impact of the PIMES evaluation on the managerial accountability of the public institutions. As explained earlier, the ratio of profit to assets (*PROFIT*) is used in order to operationalise the managerial accountability of public institutions. If *RANK* has not affected *PROFIT*, there would be no relationship between *RANK* and *PROFIT*. That implies that the coefficient (β_1) of *RANK* is not significantly different from zero. However, if the public institutions which receive better *RANK* have higher *PROFIT*, *RANK* would have a positive coefficient. and hence the coefficient (β_1) is positive. Consequently, the first null and alternative hypotheses can be stated as follows.

Null hypothesis H0: $\beta_1 = 0$

Alternative hypothesis H1: $\beta_1 > 0$

As the form of the alternative hypothesis is one-sided, the test of the first hypothesis is one-tailed test (Wooldridge, 2020).

Significance test result for the hypothesis on PROFIT

The regression coefficients, test statistics (t-values) and P-values of the β_1 in the bivariate and multiple regression models are calculated in order to test above hypotheses. Table 5-5 and Table 5-6 present the bivariate and multiple linear regression results with the 669 subjects through the Stata program. In the bivariate regression result, the regression coefficient (β_1) is 0.00010 and P-value of it is 0.247 under one-tailed test where the variable is assumed to be distributed as 't', which is not statistically significant at the 0.05 level. In the multiple regression result, the regression coefficient (β_1) is 0.00008 and P-value of it is 0.311 under one-tailed test where the variable is assumed to be

distributed as 't', which is not statistically significant at the 0.05 level. Therefore, this study is unable to reject the null hypothesis ($\beta_1 = 0$) for *PROFIT* at the 0.05 level. Wooldridge (2020) suggests that although a variable is not statistically significant at the usual levels such as 0.1, 0.5 or 0.01, it might still be necessary to ask whether the variable has the expected effect on the dependent variable and whether the effect is practically large.

This study interprets the magnitude of each coefficient in practical terms as follows. Firstly, the distribution of values of each variable is calculated under the assumption that the variable has a normal distribution. Under the assumption of normal distribution, about 95% of values of the variable are within two standard deviations of mean of the variable (Acock, 2018). Consequently, maximum and minimum values within 95% of values of the variable are calculated by adding two standard deviations and the mean and subtracting two standard deviations from the mean respectively. Secondly, the coefficient of variable is multiplied by the maximum and minimum values, and the *scope* of the coefficient depending on 95% of values of the variable is defined as 'the coefficient of variable x the maximum – the coefficient of variable x the minimum' in this study. The *scope* of the coefficient refers to change in *PROFIT*, *SATIS*, *WELFARE* depending on a 95% change in the variable of *RANK* or control variables, holding other variables constant. Lastly, this study interprets the the *scopes* of the coefficients of *PROFIT*, *SATIS*, *WELFARE* according to 95% change in the variable.

The regression coefficients of *RANK* are positive (0.00010 and 0.00008) in the bivariate and multiple regressions respectively as the alternative hypothesis expects. In other words, the public institutions with better *RANK* are predicted to have higher *PROFIT*. Table 5-2 presents that the mean and standard deviation of *RANK* are 47.83528 and 28.81297 respectively. As a result, about 95% of the values of

RANK are between -9.79066 (47.83528 - 2 × 28.81297) and 105.46122 (47.83528 + 2 × 28.81297).

As *RANK*'s range is from 0 to an upper limit of 100, minimum value (0) and maximum value (98.2) of *RANK* are used as minimum and maximum values within 95% of values of *RANK*. Consequently, in the bivariate and multiple regressions, the *scopes* of coefficients of *PROFIT* depending on 95% of values of *RANK* are $\quad$ 0.00982 (0.00010 × 98.2 − 0.00010 × 0) and 0.00786 (0.00008 × 98.2 − 0.00008 × 0) respectively, holding other variables constant. Therefore, the scopes of coefficients of *PROFIT* are less than 0.01 when expressed as the ratio of profit to assets. In addition, Table 5-2 shows that the standard deviation of *PROFIT* is 0.11019. As the scopes of coefficients of *PROFIT* (0.00982 and 0.00786) are much smaller than the standard deviation of *PROFIT* (0.11019), the effect of *RANK* on *PROFIT* is not practically large. Therefore, it is argued that *RANK* has not affected *PROFIT* significantly in either statistical or practical terms. Thus, this research is unable to find support for the hypothesis that the PIMES through the variable *RANK* has a significant positive effect on managerial accountability as measured by *PROFIT*.

Table 5-5: The effect of the independent variable *RANK* on the dependent variable *PROFIT*

Category	Regression coefficient	t-value	P-value	Test
Independent variable *RANK*	0.00010	0.69	0.247	One-tailed
Constant	0.00811	0.98	0.164	
F			0.47	
R-squared			0.0007	

Note: * denotes statistically significant at the 0.10 level, ** at the 0.05 level, *** at the 0.01 level

Table 5-6: The effect of the independent and control variables on the dependent variable *PROFIT*

Category		Name of variable	Regression coefficient	t-value	P-value	Test
Independent variable		X_1: *RANK*	0.00008	0.49	0.311	One-tailed
Control variables	Organisation character	D_1: *TYPE* (public corporation 1, quasi-governmental institution 0)	-0.01017	-0.87	0.191	One-tailed
		D_2: *LISTED* (listed on the South Korean Stock Exchange 1, not 0)	0.02923	1.25	0.107	One-tailed
	Personnel character	X_2: *SIZE* (number of regular employees)	-1.11e-07	-0.09	0.463	One-tailed
		D_3: *PAST* (past career of head is a government official 1, not 0)	0.01068	1.23	0.110	One-tailed
	Finance character	X_3: *WAGE* (average wage of regular employees)	4.09e-07	1.10	0.136	One-tailed
Constant			-0.02206	-0.87	0.192	
F			0.86			
R-squared			0.0077			

Note: * denotes statistically significant at the 0.10 level, ** at the 0.05 level, *** at the 0.01 level

5.4.2 Testing the effect of *RANK* on *SATIS*

Regression models

The bivariate and multiple regression models to examine the effect of *RANK* on *SATIS* are as follows.

(i) Bivariate regression: $Y_{it} = \alpha + \beta_1 X_{it} + \varepsilon_{it}$

<Dependent variable>
· Y_{it}: Customer satisfaction score of institution i at time t

<Independent variable>
· X_{it}: PIMES standardised rank of institution i at time t

$$X_{it} = \text{Standardised rank of i at t} =$$
$$\frac{number\ of\ public\ institutions\ in\ i's\ category\ at\ t - rank\ of\ i\ in\ i's\ category\ at\ t}{number\ of\ public\ institutions\ in\ i's\ category\ at\ t} \times 100$$
(i: a public institution, t: fiscal year)

· α and β are the constant and unstandardised coefficients respectively in the bivariate linear regression equation
· ε_{it}: denotes an error term

(ii) Multivariate regression: $Y_{it} = \alpha + \beta_1 X_{1it} + (\beta_2 D_{1it} + \beta_3 D_{2it} + \beta_4 X_{2it} + \beta_5 D_{3it} + \beta_6 X_{3it}) + \varepsilon_{it}$

<Dependent variable>
· Y_{it}: Customer satisfaction score of institution i at time t

<Independent variable>
· X_{1it}: PIMES standardised rank of institution i at time t

$$* X_{1it} = \text{Standardised rank of i at t} =$$
$$\frac{number\ of\ public\ institutions\ in\ i's\ category\ at\ t - rank\ of\ i\ in\ i's\ category\ at\ t}{number\ of\ public\ institutions\ in\ i's\ category\ at\ t} \times 100$$
(i: a public institution, t: fiscal year)

<Control variables>

Organisation character-related factors
· D_{1it}: Type (public corporation/quasi-governmental institution: public corporation 1, quasi-governmental institution 0)
· D_{2it}: Stock (listed on the South Korean Stock Exchange or not: listed 1, not 0)

Personnel character-related factors
· X_{2it}: Number of regular employees of public institution of i at t
· D_{3it}: Past career of public institution's head (government official or not: government official 1, not 0)

Finance character-related factor
· X_{3it}: Average wage of regular employees of i at t

· α and β_j, j=1, 2, 3, 4, 5, 6 are the constant and unstandardised coefficients respectively in the multiple linear regression equation
· ε_{it}: denotes an error term

The hypothesis about the effect of RANK on SATIS

The second hypothesis is about the impact of the PIMES evaluation on the social accountability of public institutions. As explained earlier, the customer satisfaction score (*SATIS*) is used in this study to operationalise the social accountability of public institutions. If *RANK* has not affected *SATIS*, there would be no relationship between *RANK* and *SATIS*. That implies that the coefficient (β_1) of *RANK* with respect to *SATIS* is not significantly different from zero. However, if the public institutions which receive better *RANK* have higher *SATIS*, *RANK* would have a positive coefficient and hence the coefficient (β_1) of *RANK* with respect to *SATIS* is positive. Therefore, the second null and alternative hypotheses can be stated as follows.

Null hypothesis H0: $\beta_1 = 0$

Alternative hypothesis H1: $\beta_1 > 0$

As the form of the alternative hypothesis is one-sided, the test of the second hypothesis is one-tailed test (Wooldridge, 2020).

Significance test result for the hypothesis on SATIS

Table 5-7 and Table 5-8 present the bivariate and multiple linear regression results through the Stata program to test the above hypotheses. While the regressions for estimated *PROFIT* were estimated with the 669 subjects, the regressions for *SATIS* are estimated with the 668 subjects because the outlier (*RANK* is '0' and *SATIS* is '0') which was identified earlier was removed. Table 5-9 presents the descriptive statistics for *SATIS*, *RANK* and control variables with the 668 subjects. In the bivariate regression for *SATIS*, the regression coefficient (β_1) of *RANK* is positive (0.02299) and P-value of it

is 0.003 under one-tailed test where the variable is assumed to be distributed as 't', which is statistically significant at the 0.05 level. Moreover, in the multiple regression for *SATIS*, the coefficient (β_1) of *RANK* is positive (0.02433) and P-value of it is 0.002 under one-tailed test where the variable is assumed to be distributed as 't', which is statistically significant at the 0.05 level. Therefore, this study rejects the null hypothesis ($\beta_1 = 0$) for *SATIS* in favour of the alternative hypothesis ($\beta_1 > 0$) for *SATIS* at the 0.05 level (Wooldridge, 2020).

This study interprets the coefficients (0.02299 and 0.02433) of *SATIS* in practical terms in accordance with the same procedure as above. As explained earlier, the minimum value (0) and maximum value (98.2) of *RANK* are used as minimum and maximum values within 95% of values of *RANK*. Consequently, the *scopes* of coefficients of *SATIS* depending on 95% of values of *RANK* are 2.25762 (0.02299 x 98.2 − 0.02299 x 0) and 2.38921 (0.02433 x 98.2 − 0.02433 x 0) respectively, holding other variables constant. Table 5-9 shows that the mean of *SATIS* is 91.98235 and the maximum value of *SATIS* is 100. Considering the mean and maximum value of *SATIS*, the scopes of coefficients of *SATIS* (2.25762 and 2.38921) are practically large. Therefore, it is argued that *RANK* has had a significant positive effect on *SATIS* in both statistical and practical terms. The implication of this finding will be discussed in detail in Chapter 7.

Table 5-7: The effect of the independent variable *RANK* on the dependent variable *SATIS*

Category	Regression coefficient	t-value	P-value	Test
Independent variable *RANK*	0.02299[***]	2.85	0.003	One-tailed
Constant	90.88101[***]	201.37	0.000	
F	8.10[***]			
R-squared	0.0120			

Note: [*] denotes statistically significant at the 0.10 level, [**] at the 0.05 level, [***] at the 0.01 level

Table 5-8: The effect of the independent and control variables on the dependent variable *SATIS*

Category		Name of variable	Regression coefficient	t-value	P-value	Test
Independent variable		X_1: *RANK*	0.02433[***]	2.93	0.002	One-tailed
Control variables	Organisation character	D_1: *TYPE* (public corporation 1, quasi-governmental institution 0)	2.61523[***]	4.18	0.000	One-tailed
		D_2: *LISTED* (listed on the South Korean Stock Exchange 1, not 0)	-1.82726	-1.44	0.076	One-tailed
	Personnel character	X_2: *SIZE* (number of regular employees)	0.00006	0.91	0.181	One-tailed
		D_3: *PAST* (past career of head is a government official 1, not 0)	-0.06636	-0.14	0.444	One-tailed
	Finance character	X_3: *WAGE* (average wage of regular employees)	-0.00002	-0.75	0.226	One-tailed
Constant			91.22467[***]	67.14	0.000	

F	4.89***
R-squared	0.0425

Note: * denotes statistically significant at the 0.10 level, ** at the 0.05 level, *** at the 0.01 level

Table 5-9: Descriptive statistics for *SATIS*, *RANK*, *TYPE*, *LISTED*, *SIZE*, *PAST*, *WAGE* (N=668)

Variable	Mean	Standard deviation	Minimum	Maximum	VIF
SATIS	91.98235	6.03475	62.4	100	-
RANK	47.90689	28.77492	0	98.2	1.08
TYPE	0.23802	0.42619	0	1	1.35
LISTED	0.04341	0.20394	0	1	1.27
SIZE	1704.56	3671.48	61	29973	1.15
PAST	0.45359	0.49821	0	1	1.03
WAGE	68172.13	12396.44	36577	99557	1.16

5.4.3 Analysis of quantitative results in relation to the qualitative research

As explained in Chapter 4, this study follows the convergent mixed methods research design which collects both quantitative and qualitative data concurrently, analyses both sets separately and compares different findings produced through quantitative and qualitative methods (Creswell, 2015). The convergent mixed methods research design can present complementary evidence about research problem (Plano Clark and Ivankova, 2016). Following the convergent mixed methods research design, the interviews with key participants would need to examine both why, in the above statistical analysis, *RANK* has not affected *PROFIT* significantly, but why, in contrast, it has affected *SATIS* significantly.

This study therefore now examines these questions through the qualitative data. The insignificant effect of *RANK* on *PROFIT* means that the PIMES evaluation process does not appear to improve the managerial accountability of public institutions measured by *PROFIT* in this study. The interviews with employees, evaluators, and government officials provide empirical evidence that supports the insignificant effect of *RANK* on *PROFIT*. Firstly, this may be because the weight of performance indicators in the PIMES in relation to financial performance has decreased noticeably.

> *"In the early stages of the PIMES, only government-invested institutions were evaluated and financial performance was prioritised in the evaluation. However, the weight of performance indicators that evaluate financial performance and budget is only five percent currently and hence the function for government as a shareholder of the public institutions to evaluate the financial performance of public institutions has weakened." [G4]*

Secondly, it was found that both the employees and evaluators show indifference to financial performance especially for the quasi-governmental institutions. For example, two evaluators highlighted that financial performance does not make difference when evaluating quasi-governmental institutions.

> *"The majority of quasi-governmental institutions are providing public services on behalf of government and are operated entirely by government budget. Therefore, the financial performance of quasi-governmental institutions is not evaluated significantly." [V8]*

> *"There is only one performance indicator in the PIMES in relation to financial performance for quasi-governmental institutions. Moreover, all the quasi-governmental institutions receive perfect scores in the performance indicator. As a result, there is no difference caused by financial performance of quasi-governmental institutions in the PIMES evaluation results." [V9]*

Lastly, the South Korean government's priorities in measuring the performance of public institutions

changed and the change of priority was reflected in the performance indicators in the PIMES. For example, one interviewee said that the weight of performance indicators in relation to social value such as job creation, safety and environment increased, however, the weight of performance indicators regarding financial performance decreased.

> *"In the present government, the evaluators evaluate the public institutions prioritising safety and environment. Therefore, I think that the proportion of financial performance in evaluation decreased in the present government." [E22]*

Another interviewee suggested why public institutions do not focus on improving financial performance.

> *"Evaluators evaluate the ratio of profit to assets and debt of my institution in the PIMES evaluation process. However, weight of performance indicator regarding financial performance is only 1%. As a result, most public institutions do not focus on increasing the ratio of profit to assets." [E18]*

The significance test result for the effect of *RANK* on *SATIS* is that *RANK* has significantly affected *SATIS*. The significant effect of *RANK* on *SATIS* supports a view that the PIMES evaluation process has induced the public institutions to improve social accountability as measured by *SATIS* in this study. The interviews with employees and government officials provide several reasons why the PIMES evluation process induces public institutions to improve their social accountability as measured by *SATIS*. The first reason is that public institutions' missions which are carefully evaluated in the PIMES evaluation are closely linked to the questions of the customer satisfaction survey. As a result, better *RANK* positively influences *SATIS*.

> *"I know that customer satisfaction survey questions consist of customers' evaluation on supporting programs of my institution. Hence, I think that customer satisfaction score is determined by the efforts to provide the public service*

168

effectively." [E9]

The second reason is that most public institutions try to manage their divisions to achieve higher customer satisfaction scores. For example, an employee said that each division within his institution is evaluated according to its customer satisfaction results and sanctioned if its results are poor.

> *"My institution makes a customer satisfaction plan in order to increase customer satisfaction level and requires underperforming divisions in customer satisfaction results to make customer satisfaction improvement plans." [E22]*

The evidence of PIMES causing a focus on *SATIS* is that public institutions use the customer satisfaction results for internal evaluation of their employees. The employees of public institutions have no choice but to try to respond to their customers' requirements under the structure of evaluation.

> *"I think that the employees of public institutions do their best to respond to the customer satisfaction survey. It's because the public institutions aim to advertise their good customer satisfaction level and use the result when they evaluate their employees internally." [G2]*

5.4.4 Testing the effect of the character of public institutions on their accountability

In the regression models, organisation, personnel and finance character-related factors of the public institutions are used as control variables to clarify the effect of *RANK* on *PROFIT* and *SATIS*, holding these control factors constant. The control factors are determined by external actors such as government, and hence the public institutions cannot themselves control these external factors. If the control variables affect *PROFIT* and *SATIS* significantly, this could have an apparent effect on accountability outside of the PIMES process. Therefore, this study examines whether the control

variables affect *PROFIT* and *SATIS* significantly using the multiple regression models set out earlier.

Table 5-6 presents the multiple regression results that include the effect of control variables on *PROFIT*. The regression coefficients (β_3 and β_5) of *LISTED* and *PAST* are positive (0.02923 and 0.01068) as they are expected to have a positive effect on *PROFIT* (see Table 4-1). However, the regression coefficients (β_2 and β_4) of *TYPE* and *SIZE* are negative (-0.01017 and -1.11e-07), whereas the expected signs of their effect are positive. *WAGE* is expected to have a negative effect on *PROFIT*, however, the regression coefficient (β_6) of *WAGE* is positive (4.09e-07). None of the P-values (0.191, 0.107, 0.463, 0.110 and 0.136) of regression coefficients of *TYPE*, *LISTED*, *SIZE*, *PAST* and *WAGE* under one-tailed test are statistically significant at the 0.05 level. The effects of being listed on the South Korean Stock Exchange (*LISTED*) and a head's past career as a government official (*PAST*) on *PROFIT* are 0.02923 and 0.01068. They are less than 0.03 when expressed as the ratio of profit to assets and much smaller than the standard deviation of *PROFIT* (0.11019). Consequently, the effects of *LISTED* and *PAST* on *PROFIT* are not practically large. Therefore, it is argued that the control variables have not affected *PROFIT* significantly in statistical and practical terms.

Table 5-8 presents the multiple regression results that include the effect of control variables on *SATIS*. While all the control variables are expected to have a positive effect on *SATIS* (see Table 4-1), only the regression coefficients (β_2 and β_4) of *TYPE* and *SIZE* are positive (2.61523 and 0.00006). The P-value of regression coefficient of *TYPE* is 0.000 under one-tailed test and is statistically significant at the 0.05 level. However, the P-values (0.076, 0.181, 0.444, 0.226) of regression coefficients of *LISTED*, *SIZE*, *PAST*, *WAGE* under one-tailed test are not statistically significant at the 0.05 level. The effect of being a public corporation (*TYPE*) on *SATIS* is 2.61523 and is practically large

considering that the mean of *SATIS* is 91.98235 and the maximum value of *SATIS* is 100. Therefore, it is argued that *TYPE* has affected *SATIS* significantly in statistical and practical terms. However, this research can note that the public corporations and quasi-governmental institutions are evaluated and ranked in different evaluation groups as explained in Chapter 3 and hence, *RANK* would not be affected by *TYPE*. The practical significance of regression coefficient of *SIZE* can be examined in accordance with the procedure explained earlier. Table 5-9 indicates that the mean and standard deviation of *SIZE* are 1704.56 and 3671.48 respectively. As a result, about 95% of the values of *SIZE* are between -5638.4 (1704.56 - 2 × 3671.48) and 9047.52 (1704.56 + 2 × 3671.48). As *SIZE* has a lower bound of zero, zero and 9047.52 are used as minimum and maximum values within 95% of values of *SIZE*. Consequently, the *scope* of coefficient of *SATIS* depending on 95% of values of *SIZE* is 0.54285 (0.00006 × 9047.52 − 0.00006 × 0), holding other variables constant. The *scope* of coefficient of *SATIS* (0.54285) is under one out of perfect score of 100 in customer satisfaction score, it is argued that *SIZE* has not affected *SATIS* significantly in statistical and practical terms.

The regression results about the effects of organisation, personnel and finance character-related factors on *PROFIT* and *SATIS* suggest that only the control factor *TYPE* has affected accountability. The regression results indicate that public corporations are expected to have better customer satisfaction score than quasi-governmental institutions. It is useful to bear that in mind when considering accountability in customer-satisfaction lens. Although *TYPE* has affected *SATIS* significantly, *TYPE* does not affect *RANK* between organisation types because public corporations and quasi-governmental institutions are evaluated and ranked in different evaluation groups in the PIMES.

171

5.5 Analysis of the relationship between the PIMES and the strategic behaviour of public institutions

The purpose of this section is to analyse working of the PIMES in relation to strategic behaviour resulting from moral hazard. This study therefore addresses the third research question: *what is the impact of performance management systems on the strategic behaviour of public organisations?* The analysis follows the significance test process employed in the previous section. Firstly, bivariate and multiple regression models are set out based on the regression model explained in Chapter 4. The hypothesis set out in Chapter 4 that South Korean public institutions having better PIMES results are more likely to have lower average welfare benefit of regular employees -is stated in terms of the regression coefficient of *RANK* with respect to *WELFARE*.   The regression coefficients, test statistics and P-values in the regression models are calculated by using the Stata program. Subsequently, the hypothesised negative impact of *RANK* on *WELFARE* is tested at the 0.05 level and the test result is interpreted alongside the qualitative findings. Lastly, the effects of the control variables on *WELFARE* are examined.

5.5.1 Regression models

As the bivariate and multiple regression models were used to examine the effects of *RANK* and control variables on *PROFIT* and *SATIS*, bivariate and multiple regression models are set out to examine the effect of *RANK* and control variables on *WELFARE*. The bivariate and multiple regression models are as follows.

(i) Bivariate regression: $Y_{it} = \alpha + \beta_1 X_{it} + \varepsilon_{it}$

<Dependent variable>
· Y_{it}: Average welfare benefit of regular employees of institution i at time t

<Independent variable>
· X_{it}: PIMES standardised rank of institution i at time t

$$* X_{it} = \text{Standardised rank of } i \text{ at } t =$$
$$\frac{number\ of\ public\ institutions\ in\ i's\ category\ at\ t - rank\ of\ i\ in\ i's\ category\ at\ t}{number\ of\ public\ institutions\ in\ i's\ category\ at\ t} \times 100$$
(i: a public institution, t: fiscal year)

· α and β are the constant and unstandardised coefficients respectively in the bivariate linear regression equation
· ε_{it}: denotes an error term

(ii) Multivariate regression: $Y_{it} = \alpha + \beta_1 X_{1it} + (\beta_2 D_{1it} + \beta_3 D_{2it} + \beta_4 X_{2it} + \beta_5 D_{3it} + \beta_6 X_{3it}) + \varepsilon_{it}$

<Dependent variable>
· Y_{it}: Average welfare benefit of regular employees of institution i at time t

<Independent variable>
· X_{1it}: PIMES standardised rank of institution i at time t

$$* X_{1it} = \text{Standardised rank of } i \text{ at } t =$$
$$\frac{number\ of\ public\ institutions\ in\ i's\ category\ at\ t - rank\ of\ i\ in\ i's\ category\ at\ t}{number\ of\ public\ institutions\ in\ i's\ category\ at\ t} \times 100$$

(i: a public institution, t: fiscal year)

<Control variables>

Organisation character-related factors

173

· D_{1it}: Type (public corporation/quasi-governmental institution: public corporation 1, quasi-governmental institution 0)
· D_{2it}: Stock (listed on the South Korean Stock Exchange or not: listed 1, not 0)

Personnel character-related factors
· X_{2it}: Number of regular employees of public institution of i at t
· D_{3it}: Past career of public institution's head (government official or not: government official 1, not 0)

Finance character-related factor
· X_{3it}: Average wage of regular employees of i at t

· α and β_j, j=1, 2, 3, 4, 5, 6 are the constant and unstandardised coefficients respectively in the multiple linear regression equation
· ε_{it}: denotes an error term

5.5.2 The hypothesis about the effect of *RANK* on *WELFARE* and significance test result

The hypothesis examined in this section is that high *WELFARE* is the result of managers failing to prevent strategic behaviour of employees leading to moral hazard that can result in a bad performance. Better *RANK* is expected to be reflected in a lower *WELFARE*. This hypothesis was set out in more detail in Chapter 4. This hypothesis of a negative effect of *RANK* on *WELFARE* can be stated as follows.

Null hypothesis H0: $\beta_1 = 0$

Alternative hypothesis H1: $\beta_1 < 0$

As the form of the alternative hypothesis is one-sided, the test of the first hypothesis is one-tailed test (Wooldridge, 2020).

Table 5-10 and Table 5-11 present the bivariate and multiple linear regression results with the 669 subjects through the Stata program. In the bivariate regression result, the regression coefficient (β_1) of *RANK* is 0.73679, whereas *RANK* is expected to have a negative effect on *WELFARE*. This bivariate regression result corresponds to the preliminary statistical analysis result that the Pearson bivariate correlation between *RANK* and *WELFARE* is positive (0.0180) (see Table 5-3). The P-value of the regression coefficient (β_1) is 0.321 under one-tailed test where the variable is assumed to be distributed as 't', which is not statistically significant at the 0.05 level. In the multiple regression result, the regression coefficient (β_1) of *RANK* is negative (-0.14814) as *RANK* is expected to have a negative effect on *WELFARE*. The P-value of it is 0.453 under one-tailed test where the variable is assumed to be distributed as 't', which is not statistically significant at the 0.05 level. Therefore, this study is unable to reject the null hypothesis ($\beta_1 = 0$) for the effect of *RANK* on *WELFARE* at the 0.05 level.

The practical significance of regression coefficient (β_1) of *RANK* is examined in accordance with the procedure in previous section. As explained earlier, the minimum value (0) and maximum value (98.2) of *RANK* are used as minimum and maximum values within 95% of values of *RANK*. Consequently, the *scope* of coefficient of *WELFARE* depending on 95% of values of *RANK* is 14.54735 (-0.14814 × 98.2 – (-0.14814) × 0), holding other variables constant. Table 5-2presents that the mean and standard deviation of *WELFARE* are 2161.862 (1350 pound) and 1176.498 (710 pound) respectively. As the *scope* of coefficient of *WELFARE* (14.54735) is much smaller than the mean (2161.862) and standard deviation (1176.498), the effect of *RANK* on *WELFARE* is not practically large. Therefore, it is argued that *RANK* has not affected *WELFARE* significantly in statistical and practical terms.

Table 5-10: The effect of the independent variable *RANK* on the dependent variable *WELFARE*

Category	Regression coefficient	t-value	P-value	Test
Independent variable *RANK*	0.73679	0.47	0.321	One-tailed
Constant	2126.618***	24.10	0.000	
F			0.22	
R-squared			0.0003	

Note: * denotes statistically significant at the 0.10 level, ** at the 0.05 level, *** at the 0.01 level

Table 5-11: The effect of the independent and control variables on the dependent variable *WELFARE*

Category		Name of variable	Regression coefficient	t-value	P-value	Test
Independent variable		X_1: *RANK*	-0.14814	-0.12	0.453	One-tailed
Control variables	Organisation character	D_1: *TYPE* (public corporation 1, quasi-governmental institution 0)	969.7018***	10.22	0.000	One-tailed
		D_2: *LISTED* (listed on the South Korean Stock Exchange 1, not 0)	111.0856	0.58	0.281	One-tailed
	Personnel character	X_2: *SIZE* (number of regular employees)	-0.00555	-0.57	0.283	One-tailed
		D_3: *PAST* (past career of head is a government official 1, not 0)	233.4006***	3.29	0.001	One-tailed
	Finance character	X_3: *WAGE* (average wage of regular employees)	0.04003***	13.20	0.000	One-tailed
Constant			-893.1413***	-4.33	0.000	

F	80.09***
R-squared	0.4206

Note: * denotes statistically significant at the 0.10 level, ** at the 0.05 level, *** at the 0.01 level

5.5.3 Analysis of the significance test results

This study now examines, through the qualitative data, reasons why *RANK* has not affected *WELFARE* significantly as demonstrated in the quantitative analysis. For example, some interviewees provided relevant comment on this insignificant effect of *RANK* on *WELFARE*.

> *"The former governments used the performance indicators that control expenses such as business performance efficiency, administrative expense and average welfare benefit when they evaluate quasi-governmental institutions. However, the performance indicators disappeared in evaluation of quasi-governmental institutions in the present government." [V9]*

The interview from above evaluator indicates that the weight of performance indicators controlling excessive welfare benefits has decreased like performance indicators regarding financial performance. An employee highlighted that priority in controlling welfare benefits changed in the PIMES evaluation.

> *"Regarding the welfare benefit, the former governments evaluated the size of it significantly. On the other hand, the difference in welfare benefit between regular and non-regular employees, not the size of welfare benefit, is an important assessment factor in the present government." [E15]*

Considering the interview data, it is argued that the focus of evaluation regarding performance indicators controlling excessive pay and welfare benefits was changed. When the MOEF announced

the measure to normalise the management of public institutions in 2013, the focus was to reduce the amount of excessive welfare benefit. However, the previous focus on reducing the amount of welfare benefit was changed into emphasis on decreasing the difference of welfare benefit between regular and non-regular employees within a public institution. Consequently, realising the lack of importance of *WELFARE* in the more-recent PIMES evaluation process, the impact of *WELFARE* has become unsignificant, therefore, the PIMES does not now give an incentive to keep *WELFARE* low.

5.5.4 Testing of the effect of the character of public institutions on their strategic behaviour

In the above multiple regression model, organisation, personnel and finance character-related factors of the public institutions are used as control variables to clarify the effect of *RANK* on *WELFARE*, holding theses control factors constant. This study examines whether the control variables affect *WELFARE* significantly with the above multiple regression model.

Table 5-11 presents the multiple regression results about the effect of control variables on *WELFARE*. The regression coefficients (β_2, β_3, β_5) of *TYPE*, *LISTED*, *PAST* are positive (969.7018, 111.0856, 233.4006), whereas the expected signs of their effects are negative (see Table 4-1). While the P-values (0.000 and 0.001) of regression coefficients of *TYPE* and *PAST* under one-tailed test are statistically significant at the 0.05 level, the P-value (0.281) of regression coefficient of *LISTED* under one-tailed test is not statistically significant at the 0.05 level. Table 5-2shows that the mean of *WELFARE* is 2161.862 (1350 pound). The effects of being a public corporation (*TYPE*) and a head's past career as a government official (*PAST*) on *WELFARE* are 969.7018 (606 pound) and 233.4006 (146 pound), and both are practically large considering the mean of *WELFARE*. Therefore, it is argued

that *TYPE* and *PAST* have affected *WELFARE* positively with statistical and practical significance. Hence, it can be argued that an organisation being a public corporation and having a past government official in charge attenuate the already insignificant impact of the PIMES evaluation process on their strategic behaviour as measured by *WELFARE*.

The regression coefficient (β_4) of *SIZE* is negative (-0.00555), whereas the expected sign of its effect is positive (see Table 4-1). The P-value (0.283) of regression coefficient of *SIZE* under one-tailed test is not statistically significant at the 0.05 level. As explained earlier, minimum and maximum values within 95% of values of *SIZE* are zero and 9047.52. Consequently, the *scope* of coefficient of *WELFARE* depending on 95% of values of *SIZE* is 50.21374 (-0.00555 x 9047.52 – (-0.00555) x 0), holding other variables constant. Considering the mean of *WELFARE* is 2161.862, the *scope* of coefficient of *WELFARE* (50.21374) is not practically large. Therefore, it is argued that *SIZE* has not affected *WELFARE* significantly.

The regression coefficient (β_6) of *WAGE* is positive (0.04003), as the expected sign of its effect is positive (see Table 4-1). The P-value (0.000) of regression coefficient of *WAGE* under one-tailed test is statistically significant at the 0.05 level. Table 5-2presents that the mean and standard deviation of *WAGE* are 68174.54 (42610 pound) and 12387.31 (7740 pound) respectively. As a result, about 95% of the values of *WAGE* are between 43399.92 (68174.54 - 2 x 12387.31) and 92949.16 (68174.54 + 2 x 12387.31). Consequently, the *scope* of coefficient of *WELFARE* depending on 95% of values of *WAGE* is 1983.46 (0.04003 x 92949.16 – 0.04003 x 43399.92), holding other variables constant.

Considering the mean of *WELFARE* is 2161.862 (1350 pound), the *scope* of coefficient of *WELFARE* (1983.46) is practically large. Therefore, it is argued that *WAGE* has affected *WELFARE* positively with statistical and practical significance. Hence, it can be argued that the insignificant impact of *RANK* on *WELFARE* is, if anything, attenuated by the impact of high *WAGE* on *WELFARE*. However, it is not unexpected that high wage would be related to *WELFARE*.

To sum up, the regression results about the effects of organisation, personnel and finance character-related factors on *WELFARE* suggest that *TYPE*, *PAST* and *WAGE* have a significant positive effect on *WELFARE* and that the inclusion of these control variables does switch the impact of *RANK* to the hypothesised negative sign, although its effect remains insignificant.

5.6 Conclusion

This chapter has analysed the impact of *RANK* and control variables on *PROFIT, SATIS, WELFARE* using the data set for 2014-2019. Before conducting the   regressions, missing data in the customer satisfaction score were deleted considering the mechanism of MCAR and small amounts of the missing data. The preliminary statistical analysis showed the relationship between dependent variables (*PROFIT, SATIS, WELFARE*) and independent variable (*RANK*) and control variables (*TYPE*, *LISTED*, *SIZE*, *PAST*, *WAGE*) through the Pearson correlation coefficient and scattergram. After checking the VIF values of the independent variable and control variables, it was concluded that there was no multicollinearity problem in the regression models.

With respect to the impact of the PIMES evaluation results on the accountability of public institutions, the chapter suggested two main findings. First, the PIMES evaluation result (*RANK*) has not significantly affected the managerial accountability of public institutions measured by *PROFIT* in this study significantly in the PIMES. The interview results provided cause for the quantitative finding- a decrease in the weight of performance indicators regarding financial performance. Second, the PIMES evaluation result (*RANK*) has affected the social accountability of public institutions measured by *SATIS* in this study positively with statistical and practical significance in the PIMES. This means that the public institutions with better *RANK* are more likely to have higher *SATIS*. This quantitative finding was supported by the interview results that suggest the link between the questions of customer satisfaction survey and the missions of public institutions and use of the customer satisfaction results in management by public institutions. The regression results of control variables suggested that all the control variables have not affected *PROFIT* and *SATIS* significantly except the positive effect of *TYPE* on *SATIS*. With respect to the strategic behaviour of public institutions, this study suggested that *RANK* has not affected *WELFARE* significantly in the PIMES. This finding was supported by the interview result that focus of evaluation of *WELFARE* was changed from reducing the amount of *WELFARE* to decreasing the difference of *WELFARE* between employees within a public institution.

The next chapter analyses the effectiveness of incentive schemes, performance indicators and target-setting method of the PIMES and explores the strategic behaviour of public institutions in practice.

Chapter 6 An examination of incentive schemes, performance indicators and strategic behaviour

6.1 Introduction

While the previous chapter analysed how the PIMES has affected the accountability of public institutions over time, this chapter examines the effectiveness of current PIMES in incentivising performance and explores the strategic behaviour of public institutions under the PIMES. As explained in Chapter 2, the design and use of performance management systems are a main agenda in performance management research (Ferreira and Otley, 2009; Bracci et al., 2017). The design and use of performance management systems involve reward systems, performance indicators and target-setting methods (Agostino and Arnaboldi, 2012; Bracci et al., 2017). Therefore, the examination of operational methods of PIMES such as financial and non-financial incentive schemes and the target-setting method sheds light on the main agenda of research on performance management systems. In this chapter interview quotes from key participants in the PIMES evaluation are used to examine the effectiveness of incentive schemes, performance indicators and target-setting method in the PIMES. The interview quotes provide empirical evidence on why the non-financial incentives scheme is effective in increasing the accountability of public institutions, and also show why the financial incentive scheme is ineffective in the PIMES. Moreover, the interview quotes present empirical evidence on why performance indicators can at times paradoxically decrease the accountability of public institutions and how the strategic behaviour of public institutions occurs under the target-

setting method in the PIMES.As explained in Chapter 3, the PIMES aims to improve accountability and performance of public institutions by using incentive schemes. Incentive schemes may improve accountability to citizens by increasing the quality of service and the efficiency of the public sector as the NPM reforms suggest (Burgess and Ratto, 2003; Lapuente and Van de Walle, 2020). However, too many incentives may reduce public accountability or performance rather than increase it (Van Dooren et al., 2015). In order to shed light on the academic debate about the relationship between performance-based incentive schemes and public accountability and performance, the chapter examines how the performance-based incentive schemes work in practice in the PIMES, presenting findings from interviews with the key participants in the PIMES evaluation process.

As mentioned in Chapter 2, performance indicators are relevant to incentivisation because good performance can be rewarded and poor performance can be identified through performance indicators (Bird et al., 2005). However, using many performance indicators in performance management systems can create high costs on those providing the information that informs the performance indicators (Bird et al., 2005; Van der Geer, Van Tuijl and Rutte, 2009). Moreover, strategic behaviour and gaming can occur in achieving targets in performance indicators in performance management systems (Hood, 2006; Taylor, 2021). This chapter examines how the structure of performance indicators and target-setting method in the PIMES evaluation process affect the performance and accountability of public institutions, presenting findings from the interviews and analysis of official documents from the South Korean government.

As explained in Chapter 2, there are many types of strategic behaviour and gaming that can arise under performance management systems in which targets are accompanied by a large set of

performance indicators (Hood, 2006; Terman and Yang, 2016; Taylor, 2021). On the basis of principal-agent theory, this chapter explores the existence and extent of strategic behaviour and gaming in the PIMES evaluation process. More specifically, the chapter explores the strategic behaviour and gaming of public institutions in three categories of activities: producing the management performance report, conducting tasks and preparing for the customer satisfaction survey. These three activities were found to be subject to strategic behaviour and gaming according to the interviews and articles from mass media. Focusing on the three activities, therefore, is more appropriate to explore the strategic behaviour and gaming of public institutions.

The chapter commences with an examination of incentive schemes in the PIMES through the overall grade given to an organisation in terms of public accountability and performance. The main finding is that the impact of incentive schemes in the PIMES on the performance of public institutions is different between the financial (additional pay) and non-financial (announcement of overall grades) incentive schemes. Subsequently, the structure of performance indicators and the target-setting method in the PIMES evaluation process are examined with respect to incentivisation. The chapter presents findings on several limitations: heavy burden on the relatively small public institutions and somewhat irrelevant and manageable target-setting. Lastly, the chapter explores the strategic behaviour and gaming of public institutions in the three categories of activities: production of the management performance report, the process of work and the preparation of the customer satisfaction survey. The chapter finds noticeable asymmetric information between the public institutions and the evaluators and considerable strategic behaviour and gaming of the public institutions in the PIMES evaluation process.

6.2 Incentive schemes in the PIMES

As seen in Chapter 3, the PIMES has several performance-based incentive schemes to induce public institutions to improve their accountability and performance: annual announcement of overall grade and customer satisfaction level and provision of a reward according to overall grade. The overall grade plays an important role in the performance-based incentive schemes in the PIMES. Most interviewees were of the opinion that the overall grades of South Korean public institutions are primarily determined by the evaluation of the public institutions' submitted management performance reports. Thus, preparation for the management performance reports is firstly analysed, which is followed by the findings and analysis regarding the overall grade and performance-based incentive schemes.

6.2.1 Not a level playing field

In order for performance-based incentive schemes to function properly, overall grades need to be determined by the extent of performance and accountability of public institutions through fair evaluation of management performance reports. As explained in Chapter 3, the South Korean public institutions are required to submit their management performance report to the MOEF and responsible ministries by 20th March each year. The report consists of metrics against each of the PIMES' performance indicators. The evaluators evaluate the report and determine scores of all performance indicators. The evaluators produce evaluation reports based on their evaluation and the MOEF finalises the evaluation results and announces overall grades of public institutions in June.

Many interviewees argued that the size of a public institution strongly influences the quality of management performance reports, which hampers a fair evaluation of the reports and affects the evaluation result of overall grade.

> *"I think that the difference between size of the public institutions even in the same category makes a big difference in ability to prepare for the management performance report and eventually affects the result of evaluation. For example, a large public institution can put many employees in producing the report and use consultation with experts on evaluation." [V11]*

> *"As performance indicators increase in the PIMES evaluation process, more activities are required to prepare for the evaluation. As a result, small public institutions like my institution have difficulty in responding to the evaluation. No more than three hundred employees are working for my institution and two employees are responsible for the PIMES evaluation and about twenty employees participate in producing the management performance report." [E4]*

As discussed in Chapter 4, this study defines size of public institution by the number of employees of an institution In this study, a small institution is defined as an institution which has less than three hundred employees and a large institution as an institution which has more than one thousand employees. As of 2019, the small and large public institutions respectively comprise 24% and 36% of all the public institution evaluated in the PIMES. The majority of interviewees suggested that the larger a public institution, the better overall grade it is likely to receive. This finding suggests that there is a non-level playing field according to size of a public institution in the PIMES evaluation process.

The non-level playing field also applies to employees' burden of preparation for the management performance report. Many employees of small public institutions said that they have heavy burden of

preparation for the management performance report.

> *"Hundreds of employees are working in my institution and only three employees are responsible for the PIMES evaluation. The three employees are involved in majority of the work in producing the management performance report over 100 pages and the preparation for evaluation."* [E24]

Producing the management performance report has an element of indivisibility and requires a great amount of time and effort because the report consists of many performance indicators and requires all public institutions to produce heavy paperwork of over 100 pages. Thus, the smaller a public institution, the heavier the proportional burden of producing the report placed on the employees who are responsible for it. The finding suggests that the PIMES evaluation process hinders the improvement of performance of small public institution due to the heavy burden of producing the report. For example, one evaluator said that the cost is larger than the benefit in the PIMES evaluation due to the burden of heavy paperwork.

> *"Many public institutions are investing excessive time and resources in producing the management performance report. I think that the cost of evaluation outweighs the benefit of evaluation and hence we should find a way to reduce the public institutions' burden."* [V8]

This finding implies that it is important to reduce the burden of heavy paperwork for small public institutions.

6.2.2. Overall grade

The overall grade in the PIMES evaluation is determined by ranking according to the total scores of public institutions and consists of S, A, B, C, D, E in diminishing order of excellence as explained in

Chapter 3 (see Table 4-2). Thus, the PIMES is operating a relative evaluation using the overall grade to incentivise the public institutions to improve performance and accountability. Many interviewees said that the overall grade is perceived as a publicly endorsed evaluation result of the public institutions' performance because the South Korean government evaluates the public institutions and announces the evaluation result as the overall grade to the public annually. Given the overall grade's characteristics of relative evaluation and publicly endorsed evaluation result, the overall grade is similar to a published performance league table based on administrative data (Propper and Wilson, 2003). The South Korean government aims to provide public institutions with a motivation for improving performance by using the overall grade that is an example of performance league table. As highlighted by one evaluator, the overall grade in the PIMES functions as a motivation to improve the quality of public services.

> *"The Ministry of Economy and Finance can lead the public institutions to compete to increase the quality of public service through the overall grade. The public might benefit from increased quality of public services through the competition between the public institutions. I think that the overall grade mechanism is similar to release of the Ofsted rating for improvement of educational service in the UK" [V1]*

This study examined whether the overall grade as a performance league table incentivises the public institutions to improve performance and accountability by asking the interviewees the question: to what extent do you think the overall grade as a performance league table contributes to improvement of the performance and accountability of the South Korean public institutions? Most of the interviewees suggested that the overall grade induces the public institutions to compete in order to receive better overall grade.

> *"My institution has received the poor overall grade of 'D' or 'E' for a long time. In*

188

order to escape from the poor overall grades, management pay much attention to the evaluation and employees are actively cooperating when preparing for the evaluation. As a result, my institution's integrity grade has significantly increased from fifth grade in previous year to first grade in this year." [E21]

"I think that the reason for the public institutions to compete for better overall grade is that the heads of public institutions relate the overall grade to their pride. The competition between the public institutions enabled my institution to adopt a new pay system in which pay is determined by the characteristic and difficulty of a job." [E15]

The interview results suggest that overall grade in the PIMES leads the public institutions to improve accountability and performance. The integrity grade in interview *[E21]* means how honestly an institution carries out its tasks. Therefore, the increase of integrity grade of an institution implies that the institution becomes more credible and accountable. In South Korea, the majority of employees' pay in the public sector is determined by the employees' working years not their merit. As the pay system is criticised for not incentivising employees in the public sectors to improve performance, the South Korean government has tried to introduce the new pay system in which pay is determined by the characteristics or difficulty of a job (Bregn, 2008). Many employees said that their institutions try to introduce the new pay system to receive better overall grade in the PIMES despite the opposition of majority of employees. Therefore, it is argued that overall grades in the PIMES as a performance league table contributes to leading the public institutions to improve accountability and performance. The effect of overall grade on public accountability and performance can be clarified by examining the effectiveness of the financial and non-financial incentive schemes according to the overall grade in the PIMES.

6.2.3 Financial incentive scheme in the PIMES

As mentioned in Chapter 2, financial incentives are a salient factor to improve the efficiency and accountability in public sector in the NPM reforms (Perry et al., 2009; Lapuente and Van de Walle, 2020). Based on the principle of NPM, performance-related pay has been adopted extensively to improve performance of organisations in the public sector (Burgess and Ratto, 2003; Belle and Cantarelli, 2015). As explained in Chapter 3, the PIMES is also operating a performance-related pay scheme under which the employees of public institutions receive 0~250% of monthly pay additionally as a bonus every year according to their institutions' overall grades (see **Error! Reference source not found.**).

The effect of the financial incentive scheme in the PIMES on performance of public institutions was explored through the interviews. The government officials and many evaluators in the interviews suggested that the financial incentive scheme is the most important tool to incentivise public institutions to improve their performance. For example, one evaluator argued for the importance of financial incentives as follows.

> *"I participated in an evaluation system other than the PIMES. In the evaluation system, the evaluation result was not linked to financial incentives of employees. As a result, the quality of evaluation reports was poor and the organisations under the evaluation system did not compete for a good evaluation result. However, the quality of management performance reports is good and the public institutions compete for better overall grade in the PIMES because of the impetus of financial incentives."* [V3]

However, the majority of employees have a different viewpoint from the opinions of government officials and evaluators. First, most employees show indifference to the financial incentives embodied in the overall grades because both the amount of financial incentives and difference in incentive payments according to the overall grades are too small for employees to recognise.

> *"The overall grade of my institution rose to 'B' from 'C' and hence financial incentives of employees increased. However, most employees did not pay attention to the increase of financial incentives. They seem to consider the increase small. I think that most employees prefer more leisure time to small increase of financial incentives."* [E25]

> *"If the overall grade of my institution falls by one grade, the annual financial incentive of an employee decreases by 10% of his or her monthly pay. The decrease is so small that the employees tend to show indifference to the financial incentives."* [E10]

Second, many employees of quasi-governmental institutions said that there is a paucity of the financial incentives, comparing their financial incentives with those of the employees of public corporations. This finding relates to the structure of financial incentives in the PIMES. Whilst the employees of public corporation receive 0~250% of monthly pay as a bonus, the employees of quasi-governmental institution receive 0~100% of that as a bonus every year (see **Error! Reference source not found.**). For example, one employee of quasi-governmental institution said that the financial incentive scheme does not seem to be fair.

> *"Employees of my institution are putting the same or sometimes more time and efforts as the employees of public corporations into the preparation for the PIMES evaluation. However, the employees' bonuses are not even half of those of the employees of public corporations."* [E4]

191

The lower bonus for quasi-governmental institutions decreases the effectiveness of financial incentives for those bodies.

Third, the employees of a few public institutions are sceptical about the financial incentives because they have received little or no bonus due to their continuously poor overall grades. For example, one team leader of public institution said that the difficulty in preparing for the PIMES evaluation was due to continuous poor overall grades of the institution.

> *"Some public institutions including my institution usually receive poor overall grades of 'D' or 'E'. Although I try to lead my institution to receive a good overall grade, most employees are pessimistic about chances of receiving the bonus according to a good overall grade. As the employees think that there is no chance to receive bonus, it is difficult for me to encourage them to improve performance in performance indicators." [E14]*

The financial incentives struggle to motivate public institutions with sceptical view of financial incentives to improve their performance.

Finally, most employees argued that only management is motivated to improve the performance and accountability of its institution by the financial incentives. For example, one employee of a public corporation said that the difference of financial incentives between good and poor overall grades for management is large enough to motivate management.

> *"The difference of management's financial incentives between good and poor overall grades is large. In my institution, the difference amounts to tens of millions won (over £ 10,000) and hence the financial incentive absolutely motivates the management to improve the performance and accountability of my institution."*

[E20]

Consequently, most employees except management in public institutions are not motivated to improve performance and accountability by the financial incentives in the PIMES.

To sum up, it was found from the interview data that whilst government officials and evaluators advocate the effectiveness of financial incentive scheme in the PIMES, most employees do not support the effectiveness of it. As much research on the effect of financial incentive schemes has been undertaken in the literature, this thesis will discuss the effectiveness of financial incentive scheme in the PIMES with the research in Chapter 7 in detail.

6.2.4 Non-financial incentive scheme in the PIMES

The majority of the South Korean public institutions are not evaluated by the market, unlike private companies because most of public services provided by them are not traded in a market and most of the public institutions are not listed on the South Korean Stock Exchange. However, public institutions are evaluated by the South Korean government and the public under the PIMES and the South Korean government announces the evaluation results to the public annually. This announcement can be considered as a non-financial incentive scheme in the PIMES because the announcement does not involve money, however, a good overall grade provides the public institutions with a short run of 'prestige' benefit (Mathauer and Imhoff, 2006; Kefay and Kero, 2019; Propper and Wilson, 2003). Most interviewees gave the opinion that the annual announcement of overall grades in the PIMES influences employees' pride and morale.

> *"I think that the announcement of good overall grades gives the employees of*

public institutions motivation to improve performance because most employees are proud of a good overall grade. For example, three public institutions including my institution in this city received the overall grade of 'A' last year. Thus, employees of the three public institutions were very proud of the overall grade and the three institutions celebrated it by hanging a placard." [E7]

"I think that the announcement of overall grades affects the morale of employees of public institution strongly. As my institution does not usually receive a good overall grade, employees' morale becomes very low. However, when it was announced that my institution received the overall grade of 'C', employees' morale became high and management actively made new business plans because we escaped from poor overall grades of 'D' and 'E'." [E14]

The majority of interviewees suggested that both management and employees of public institutions are affected by the announcement of overall grades rather than the financial incentives embodied in the overall grades. There are three main reasons for this arising from the interview data. The first reason is that the management and employees of public institutions are very concerned about the evaluation of the South Korean government and the public. The management and employees of public institutions said that the announced overall grades determine the reputation of their institutions. For example, the management and employees who received the overall grade of 'A' tended to regard their institution as a publicly recognised good institution. As a result, the management and employees were very proud of their institution and were trying to maintain the good reputation via a good overall grade in next year. As explained in Chapter 2, the positive feedback of an announced good overall grade was found to increase the intrinsic motivation of management and employees to improve their performance and accountability (Sansone and Harackiewicz, 2000; Ryan and Deci, 2017). On the other hand, the management and employees who had received the poor overall grades tended to think that their institution was criticised by the South Korean government and the public. In order to escape

the criticism, the management and employees were found to endeavour to increase their performance.

The second reason is that the public institutions try to follow their reference groups which are all groups that influence behaviour of individuals due to the announcement of overall grades (Dawson and Chatman, 2001). For example, one evaluator suggested that public institutions have their reference groups according to their overall grades.

> *"If a public institution received the overall grade of 'D' or 'E' last year, the public institution is included in low-graded group in this year. Similarly, the public institutions which received the overall grades of 'B' or 'C' and 'A' or 'S' last year are respectively included in middle and high-graded groups in this year. The public institutions in the low-graded group really endeavour to escape from the low-graded group by following the middle and high-graded group as their reference group." [V6]*

The public institutions' effort to follow their reference groups in the PIMES evaluation process can be explained by using reference group theory. Reference group theory suggests that people take the standards of significant others as a basis for making self-appraisals and comparisons (Hyman and Singer, 1968; Dawson and Chatman, 2001; Panagiotou, 2007). Most employees said that their institutions invest many resources in the preparation for the evaluation to catch up with their reference groups.

> *"Management pays much attention to the evaluation and employees actively provide data about their performance when preparing for the evaluation to escape the low overall grade of 'D' or 'E'. For example, we continue to monitor the result of quantitative performance indicators and do our best to prepare for qualitative performance indicators." [E21]*

> *"My institution has received the overall grade of 'B' since 2017. I think that*

employees of my institution do their best to receive the overall grade of 'A' by working with accountability and increasing their performance." [E19]

The third reason is that the announcement of overall grade affects the working conditions of public institutions by influencing cooperation between the public institutions and the responsible ministries and the institution's recruiting. The public institutions liaise with the responsible ministry for their funding and management. Many evaluators said that the better the overall grade a public institution receives, the easier it is for the institution to cooperate with its responsible ministry. For example, one evaluator noted that the public institutions which received the overall grade of 'D' or 'E' might get less funding next year.

"If a public institution receives the poor overall grade of 'D' or 'E', the public institution would be disregarded by its responsible ministry and hence have difficulty in liaising with it. For example, responsible ministries tend to decrease funding of public institutions that were announced to receive poor overall grades." [V6]

Some evaluators said that the announced overall grades of public institutions are one of the factors that job hunters consider when they decide where they get a job. Therefore, public institutions with a good overall grade are more likely to recruit more qualified new staff.

"The announced overall grades affect not only the reputation of public institutions but also the decision of job hunters to get a job at which institution." [V2]

As explained in Chapter 3, the South Korean government conducts a customer satisfaction survey of public institutions annually and the survey result is used as a performance indicator in the PIMES evaluation process pursuant to the Act on the Management of Public Institutions. Moreover, the South Korean government announces the survey result annually separately from the overall grade. The

announcement of the customer satisfaction survey results relates to a non-financial incentive scheme in the PIMES because the announcement of survey results is the release of the result of one of the performance indicators in the PIMES. The interview data suggested that most public institutions are concerned about the announced customer satisfaction survey results and are trying to increase their customers satisfaction level. Two interviewees gave examples.

> *"My institution is operating a paid camping site. My institution made a website for the camping site users and created themed streets to increase the customer satisfaction level."* [E23]

> *"My institution endeavours to increase quality of facilities and education for trainees to maintain a good customer satisfaction level. Trainees evaluate my institution focusing on education contents and facilities."* [E4]

The interview results support the influence of the announcement of customer satisfaction results on the quality of public services. Moreover, the interview results are aligned with the NPM theory's suggestion that the public can function as customers to provide information on social accountability of public organisations (Grandvoinnet et al., 2015).

6.3 Performance indicators in the PIMES

There are many performance indicators in the PIMES to lead the public institutions to improve performance and accountability. This section explores the current structure and target-setting method of the performance indicators to examine the effectiveness of performance indicators in incentivising the performance of public institutions in practice.

6.3.1 Structure of performance indicators in the PIMES

As explained in Chapter 2, there are several ways performance indicators can be used to incentivise performance and accountability of organisations efficiently: (i) measurement of performance controlled by an organisation, (ii) clear relevance to objectives of organizations, (iii) consistency of indicators over time, (iv) clarity of interpretation of monitored performance, and (v) low cost on those providing information (Carter et al.,1995; Bird et al., 2005; Van der Geer et al., 2009). These ways are important for implementing performance management systems through performance indicators. This section focuses on the costs imposed on those providing information to explore the current structure of performance indicators in the PIMES. The cost on those providing information through performance indicators is mainly determined by the number of performance indicators (Van Dooren et al., 2015). In other words, the more performance indicators, the higher is the cost on those providing information. Considering the negative impact of excessive performance indicators on the performance of public institutions, the MOEF (2010) announced that it would decrease the number of performance indicators in the PIMES from 28 to 18 to reduce the burdens on public institutions in 2011. However, despite this, most employees argued that the number of performance indicators in the PIMES had increased continuously for the past few years. For example, one employee noted that performance indicators are too many to be managed efficiently.

> *"I think that the performance indicators are too many and a performance indicator has seven subdivided performance indicators even though the weight of performance indicator is small in the evaluation. Therefore, my institution has difficulty of writing too much on the management performance report." [E2]*

The increase of performance indicators in the PIMES implies there is a heavy paperwork from

management performance reports.

This study analysed the management evaluation reports that are made by the evaluators and released by the South Korean government. As explained in Chapter 3, the performance indicators in the PIMES are sorted into two categories: (i) business administration and (ii) main businesses. In the second category of main businesses, the number of performance indicators is different across all public institutions and hence it is not feasible to observe the change of number of performance indicators in the second category. Therefore, calculation and comparison were made of the number of performance indicators in the first category of business administration in 2014 when the two categories were established and 2019. In the first category, the number of performance indicators applied to the public corporations was 15 in 2014 and 17 in 2019. The number of performance indicators applied to quasi-governmental institutions was 13 for fund-management-type and commissioned-service-type institutions or nine for strong and small institutions in 2014. However, the number increased to 15 across all the quasi-governmental institutions in 2019. Thus, the number of performance indicators applied to public corporations increased by two and the number of performance indicators applied to quasi-governmental institutions increased by two for fund-management-type and commissioned-service-type institutions or six for strong and small institutions during the last five years. Moreover, additional scores regarding job creation and innovative growth were found to be operated apart from the two categories since 2017. The employees argued that the additional scores make preparation for the evaluation more difficult. So, as noted by one interviewee, one additional score still remains in the PIMES evaluation and causes difficulty in preparing for the evaluation.

> *"I think that the additional scores should be included ,,, one or two years after they were introduced. Although the additional score regarding job creation was included in the first category, the additional score regarding innovative growth still*

> *remains and hence imposes additional burden on employees who are responsible*
> *for the evaluation." [E7]*

Additionally, it is every year that South Korean public institutions are evaluated by the many performance indicators. This creates a higher evaluation burden than other comparable systems in other countries such as the UK and France, which evaluate public organisations every 5 years or every 3-4 years. Therefore, it is argued that the increased number of performance indicators in the last five years and the annual evaluation in the PIMES impose heavy burden. The heavy burden takes up the employees' time to work and hence it hinders the public institutions in their efforts to increase performance. The heavy burden of preparing for the PIMES evaluation was noted by interviewees.

> *"The public institutions should submit the management performance report in*
> *March and receive the due diligence by evaluators in May. As a result, the*
> *employees responsible for the PIMES evaluation cannot do anything other than*
> *preparing for the evaluation in February and April every year." [E18]*

> *"I think that it is reasonable for the public institutions to be evaluated every two*
> *years. As the public institutions are evaluated every year, most public institutions*
> *start to prepare for the evaluation in July after they receive the evaluation results*
> *in June. Therefore, most public institutions spend a lot of time preparing for the*
> *evaluation." [E6]*

The interview data shows that many performance indicators and annual evaluation in the PIMES can cause dysfunction of evaluation in that public institutions can spend more time preparing for the PIMES evaluation than improving performance.

Lastly, the problem of redundancy of performance indicators was found to occur as the number of performance indicators in the PIMES increased. For example, one interviewee drew attention to the increase of performance indicators during last few years and the redundancy of some performance

indicators.

> *"The number of qualitative performance indicators has increased significantly since 2016. Consequently, many redundant factors occurred. For example, certain factors are included in several performance indicators simultaneously. As a result, employees should invest more time in producing the management performance reports." [E9]*

 Redundant performance indicators cause unnecessary cost on the public institutions because they are providing the same information several times in the management performance report. This unnecessary cost on the public institutions by redundant performance indicators can be considered as an administrative burden which means an individual experience of policy implementation as onerous (Herd, DeLeire, Harvey and Moynihan, 2013). In social welfare programs, administrative burden may be made to limit payment to ineligible claimants, but it also serves to limit take-up of benefits by eligible claimants (Herd et al., 2013). Similarly, in the PIMES evaluation, redundant performance indicators may be made to increase performance of public institutions, however, they were found to cause unnecessary paperwork and decrease performance of public institutions.

6.3.2 Target-setting method in the performance indicators

Targets are salient factor for the performance indicators to incentivise performance because monitoring organization's progress towards the targets contributes to improving performance (Carter et al., 1995; Treasury, 2004, cited in Townley, 2008). As explained in Chapter 3, whilst the MOEF set targets in the first category of business administration, the public institutions can propose new targets in the second category of main businesses and discuss it with the evaluators. This section

focuses on the target-setting method in the second category. Although there are several target-setting methods in the PIMES evaluation, the most frequently used method of giving targets is according to standard deviation of previous performance. The method sets targets based on the standard deviation of performance for three or five years. Employees said that most targets of quantitative performance indicators increase by about 20% than the targets in the previous year. Government officials and evaluators suggested that a target based on the standard deviation of last year's performance provides the public institutions with challenging targets and hence is useful for improving performance. As the government officials and evaluators expected, the employees said in the interviews that the targets set by the target-setting method are challenging and require them to invest much resources and effort in their tasks. However, employees suggested that the increase in the levels of targets means thattheir institutions can face impossible targets two or three years after targets of new performance indicators are set. They named a target that cannot be achieved by their efforts an 'impossible target' and argued that their institutions have some 'impossible targets' in the performance indicators in the second category of main businesses.

> *"There are ten quantitative performance indicators in the second category in my institution. I think that two or three performance indicators among them are 'impossible targets'." [E3]*

> *"The volume of goods and productivity in my institution faced a limit to growth. Therefore, I think that it is impossible to achieve the target about the volume of goods and productivity by increasing them by 20% every year by our efforts alone." [E24]*

The employees said that the government officials and evaluators have recently accepted a practice of changing performance indicators three years after new performance indicators are set, as they know the difficulty of achieving the targets as time passes. Consequently, it was found that public

institutions tend to propose new targets and performance indicators and discuss them with evaluators every three years to remove unachievable targets. Many evaluators argued that they have difficulty in refuting the employees' opinions about the necessity of new targets and performance indicators when they discuss the change of performance indicators with the employees. The interview with an evaluator below shows well how the evaluators suffer from information inferiority.

"When the evaluators discuss change of the performance indicators, the employees try to ensure that new targets and performance indicators are adopted. The employees are experts in the field of their public services, however, evaluators do not have the same knowledge about the public services and public institutions as the employees. As a result, evaluators are likely to be controlled by the employees' arguments." [V2]

With respect to the change of targets and performance indicators in the PIMES evaluation process, one aim of this thesis is to examine whether strategic behaviour and gaming occur between employees and evaluators. If the target-setting method in the PIMES efficiently prevents the ratchet effects (see Chapter 2) from occurring, as suggested by government officials and evaluators, the targets of performance indicators can contribute to improving accountability and performance of the public institutions.

Government officials and some evaluators suggested that the current target-setting method induces the public institutions to improve performance.

"The current target-setting method gives the public institutions challenging targets. Therefore, the target-setting method is more helpful for the institutions to improve performance than previous target-setting method." [G1]

203

However, the interview data with employees and another evaluators provided a string of empirical evidence on the ratchet effects under the current target-setting method in the PIMES. The first evidence is that public institutions are trying to set new targets and performance indicators that are as manageable as possible. For example, one interviewee said that public institutions tend to set easily achievable targets.

"Unusual things that some public institutions always receive almost perfect scores in all the quantitative performance indicators happen. I think that the public institutions may set the quantitative performance indicators and the targets as manageable as possible to achieve them easily." [E2]

In addition, one evaluator noted that public institutions convince evaluators to accept manageable targets.

"The employees' most important mission in preparatory stage of the evaluation is to make new targets and performance indicators as manageable as possible and discuss them with the evaluators. In the process of discussion, the employees try to persuade the evaluators to accept the new targets and performance indicators." [V1]

A second piece of evidence is that public institutions tend to set targets and performance indicators which are somewhat irrelevant to their main missions. This tendency is noticeable when it is difficult to achieve the target and performance indicator relevant to a main mission. For example, one interviewee suggested that the majority of public institutions do not employ targets which are relevant to their main missions.

"My institution is reluctant to set performance indicators that show my institution's performance better than the present performance indicators. I think that almost all the public institutions are using manageable performance indicators to achieve easily rather than the performance indicators which show their missions best." [E9]

Lastly, a government official suggested that in practice most public institutions tried to set lower targets than their abilities in certain performance indicators.

> *"I think that almost all the public institutions propose their targets conservatively. For example, I adjusted the targets that the public institutions submitted to the MOEF after analysing their plans. The public institutions seemed to have buffer in their plan to achieve their targets easily." [G4]*

Considering the interview data, it is argued that a target-setting method based on the standard deviation of previous performance fails to prevent ratchet effects when new targets and performance indicators are set. This thesis will discuss the ratchet effects which occur under the current target-setting method in the PIMES in more detail in Chapter 7.

6.4 Strategic behaviour and gaming under the PIMES

This section explores strategic behaviour and gaming of the public institutions in the PIMES evaluation process using principal-agent theory (see Chapter 2). The exploration of the strategic behaviour and gaming in the PIMES evaluation process relates to the second research question (*How do the conditions and processes of performance management systems affect the accountability of public organisations*) as well as the third research question (*What is the impact of performance management systems on the strategic behaviour of public organisations?*). The MOEF endeavours to mitigate the strategic behaviour and gaming of the agents of the public institutions through the PIMES. However, the interview data and articles in the mass media suggest that the public institutions make the most of asymmetric information and exhibit moral hazard such as manipulation and threshold

effects. This section firstly explores significant strategic behaviour and gaming of the public institutions when they produce management performance reports. Secondly, the section examines whether the threshold effects occur when the public institutions conduct their tasks. Lastly, the strategic behaviour and gaming of public institutions in preparing for the customer satisfaction survey are explored.

6.4.1 Adverse selection in a management performance report

As mentioned earlier, all the public institutions are required to submit management performance report to the MOEF by March and are evaluated based on the reports every year. The public institutions include all performance for a year in the reports in accordance with the performance indicators. The information of the reports presented by the public institutions is an example of signalling, which means an action taken by informed agent to send information to a less informed principal (Perloff, 2018). If the signalling works properly in the PIMES evaluation process, it is argued that adverse selection is mitigated in the PIMES. Adverse selection means that an agent has information that a principal cannot have and hence the information can be used for the agent's private benefit (Gauld, 2007; Bertelli, 2012; Zwalf, 2022). Thus, this section firstly explores how the signalling in the PIMES evaluation process works in practice. Interviews explored the way information was used in compiling the report. Most employees argued that they include good and poor performances for a year in the management performance reports because the performance indicators in the PIMES evaluation require specific and detailed performance. For example, one leader of large public institution presented how all performance of his institution is included in its management performance report.

206

"My institution conducts a total of 36 tasks and all the tasks are distributed to divisions in my institution. All the employees in the divisions are responsible for including their performance in the management performance report and the performance is also connected to internal evaluation. Therefore, my institution is including all performance in the report." [E19]

The performance indicator of 'equal opportunity and social integration' requires all performance about hiring the disabled, youth, regional talent, immigrants and achieving more than a certain percentage of female executives. Another interviewee showed how poor performance in relation to the performance indicator of 'equal opportunity and social integration' was included in her institution's management performance report.

"I have several experiences to show my institution's poor performance. For example, the performance indicator of 'equal opportunity and social integration' requires the performance about employing North Korea defectors or immigrants. However, my institution had no performance in the performance indicator and hence my institution had to show the poor performance in the performance indicator." [E14]

According to the interviews with employees, signalling appears to work well in the PIMES evaluation process. However, the evaluators' opinions about sufficiency of information were different from the employees' opinions. Some evaluators suggested that they do not have enough information to evaluate the performance of public institutions and hence have to find further information in the mass media or require the public institutions to provide further information. For example, one evaluator said that public institutions do not provide information about issues previously criticised by evaluators and improvement related to such issues.

"I consider the management performance report a communication between the evaluators and the public institutions. However, I think that it is difficult to fully

understand what the public institutions are doing through the report. For example, the public institutions should present points criticised by evaluators last year and improvement in the report, however, some public institutions intentionally remove the points from the reports." [V10]

Moreover, one evaluator noted that public institutions had intentionally omitted poor performance several times.

"I found some cases that the public institutions had announced plans in the last year's management performance reports, however, they wrote nothing about the plans in this year's reports. For example, some public institutions presented plans to increase their revenue by expanding paid facilities or to decrease debt by selling their assets in last year's reports. However, I found nothing about the performance regarding the plan in their reports in this year." [V9]

The interview data from above evaluators suggests that public institutions take steps to hide their poor performance to receive a good overall grade.

Furthermore, an evaluator presented empirical evidence that the signalling in the PIMES evaluation process has a significant problem.

"There are some cases that the public institutions intentionally omit poor performance or significant accidents to receive a good overall grade. For example, when I evaluated last year, a public institution intentionally omitted a physical accident from its management performance report. The accident was so critical that it could decrease the overall grade of the public institution. The evaluation team including me found the physical accident in the evaluation process and reflected it on the management evaluation report." [V11]

The evaluators argued that poor performance and physical accidents that significantly affect the

evaluation result should be included in the management performance reports pursuant to guidance on the management evaluation. For example, casualties in process of education should be reported in the performance indicator of 'safety and environment'. However, evaluators said that casualties could be intentionally omitted in the type of public service of providing students with training services (see Chapter 3). If the omission of poor performance and physical accidents from the management performance reports occurs frequently, as the evaluators suggested, it is argued that the signalling fails to work properly in the PIMES evaluation process. The failure of signalling causes the problem of adverse selection in the PIMES evaluation process. In other words, a public institution with poor performance or physical accident might receive a good overall grade by hiding the poor performance or physical accident. Therefore, it is argued that the problem with signalling such as omission of poor performance and physical accidents can cause adverse selection in the PIMES evaluation process and decrease of the accountability and performance of public institutions.

6.4.2 Moral hazard in a management performance report

Based on the information of the management performance reports, the evaluators invest their time and efforts in verifying and evaluating the performance of public institutions. The efforts of the evaluators are an example of screening according to principal-agent theory, which means an action taken by an uninformed principal to determine the information possessed by an informed agent (Perloff, 2018). If the screening in the PIMES evaluation process works properly in practice, it is argued that the problem of adverse selection is mitigated in the PIMES. Moreover, the screening in the PIMES evaluation process can address the problem of moral hazard, which means the tendency of an agent who is imperfectly monitored to engage in dishonest or otherwise undesirable behaviour

(Bertelli, 2012; Mankiw and Taylor, 2020; Zwalf, 2022).

The interviews with the evaluators provided empirical evidence on the limitations of screening in the PIMES evaluation process, as evaluators often lack the time to do a thorough job.

> *"Based on my experiences in the PIMES evaluation process over the past few years, I advised new evaluators in my team to read the management performance reports over the past three years. However, evaluators are usually selected in February or March and the PIMES evaluation process starts in March, therefore, selected evaluators lack time to read the reports over the past three years." [V8]*

> *"Evaluators lacked time to evaluate the public institutions and hence my evaluation team had meetings even at dawn. I think that it is important to give evaluators more time to evaluate. Moreover, there are evaluators who have had no experience to evaluate among evaluators in the PIMES, therefore, more education is necessary for them." [V3]*

The lack of time to prepare for the evaluation and evaluate the performance of public institutions in the PIMES evaluation process hinders the evaluators' screening. Consequently, the interview data and articles in the mass media provided findings that the public institutions commonly commit moral hazard such as output distortion and great exaggeration in the PIMES evaluation process. For example, an evaluator highlighted that output distortions like accounting fraud are occurring in the management performance reports.

> *"When I evaluated the public institutions, there were some cases that the public institutions distorted output by deliberately adjusting period or performance according to their targets like accounting fraud." [V8]*

In addition, an evaluator said that exaggeration of performance is commonly found in the PIMES evaluation process.

> *"I think that there is a possibility that the public institutions distort output or greatly exaggerate performance. For example, when I analysed the public institutions' release of their performance, I found that some release is greatly exaggerated because their performance was poor in fact."* [V10]

However, output distortions are under-reported because of the awareness of being sued. For example, one interviewee noted that there was a case of lawsuit against the South Korean government for a poor overall grade. As a result, the evaluators may be reluctant to find output distortions by the public institutions and reflect them in the PIMES result due to the danger of lawsuit. One evaluator showed how output distortion is under-reported because of the danger of lawsuit in the PIMES evaluation process.

> *"I think that there are many cases that the public institutions distort their output. However, it is difficult for the evaluators to find all the output distortions and lower the PIMES evaluation results by reflecting the output distortions. It is because if a public institution's evaluation result is lowered due to the reflection of output distortion by the evaluators, the public institution would strongly claim that its evaluation result is not correct and may sue the evaluators for an incorrect evaluation."* [V6]

As explained in Chapter 2, signalling and screening are useful methods to address the problem of adverse selection and moral hazard that occur in a principal-agent relation (Perloff, 2018; Pindyck and Rubinfeld, 2018). However, the screening in the PIMES evaluation process is found to have limitations in that the evaluators lack the time to prepare for the evaluation and to evaluate public institutions. As a result, the problem of moral hazard such as output distortion and exaggeration of performance is found to occur frequently in the PIMES evaluation process. Therefore, it is argued that the PIMES does not fully mitigate the moral hazard of public institutions and can fail to increase

the accountability and performance of public institutions due to the limitation of screening in the PIMES evaluation process.

6.4.3 Threshold effects in the performance of public institutions

As explained in Chapter 3, all the public institutions are evaluated based on similar level of targets in the first category of business administration, which are set by the MOEF. This research explored how the public institutions respond to the similar level of targets when they perform the tasks such as social value and finance by employing the theory of threshold effects (see Chapter 2). The interview data found that some public institutions performing above the similar level of targets decrease their performance to the similar level of targets and hence threshold effects were identified. For example, an interviewee highlighted that there was a common practice in public institutions of decreasing performance.

> *"There are some cases that the public institutions deliberately reduce their performance in this year in order to receive good overall grades in next year as well as in this year." [V9]*

Moreover, an interviewee suggested how threshold effects occur in practice, employing a case of public institution.

> *"I think that a public institution always receives a good overall grade not only by their good performance but also by their many overseas branch offices because the public institution has room for adjustment in its performance through the overseas branch offices every year. If some overseas branch offices are expected to have very good performance, the institution can defer some of expected performance to next year." [V11]*

As literature on gaming suggest that threshold effects occur in top performers (Terman and Yang, 2016; Taylor, 2021), the similar level of targets in the PIMES evaluation process was found to decrease the performance and accountability of public institutions performing above the similar level of targets. However, contrary to the account of threshold effects in literature, the similar level of targets was found to decrease the managerial accountability of the public institutions performing below the similar level of targets. This is because the public institutions' efforts to achieve the similar level of targets might result in decreasing the efficiency of public institutions. One evaluator showed how the similar level of targets decreases the managerial accountability of the public institutions performing below the targets.

> *"The present government emphasises job creation in the public sector and gives the public institutions the similar level of targets about job creation. Therefore, the public corporations noticeably have increased employees to achieve the targets. As a result, the public corporations which have significantly increased employees will have difficulty in making a profit because of the dramatic increase of labour cost."*
> *[V4]*

Furthermore, the interview data suggested that some public institutions intentionally keep their performance low to achieve easily their self-set targets in the second category of main businesses for next few years. As explained earlier, the targets of performance indicators in the second category are set based on the performance over the past three or five years. In other words, the better the performance over past few years, the higher level of the target in next years. Therefore, the public institutions can lower their level of targets by keeping their performance low. The finding of keeping performance low suggests that threshold effects are identified when the public institutions not only face the similar level of targets but also set their targets in the PIMES evaluation process. This thesis will discuss the threshold effects which occur under similar level of targets and individual targets set by the public institutions in the PIMES evaluation process in more detail in Chapter 7.

6.4.4 Manipulation of the customer satisfaction survey

As mentioned earlier, the customer satisfaction survey of the public institutions is conducted every year and the result of survey is used as a performance indicator in the PIMES evaluation process. As analysed in Chapter 5, public institutions try to increase their customer satisfaction level.

> *"My institution is cooperating with the customer companies in marketing activities and holds seminars on common topics with them to increase the customer satisfaction level." [E16]*

As the customer satisfaction results are determined by the customers' evaluation of public services, the selection of customers who participate in the evaluation can affect these results. For example, if a public institution is reported on in its customer survey by only its favoured customers, it is likely to receive a good customer satisfaction grade regardless of its public service quality. Therefore, the public institutions are tempted to manipulate hich customers are consulted in the survey. Moreover, only the public institutions know their full customer list because a public institution's customers are made up of all people and all companies who use its public service. For example, the Korea Railroad Corporation's customers consist of all people who used its railroad service at least once during specific period. The Korea Railroad Corporation makes a customer list from its full customer list and submits the customer list to a customer satisfaction survey agency which designs the survey. Subsequently, the customer satisfaction survey agency extracts a sample from the customer list. In the process of making a customer list from a full customer list, a public institution may intentionally add its employees into the customer list. As a result, there is a possibility that a public institution intentionally omits hostile customers from its customer list or adds its employees to the customer list.

Therefore, the customer satisfaction survey is vulnerable to manipulation. Government officials said that the manipulation of customers by the public institutions had occurred several times and the public institutions are generally tempted to manipulate their customer lists. For example, a public institution was found to have manipulated its customer list very badly in the customer satisfaction survey and hence the public institution received a customer satisfaction score of '0' (see Chapter 5).

> *"Some public institutions were found to have manipulated their customers by putting their employees into their customer lists in addition to the significant manipulation of customers by a public institution last year." [G2]*

> *"The manipulation of result in the customer satisfaction survey by the public institutions occurs every few years." [G3]*

As explained earlier, the PIMES aims to improve social accountability by using the customer satisfaction scores as a performance indicator in the PIMES evaluation. However, the manipulation of customers by public institutions can hinder the PIMES in improving the social accountability of public institutions. Moreover, the manipulation of customers might incentivise public institutions to select favoured customers for the customer satisfaction survey rather than increase their public service quality in order to receive a good customer satisfaction grade.

6.5 Conclusion

This chapter has examined the effectiveness of the relative evaluation using the overall grade, the performance-based incentive schemes and performance indicators in the PIMES evaluation process in improving the accountability and performance of public institutions. Subsequently, the chapter has

explored the strategic behaviour and gaming of public institutions in the PIMES evaluation process. A non-level playing field according to the size of public institutions firstly was found in preparation for a management performance report. The non-level playing field prevents performance-based incentive schemes from working properly because the larger a public institution, the better overall grade the institution receives. Moreover, the non-level playing field hampers the improvement of performance of small public institution due to the heavy burden of producing the report. A relative evaluation using the overall grade in the PIMES was found to induce the public institutions to compete for better overall grade. The performance-based incentive schemes in the PIMES are based on the relative evaluation using the overall grade. It was found that government officials, evaluators and employees have different viewpoints about the effect of financial incentives on the performance of public institutions. Whilst the government officials and evaluators believed that the financial incentive scheme would have a positive effect, most employees disagreed because of their indifference to both the amount of financial incentives and the small difference in incentive payments as overall grade improved. However, most interviewees argued that both management and employees are affected by non-financial incentive schemes in the PIMES because they are concerned about the evaluation by the South Korean government and try to follow their reference groups.

As for the performance indicators in the PIMES, the structure of performance indicators and the target-setting method were analysed to examine the effect of performance indicators on the performance of public institutions. The number of performance indicators in the first category of business administration was found to have increased for 2014-2019. As the public institutions are evaluated in accordance with many performance indicators every year, it is argued that the increased performance indicators impose a heavy burden on the public institutions. Moreover, the redundancy of performance indicators occurred due to the increased numbers of performance indicators and hence

the public institutions often provided the same information several times. Therefore, it is argued that the increased number of performance indicators hinders the public institutions' improvement of performance by taking up the employees' work time. Public institutions are facing challenging targets in the PIMES evaluation process according to the target-setting method based on their performance over past few years. Government officials and employees were found to have different opinions about the effectiveness of the target-setting method. Whilst government officials argued that the target-setting method is useful for leading the public institutions to increase their performance, employees suggested that it is difficult to achieve their targets three years after new targets and performance indicators were set. Consequently, public institutions tend to try to set new targets and performance indicators every three years and hence ratchet effects occurred in the process of setting new targets and performance indicators. New targets and performance indicators were found to be somewhat irrelevant to the public institutions' main missions as well as being manipulated to be manageable.

Lastly, strategic behaviour and gaming were explored in the production of management performance reports, the process of work and the preparation for the customer satisfaction survey. The opinions of the evaluators and the employees were different about the sufficiency of information in the management performance reports. According to the opinions of evaluators, the omission of poor performance and physical accidents occurred frequently in the reports, contrary to guidance on the management evaluation. This problem of signalling suggests a chance of adverse selection that a public institution with poor performance or physical accident receives a good overall grade by hiding nformation. Furthermore, the evaluators lacked the time to prepare for the evaluation and evaluate the public institutions in the PIMES evaluation. The lack of time hindered the screening by evaluators and hence it was found that moral hazard such as output distortion and exaggeration of performance occurred more frequently than the cases reported in the mass media. When public institutions tried to

217

achieve the similar level of targets, threshold effects were identified. Threshold effects in the PIMES evaluation process were found to decrease not only the performance of public institutions above the similar level of targets but also the managerial accountability of public institutions below the similar level of targets. Moreover, threshold effects were found to decrease the performance of public institutions under individual targets as well as the similar targets. It was found that the customer satisfaction survey is vulnerable to the manipulation of customer lists by the public institutions. As a result, there were some cases that the public institutions manipulated customers surveyed by putting their employees into their customer lists. The manipulation of customers hampers the effectiveness of the PIMES evaluation in improving the social accountability of public institutions.

The next chapter interprets and discusses the key findings in Chapter 5 and 6 and discusses this research's contributions.

Chapter 7 Discussion

7.1 Introduction

The purpose of this chapter is to discuss the key findings about the relationship between the PIMES and the accountability of public institutions, the effectiveness of incentive schemes and performance indicators and strategic behaviour and gaming of public institutions. The discussion reflects on the key findings and how they contribute to understanding relevant literature. The chapter discusses empirical, theoretical and practical insights of relevance to South Korean context and beyond. The chapter also suggests this study's unique contributions in terms of theoretical, methodological and empirical dimensions.

The first section reflects on what the findings mean for the first research question: *Do performance management systems improve the accountability of public organisations?* The section discusses the finding that the PIMES evaluation result (*RANK*) has a significant positive effect on the customer satisfaction score of public institutions (*SATIS*). This is interpreted as indicating that there is an element of social accountability in place. Subsequently, two important implications of this social accountability effect are discussed. The first implication is that citizens play an important role in increasing quality of public services as customers play the role in a market system. The second implication is that the public institutions are sensitive to their incentives in the PIMES. The section also discusses the finding from this research that the PIMES evaluation result *(RANK)* does not significantly affect the ratio of profit to assets of public institutions *(PROFIT)*. This research had

taken the PIMES evaluation result to have an impact on the ratio of profit to assets as an proxy of managerial accountability, but did not find this effect in the empirical modelling. The implications of this lack of relationship between the PIMES and the managerial accountability of public institutions are discussed.

The second section discusses what the findings mean for the second research question: *How do the conditions and processes of performance management systems affect the accountability of public organisations?* The section interprets and discusses the effects of financial and non-financial incentive schemes on the accountability and performance of public institutions via the PIMES evaluation process. This study has found that most employees of public institutions profess an indifference to the financial incentives embodied in the overall grades awarded in the PIMES. This section seeks to explain the reasons for this somewhat unexpected finding. This thesis hypothesised that non-financial incentive schemes such as the announcement of overall grade and customer satisfaction level would be effective in strengthening the accountability and performance of public institutions. In this section, the support for the hypothesis is explained in the South Korean context and its contribution to the incentive design in performance management systems is discussed in wider context. In addition, this section examines why performance indicators have proliferated since 2014 and discusses the effect of the greater number of performance indicators on the PIMES evaluation process. Furthermore, the section discusses the implications of the finding that the target-setting method in the PIMES is vulnerable to the strategic behaviour of public institutions.

The third section explores what the findings mean for the third research question: *What is the impact of performance management systems on the strategic behaviour of public organisations?* The section

firstly discusses the finding that the PIMES evaluation result (*RANK*) does not significantly affect the employees' welfare benefit (*WELFARE*). This can be argued to imply that the public institutions can behave strategically when they determine their level of welfare benefit. This study has found that the strategic behaviour occurs more than is generally realised. So, when an institution's performance is assessed by means of its reported performance indicators, such indicators may be subjected to distortion. This sort of strategic behaviour can also have an effect on the performance of public institutions or customer satisfaction results via keeping their performance low or manipulating their customer lists. The section discusses the impact of such strategic behaviour on the accountability of institutions.

The fourth section discusses the thesis's contribution. In terms of theoretical contribution, this thesis finds an extended role of citizens in performance management systems in increasing quality of public services beyond the monitoring services. The thesis suggests that a non-financial incentive scheme is more effective for incentivising public organisations to improve accountability and performance than financial incentive schemes. In addition, the thesis suggests a new theoretical understanding of ratchet effects and threshold effects by identifying '*irrelevant ratchet effects*' and '*individual threshold effects*'. In terms of methodological contribution, this thesis proposes a systematic and complementary research methodology to study performance management systems with quantitative and qualitative data. In terms of empirical contribution, this thesis provides latest findings for understanding the effect of recent reform of PIMES and expands research on NPM beyond Western countries.

7.2 The effect of performance management systems on public accountability

7.2.1 The relationship between performance management systems and managerial accountability

As analysed in Chapter 5, it was found that *RANK* does not have a significant positive effect on *PROFIT*, using the data set for 2014-2019. This finding suggests that the PIMES evaluation has not improved the managerial accountability of public institutions. Why should this finding of no *PROFIT* effect be the case? Firstly, our empirical measures may not be sufficiently appropriate to capture this effect because managerial accountability consists of not only *PROFIT* but also other factors. However, assuming that the proxy measures used have some validity, the reasons for insignificant effect of *RANK* on managerial accountability as reflected by *PROFIT* probably lie in the changes made to the PIMES evaluation process over the years studied. If we compare the regression results for 2014-2016 with the results for 2017-2019 using the regression models set out in Chapter 4, we find that the coefficients estimated for the model change a lot. A new government was inaugurated in May 2017 and the methods in PIMES changed after 2017 in accordance with the new government's philosophy. For example, the weight of performance indicators regarding financial performance in the PIMES evaluation process was significantly reduced by the new government. Accordingly, this research found regression estimates of the effect of *RANK* for 2017-2019 on *PROFIT* to be much less than for 2014-2016.[3]

---

[3] In the bivariate regression model, the regression coefficient of *RANK* for 2017-2019 is -0.00005 and P-value is 0.386 under one-tailed test. However, the regression coefficient of *RANK* for 2014-2016 is 0.00026 and P-value of it is 0.141 under one-tailed test. In the multiple regression model, the regression coefficient of

The absence of this effect of the PIMES evaluation result on the ratio of profit to assets of public institutions over the years studied does not align with a straightforward NPM hypothesis on how to improve managerial accountability. As explained in Chapter 2, the key premise of NPM reforms was that performance management systems would enhance the efficiency of public organisations (Van De Walle and Cornelissen, 2014; Christensen and Lægreid, 2015; Lapuente and Van de Walle, 2020). However, this study found that the PIMES evaluation did not have a significant positive effect on the managerial accountability of public institutions for 2014-2019. Considering the finding, it is argued that the PIMES does not have an effective mechanism to improve performance via managerial accountability.

7.2.2 The relationship between performance management systems and social accountability

As analysed in Chapter 5, *RANK* has a significant positive effect on *SATIS*. This finding suggests that the PIMES evaluation has improved the social accountability of public institutions, as measured by *SATIS*. This study provided many reasons for the finding in Chapter 5. In addition to those reasons, this section discusses growth of literature on analysis of quality of service based on customer satisfaction surveys (Oña and Oña, 2015). The quality of public service has been a major topic for practitioners and researchers in the public transport sector in particular, where there has been a key focus on passengers' perspective in the last few decades (Oña and Oña, 2015). According to 'Service Quality-Satisfaction-Loyalty/ Behavioural Intentions' paradigm, customers' satisfaction is the result of good service quality and results in their positive behavioural intentions (Jen, Tu and Lu, 2011). As

---

RANK for 2017-2019 is -0.00007 and P-value is 0.347 under one-tailed test. However, the regression coefficient of *RANK* for 2014-2016 is 0.00021 and P-value is 0.204 under one-tailed test.

a result, higher *SATIS* represents better quality of public service. Therefore, this study bears out the argument that the PIMES has improved the quality of public services by using the customer satisfaction survey results in its evaluation process.

This argument has two important implications. First, citizens have an important role in improving the quality of public services via their views on customer satisfaction. Although in the private sector citizens have important role in determining the revenue of private companies by choosing which private company they will buy a service from, in the public sector, in contrast, citizens mainly face public sector monopolies (Dixit, 2002). Consequently, it can be argued that public service agencies do not have a competition incentive to provide citizens with high quality services and may lack attention to consumers' preferences (Dixit, 2002). However, in contrast to this argument, South Korean public institutions are found to be responsive to citizens' demands, and are seen to try to improve quality of their public services for better *SATIS* in the PIMES. Therefore, it is argued that citizens may play a similar role in increasing quality of public services in South Korea to that which consumers may play in a market system.

Brinkerhoff and Wetterberg (2016) argue that engaging citizens in holding public service providers accountable is a remedy for increasing the effectiveness of service delivery, reviewing four donor-funded projects with social accountability mechanisms. For example, the Quality Assurance Partnership Committee (QAPC) project in the Philippines carried out customer satisfaction surveys or suggestion/complaint boxes in three public health facilities. As a result, the effectiveness of public service delivery improved and service utilisation increased (Brinkerhoff and Wetterberg, 2016). Their

argument about effect of citizens' engagement in decision makers corresponds to South Korean citizens' role in increasing public service quality through customer satisfaction surveys. Furthermore, the South Korean citizens' role in increasing quality of public service supports the NPM reforms' suggestion that citizens can be seen as clients who reward good public service by loyalty or punish poor public service by exit (Grandvoinnet et al., 2015). On the other hand, the degree for the public institution to respond to customers' demands and endeavour to increase service quality is found to be different between public corporations and quasi-governmental institutions according to the interview data. The difference is noted by comparing the interview with an employee of public corporation with the interview with an employee of quasi-governmental institution.

> *"My institution is trying to increase the quality of restaurants in expressway services and established our own petrol stations in accordance with the government policy of 'thrifty petrol station' to decrease petrol price." [E20]*

> *"Customers' expectation increases continuously. My institution is trying to meet their demands, however, there are cases where my institution does not meet customers' demands." [E17]*

The interview *[E20]* with an employee of public corporation shows that public corporations try to increase the quality of public services actively. However, the interview *[E17]* with an employee of quasi-governmental institution suggests that quasi-governmental institutions are less active in increasing the quality of public services than public corporations. This finding results from difference of intensity of competition between group of public corporations and group of quasi-governmental institutions in the PIMES evaluation process. As public corporations are known to the public more widely than quasi-governmental institutions, the public corporations are more sensitive to *SATIS* and hence intense competition tends to occur in the group of public corporations. Therefore, the public

225

corporations are more active in increasing quality of their services than the quasi-governmental institutions to receive higher *SATIS*. This finding from the interview data is also supported by the regression result that being a public corporation (*TYPE*) has a significant positive effect on *SATIS* (see Chapter 5).

The second implication is that public institutions are sensitive to their financial and non-financial incentives in the PIMES. The significant positive effect of *RANK* on *SATIS* means that the public institutions consider the customer satisfaction survey important and respond to the survey actively. The public institutions' strong response to the survey is attributable to their implicit eagerness to receive more - incentives in the PIMES. This implication suggests that the - schemes in the PIMES can in some cases be strong tools to increase the accountability and performance of public institutions.

To sum up, it is argued that the PIMES has contributed to improving the quality of public services, via its use of the customer satisfaction survey results. Furthermore, this study suggests that the usage of customer satisfaction survey in performance management systems can lead to improvement of social accountability in the public sector.

7.3 The effect of incentive schemes and performance indicators in performance management systems on public accountability

One evaluator highlighted a fierce competition for a good overall grade between public institutions in the PIMES.

"Heads of public institutions are fiercely competing for good overall grade in the PIMES. I think that the public institutions are investing more efforts and resources in receiving a good overall grade rather than in increasing their budget." [V10]

Why are the public institutions competing for a good overall grade? In order to answer this question, it is necessary to examine the effectiveness of financial and non-financial incentives according to a good overall grade in the PIMES process.

7.3.1 The effectiveness of financial incentive scheme

Chapter 6 established the existence of a contradiction relating to the role of financial incentives in strengthening accountability and performance of public institutions in the PIMES. While government officials and evaluators argued that financial incentives were effective in stimulating performance improvement, most employees disagreed with the argument. Under the performance management system of the PIMES designed, as we have noted, to some extent on a basis of the NPM principles, the financial incentive is a salient condition to improve accountability and performance of public institutions (Dunleavy et al., 2005; Lapuente and Van de Walle, 2020). Thus, a rigorous assessment of the effectiveness of financial incentives is important to address the second research question: *How do the conditions and processes of performance management systems affect the accountability of public organisations?*

Interview results and evidence gathered in this study suggest that the effectiveness of financial incentives is limited to the management and the team which is responsible for preparing materials for

the PIMES evaluation in the public institution. The first evidence for this suggestion is that most employees expressed indifference to financial incentives that could be adduced to the PIMES evaluation results. While management receives large amount of financial incentive from achieving a good overall grade, most employees receive relatively little reward from that achievement. Many employees said in the interviews that both the amount of payments resulting from PIMES evaluation results and the difference in consequent incentive payments were too small to spur them to increase their work effort.

A second piece of evidence relates to changing labour motive and work values of many employees. Bessokirnaia (2011) argues that the time when good pay is the only very important labour motive has passed and at present working conditions and relations with colleagues on the job are important labour motives along with good pay. Work values are defined as generalized beliefs about the desirability of certain attributes of work such as pay and working conditions and work-related outcomes such as accomplishment and prestige (Lyons, Duxbury and Higgins, 2006). Work values can be divided into intrinsic work values such as personal interests and benefits that affect society and extrinsic work values such as salary and stability (Herzberg, Mausner and Snyderman, 2011). Through analysis of data from 15,013 South Korean workers, Lee, Lee, Kim and Song (2022) argue that South Korean employees with moderate values for job security, self-fulfillment and job reputation have a higher level of job satisfaction than income-oriented employees. This suggests that while income is not an important factor for South Korean employees' job satisfaction, having balanced work values is the important factor for their job satisfaction.

The third evidence relates to characteristics of South Korean society. While Western countries share

characteristics of cultural individualism, South Korea is a representative collectivistic society (Lim, Choi and Song, 2012). South Korean society has highlighted collective and social interests such as filial piety and ideals of loyalty, therefore, needs of individuals tend to be less emphasised (Lim et al., 2012). In contrast, financial incentives to improve performance are based on rationalistic approaches where individuals are extrinsically motivated and rational players (Belle and Cantarelli, 2015). Thus, financial incentives in the public sector may not resonate as much in South Korean society where collective interest has been considered more important than individual interest. Considering these arguments regarding labour motive, work values of current workers and the characteristics of South Korean society, it seems unlikely that the small amount of financial incentive in the PIMES can incentivise most employees to improve performance.

The limited influence of the financial incentive scheme in the PIMES on the employees of public institutions suggests that in practice, improving performance via financial incentive schemes may work less well than would be indicated by expectations from NPM theory. The next section discusses the effectiveness of the non-financial incentive scheme that satisfies employees' intrinsic work values.

7.3.2 The effectiveness of non-financial incentive scheme

As analysed in Chapter 6, the non-financial incentive scheme in the PIMES such as the announcement of overall grades and customer satisfaction level was found to affect not only management but also employees strongly. Employees' performance is dependent on their willingness to work diligently and carry out necessary tasks as well as their competences (Franco, Bennett, Kanfer and Stubblebine, 2004). Thus, whether incentive schemes affect employees' willingness is very important for

improvement of the performance of organisations. Whilst the financial incentive scheme in the PIMES influences only the willingness of management, the non-financial incentive scheme affects both the willingness of management and employees. Therefore, it is argued that the non-financial incentive scheme in the PIMES is an effective tool to improve the performance of public institutions.

The reasons why the non-financial incentive scheme in the PIMES is the key to improve the performance of public institutions are its impact on the job satisfaction, motivation and pride of employees and the somewhat different values of employees in the public sector compared to employees in the private sector (Franco et al., 2004; Kefay and Kero, 2019; Lyons et al., 2006). Firstly, employees' self-fulfillment and job reputation are raised by a good overall grade that is announced for their institution. As a result, the employees with moderate values for self-fulfillment and job reputation have high level of job satisfaction that is an essential component for their motivation and encouragement towards better performance (Lee et al., 2022; Raziq and Maulabakhsh, 2015). Subsequently, the announcement of a good overall grade improves recognition for a public institution's performance. When the recognition grows through the announcement of a good overall grade, employees' pride increases, therefore, the performance of a public institution improves (Franco et al., 2004; Kefay and Kero, 2019). Lastly, employees in the public sector value contributing to society more than employees in the private sector do (Lyons et al., 2006). In the PIMES evaluation, the good overall grade of public institution refers to the institution's contribution to society through better public service. Thus, when it is announced that a public institution receives a good overall grade, employees of the institution value a good overall grade more than employees in the private sector. Consequently, the job satisfaction and pride of employees of a public institution increase through an announced good overall grade and hence performance of the employees improves.

The effect of non-financial incentive schemes on the performance of public institutions has an important implication when designing an incentive scheme in performance management systems beyond the PIMES. It is that expanding non-financial incentives rather than financial incentives is a more effective way for performance management systems to improve the performance of public organisations. According to the interview data, most employees preferred enhancement of working conditions and reputation to an increase of bonus as motivation. A government official also suggested that enhancing public institutions' reputation and expanding their capacity is more effective as rewards than financial incentives based on his experience of being responsible for the PIMES evaluation for several years.

7.3.3 The effect of performance indicators and target-setting methods on the public accountability and performance

As explained in Chapter 6, in 2010 the South Korean government had recognised the public institutions' burden due to many performance indicators and announced a decrease in the number of performance indicators (MOEF, 2010), however, the number of performance indicators in the PIMES had increased for 2014-2019. Consequently, many employees argued that they struggle to produce the management performance reports owing to the increased numbers of performance indicators and they suggested that there is a degree of duplication and redundancy in the performance indicators. The interview data suggests that the increase in the number of performance indicators results from the South Korean government's perception of PIMES as being the way to solve most public-institution problems. For example, one interviewee said that a performance indicator is added whenever a problem about public institutions occurs.

> *"In the past, the MOEF made performance indicators with the cooperation of evaluators. However, nowadays other ministries tend to add performance indicators to the management performance report to tackle the problems that the ministries are responsible for."* [E6]

Employees said that they are required to make additional reports whenever other ministries than the MOEF add performance indicators to the management performance report. Duplication and redundancy in performance indicators occur because the government adds performance indicators to respond to new public-institution problems but does not remove similar indicators to the new public indicators from the report.

The interview data suggests that the increase of performance indictors in the PIMES is causing the public institutions to invest excessive resources in preparation for the PIMES evaluation instead of performing their missions. For example, one employee said that public institutions are taking on what government should do.

> *"As the government is adding the performance indicators by reflecting the government key polices on the performance indicators, my institution has difficulty implementing the government key policies for a good overall grade in addition to our main missions."* [E11]

Carter et al. (1995) suggest that a system with many performance indicators may be unwieldy and practitioners bemoan the cost of collecting and monitoring many performance indicators. The increase and redundancy of performance indicators can make the PIMES inefficient at improving the accountability and performance of public institutions because the public institutions may invest excessively many resources in producing the management performance reports. Therefore, it would be of value to reduce the number of performance indicators to make the PIMES efficient for improving the accountability and performance of public institutions.

Governance by targets and performance indicators consists of targets specified in measurable form, monitoring systems to measure performance against the targets and feedback mechanisms linked to measured performance (Bevan and Hood, 2006; Gerrish, 2016). Thus, target-setting is an important process in performance management systems. Targets can improve the performance of organisation in some conditions and cause problems in others (Hood, 2012). This thesis discusses important implications of the target-setting method in the PIMES evaluation process to address the second research question: *How do the conditions and processes of performance management systems affect the accountability of public organisations?*

Chapter 6 showed that ratchet effects occur under the current target-setting method in the PIMES evaluation process. The ratchet effects are a little different from the ratchet effects that have largely been discussed in the literature to date. The research on the ratchet effects has focused on the tendency of target setters or those having to achieve targets to set lower targets than their abilities (Hood, 2006; Terman and Yang, 2016; Taylor, 2021). However, this study argues that in the case of PIMES the public institutions tend to set new targets and performance indicators somewhat '*irrelevant*' to their main missions as well as manageable. For example, a public corporation included in the second type to create energy or dig up mineral resources and supply them (see Chapter 3) was found to have the performance indicator of 'recovery of loan' which is irrelevant to its main mission. This thesis names the ratchet effects found in the PIMES evaluation process '*irrelevant ratchet effects*'. This study identified '*irrelevant ratchet effects*' through coding following the eight steps of Tesch (1990). Following the first to seventh step from eight steps of Tesch (1990), the researcher could categorise interviewee quotes which are relevant to ratchet effects into six concepts discussed in Chapter 2. Based on the six categorised concepts, the researcher performed a preliminary analysis, which is the Tesch's eighth step. In the process of analysis, a type of strategic behaviour of public institutions

which had not been researched was found. The researcher named the new strategic behaviour *'irrelevant ratchet effects'*. Public institutions are more likely to be evaluated by somewhat irrelevant and manageable targets in the PIMES evaluation process as new targets and performance indicators are set due to *'irrelevant ratchet effects'*. As a result, the public institutions focus on achieving somewhat irrelevant and manageable targets rather than to try to increase the performance relating to their main missions under the target-setting system.

The *'irrelevant ratchet effects'* can be explained by using several motivations for gaming by public officials (Taylor, 2021). The first is instrumental motivation at an organisational level. Officials may game the target-setting process to avoid the public release of poor performance. The second is self-interested motivation. According to principal-agent theory, officials are self-serving agents, therefore, they engage in gaming in pursuit of personal gains such as a salary bonus and career advancement. The third is intrinsic motivation. A conflict between officials' personal values and their organisation's values can result in gaming. The *'irrelevant ratchet effects'* occur in the PIMES evaluation process due to the instrumental motivation at an organisation level and self-interested motivation. The employees responsible for the PIMES evaluation engage in the *'irrelevant ratchet effects'* to avoid the announcement of a poor overall grades and gain their career advancement.

The current target-setting method in the PIMES evaluation process is subject to the *'irrelevant ratchet effects'*. This has two negative impacts on the ability of the PIMES evaluation to improve the performance and accountability of public institutions. The first is that it is genuinely difficult for public institutions to propose good performance indicators reflecting their main missions in the process of their target-setting. As a result, the targets of some performance indicators are somewhat

irrelevant to the public institutions' missions. The monitoring of the institutions' progresses towards targets that are irrelevant to their missions leads the public institutions to neglect their main missions. The second negative impact is that the public institutions tend to waste much time by replacing difficult-to-achieve targets of performance indicators with manageable targets because they are allowed to do this every three years under the current target-setting method.

7.4 The effect of performance management systems on strategic behaviour and gaming

7.4.1 strategic behaviour and gaming under performance management systems

As described in Chapter 3, the MOEF (2013) announced that it would increase the weight of the performance indicator that related to controlling excessive pay and benefits from 8% to 12% in the PIMES evaluation to control inefficient management of public institutions. Considering the South Korean government's measure in 2013, this thesis analysed the relationship between the PIMES evaluation result (*RANK*) and the average welfare benefit of regular employees *(WELFARE)* for 2014-2019. The regression results showed that *RANK* does not have a significant negative effect on *WELFARE*. This suggests that on the basis of the measure employed here, the PIMES evaluation does not significantly decrease the level of moral hazard of public institutions. The analysis result is contrary to the expected effect of *RANK* on *WELFARE* according to the measure announced by the South Korean government in 2013.

Why should this be the case? In contrast to the insignificant effect of *RANK* on *PROFIT,* the analysis result cannot be explained by change of priority of PIMES evaluation. This is because regression estimates of the effect of *RANK* for 2014-2016 on *WELFARE* are similar to 2017-2019.[4] The insignificant effect of *RANK* on *WELFARE* probably occurs because the public institutions tend to receive similar scores in the performance indicators that relate to controlling excessive pay and benefits regardless of their efforts to reduce pay and benefit. This study analysed the scores in the performance indicators controlling excessive pay and benefits across five categories of public institutions explained in Chapter 3. Except in 2014, the scores for the public institutions in the same category were similar in the data set for 2015-2019. This analysis result suggests that the public institutions had received similar scores in the performance indicators controlling excessive pay and benefits regardless of their level of *WELFARE* for 2015-2019. As a result, *RANK* has no significant effect on *WELFARE.*

The finding suggests that the threshold effects are identified when the public institutions achieve the targets of the performance indicators controlling excessive pay and benefits. As explained in Chapter 6, the targets in the performance indicators in the first category of business administration are similar across the public institutions. The performance indicators controlling excessive pay and benefits are included in the first category and hence the public institutions have the similar level of targets in the performance indicators. Consequently, the public institutions having capacity to reduce pay and benefits more than the similar level of targets tend to reduce their pay and benefits to the similar level

---

[4] In the bivariate regression model, the regression coefficient of *RANK* for 2014-2016 is 0.02163 and P-value is 0.497 under one-tailed test and the regression coefficient of *RANK* for 2017-2019 is 1.24751 and P-value of it is 0.28 under one-tailed test. Both the regression coefficients are positive, whereas expected sign of their effects is negative. In the multiple regression model, the regression coefficient of *RANK* for 2014-2016 is -0.04145 and P-value is 0.491 under one-tailed test and the regression coefficient of *RANK* for 2017-2019 is -0.24723 and P-value is 0.443 under one-tailed test. Therefore, their P-values are similar.

of targets, which is lower than their capacity. As a result, the level for the public institutions to reduce their pay or benefits converge on the similar level of targets. This corresponds to the 'performance target paradox' which means that performance will cluster around the target (Van Dooren et al., 2015). Therefore, it is argued that the similar level of targets in the performance indicators controlling excessive pay and benefits give perverse incentive to the public institutions with capacity to reduce their pay and benefits more than the similar level of target (Terman and Yang, 2016; Taylor, 2021). This suggests that the insignificant effect of *RANK* on *WELFARE* occurs due to the threshold effects.

Chapter 6 showed that moral hazard, such as output distortion and exaggeration of performance occur frequently in the PIMES evaluation process when the public institutions produce the management performance reports. The output distortion and exaggeration of performance suggest that the public institutions engage in gaming of generating positive performance data regardless of achievement of objectives behind performance indicators in the PIMES (Benaine and Kroll, 2019). The gaming in producing the management performance report decreases the accountability and performance of public institutions and causes dysfunction in the PIMES. Firstly, monitoring cost by the government increases, contrary to autonomy of a public institution's management that the PIMES aims to establish. For instance, if output distortion such as converting sales to expected sales increases, the government might need to expand their team of accountants to monitor public institutions' financial statements. Secondly, many resources in a public institution may be transferred from conducting its mission to producing an impressive-looking report containing an exaggerated account of performance. This distortion of priorities in a public institution will reduce outputs (Bohte and Meier, 2000). Thirdly, deliberate omission of poor performance or physical distortion in the management performance report creates a vicious circle between employees and evaluators. Many evaluators said that they should read the management performance reports for last two years and demand additional data from

evaluated institutions to check any omission of poor performance. In contrast, some employees said that they spend much time preparing additional data that is requested. The vicious circle between employees and evaluators causes waste.

7.4.2 The strategic behaviour in performance and customer satisfaction surveys

As analysed in Chapter 6, various threshold effects occur under the similar level of targets in the first category of business administration and the individual targets in the second category of main businesses. This thesis names threshold effects under the individual targets '*individual threshold effects*'. While the literature has discussed the threshold effects under a uniform target applying to all actors (Hood, 2006; Terman and Yang, 2016; Taylor, 2021), the '*individual threshold effects*' were found under individual targets set by the public institutions. Therefore, the '*individual threshold effects*' extend the application of threshold effects from uniform targets to different targets set by actors. The '*individual threshold effects*' result in the decrease of performance of public institutions in the PIMES evaluation process.

Under the similar level of targets of performance indicators, some public institutions performing above the targets decreased their performance to the similar level of target and the similar level of targets decreased the managerial accountability of some institutions performing below the targets. These various types of threshold effects suggest that targets can have an unintentionally negative effect on the performance and accountability of public institutions by their strategic behaviour to achieve the targets (Hood, 2012). In order to mitigate the threshold effects induced by the similar level of targets in the PIMES evaluation process, it is needed for the government to make

sophisticated targets considering different level across the public institutions. For example, one evaluator highlighted the necessity of targets reflecting different working conditions of public institutions.

> *"The probability of an accident happening is very different across the public corporations. For example, the public corporations generating electricity have higher probability of an accident happening than other public corporations. Therefore, it is not reasonable to evaluate safety level of all public corporations using a uniform target." [V7]*

As explained in Chapter 6, the interviews with government officials and evaluators suggested that the customer satisfaction survey is vulnerable to manipulation of customers by the public institutions. In the process of submission of customer lists, addition of the institutions' employees or omission of hostile customers can occur. The manipulation of customer lists corresponds to data manipulation (Hood, 2006; Taylor, 2021), which is similar to a school district's cheat to exempt certain students from an exam to raise the overall pass rate of the district (Bohte and Meier, 2000). The manipulation of customer lists causes mismeasurement of the level of social accountability of public institutions because customer satisfaction results are used to measure the social accountability of public institutions. Moreover, the manipulation of customer lists results in the increase of overall grade as well as customer satisfaction results without improvement of public service quality.

7.5 Contribution to knowledge

7.5.1 Theoretical contribution

The effect of performance management systems on managerial and social accountability

This thesis has analysed and discussed the impact of the performance management system of PIMES on public accountability, focusing on the managerial and social accountability of public institutions. Since the NPM principles focused on performance management systems as a principal means to strengthen public accountability, there has been considerable research undertaken examining whether performance management systems contribute to strengthening public accountability (Lapuente and Van de Walle, 2020; Kenk and Haldma, 2016; Jones and Bouckaert, 2017; Mizrahi and Minchuk, 2019a). For example, in the Australian federal system, performance management was found to have limited capacity for strengthening accountability in federal agreements because performance management established perverse incentives and gaming to obtain the rewards for meeting targets (Jones and Bouckaert, 2017). In Estonian local governments, performance management systems resulted in weak public accountability because the local governments and their stakeholders did not act as collaborative partners in the governance network (Kenk and Haldma, 2016). In the context of the literature on performance management and public accountability, this thesis is unique because there has been little research undertaken examining the effect of performance management systems on managerial and social accountability.

With respect to managerial accountability, this thesis suggests that the PIMES does not influence the managerial accountability of public institutions, which was measured by *PROFIT*. The finding does not align with NPM principles of how to improve managerial accountability through performance

management systems (Christensen and Lægreid, 2015; Lapuente and Van de Walle, 2020). Another finding of this research is that the managerial accountability of public institutions may even decrease when they try to achieve the similar level of targets regarding job creation in the PIMES evaluation process. This finding fills the gap in literature on the effect of target on managerial accountability of organisations by focusing on the negative effect of targets on organisations performing below the targets. Previous research suggests that uniform targets encourage the organisations performing below the target to increase managerial accountability although it gives the organisations performing above the target a perverse incentive (Terman and Yang, 2016). However, this study suggests that the managerial accountability of organisations performing below target can deteriorate as a perverse effect of targets under performance management systems.

Subsequently, this thesis contributes to the literature on the role of citizens in relation to social accountability of public organisations by emphasising the role of citizens. The literature suggests that citizens may have a critical role in monitoring the accountability of public officials who provide public services (Willems and Van Dooren, 2012; Mizrahi and Minchuk, 2019a). However, this thesis does not limit the role of citizens to monitoring, but argues further that citizens have a role to increase quality of public services via their views on customer satisfaction under performance management systems, like consumers in a market system. Furthermore, performance management systems can contribute to improving the social accountability of organisations through collaborative partnership between the government and citizens. The argument corresponds to the NPM discipline that citizens as clients provide information to promote social accountability by assessing quality of public services.

More efficient non-financial incentives than financial incentives in the public sector

241

The majority of countries in the Organisation for Economic Co-operation and Development (OECD) have adopted financial incentive schemes to enhance the motivation and performance of government employees (OECD, 2011; Belle and Cantarelli, 2015). The performance-related pay system as an exemplar of financial incentive schemes applies to employees as well as management in these countries (Perry, Engbers and Jun, 2009). Much research has been undertaken to examine the effects of performance-related pay systems (Perry et al., 2009; Belle and Cantarelli, 2015). Perry et al. (2009) analysed 57 studies evaluating performance-related pay systems for 1977-2008 and found that the systems in the public sector are ineffective in improving employees' performance at the aggregate level. In addition, Belle and Cantarelli (2015) argue that monetary rewards do not enhance effort among civil servants through a vignette experiment with the sample of 296 public managers working in Italian central departments. This thesis finds that the financial incentive scheme in the PIMES impacts only the management and the employees responsible for the PIMES evaluation, but does not influence most employees to improve performance and accountability. The finding results from most employees' indifference to the financial incentives. The finding supports the suggestion that performance-related pay systems fail to fulfil their promise (Perry et al., 2009; Belle and Cantarelli, 2015) by showing the limited effect of the financial incentive scheme in the performance management system of PIMES in Korea where collective interests such as ideals of loyalty are considered more important than individual interests (Lim et al., 2012).

It is important that financial incentive schemes are integrated with non-financial incentive schemes to create a balanced approach to increase motivation and performance (Franco et al., 2004; Mathauer and Imhoff, 2006; Kefay and Kero; 2019). Literature on incentive schemes for health workers

suggests that many non-financial incentive schemes can be more effective tools to improve employees' motivation and performance than financial incentive schemes (Franco et al., 2004). For example, a study undertaken in London and Geneva finds a limited effect of financial incentives on nurses and the relevance of non-financial incentives for nurses' job satisfaction and self-esteem (Kingma, 2003; Mathauer and Imhoff, 2006). However, contrary to abundant research on financial incentive schemes, research on the effectiveness of non-financial incentive schemes has been undertaken in limited areas such as health and education (Propper and Wilson, 2003; Mathauer and Imhoff, 2006; Jalava, Joensen and Pellas, 2015; Burgess, Metcalfe and Sadoff, 2021). This thesis finds that the non-financial incentive scheme in the PIMES influences both management and employees' willingness to work diligently and carry out necessary tasks and job satisfaction. Therefore, this thesis argues that the non-financial incentive scheme is an effective tool to improve the performance of public organisations in performance management systems. The finding contributes to the literature on incentive schemes in performance management systems by showing the non-financial incentive scheme's effectiveness to the seven types of public services (see Chapter 3) that the South Korean public institutions provide beyond the areas of health and education.

Extension of the notion of strategic behaviour and gaming in the literature

This thesis has found various forms of strategic behaviour and gaming of public institutions in the PIMES evaluation process. The ratchet effects, threshold effects, output distortion and exaggeration of performance could occur when the public institutions set new targets, produce management performance reports, perform their tasks and submit their customer lists (Bohte and Meier, 2000; Hood, 2006; Terman and Yang, 2016; Benaine and Kroll, 2019). The finding supports 'Goodhart's

Law', which suggests that numbers used for evaluation will tend to become ineffective over time due to strategic behaviour or gaming by actors evaluated by the numbers (Goodhart, 1984; Hood, 2012). In the PIMES evaluation process, the strategic behaviour and gaming of the public institutions are found to increase over time because the public institutions adapt to the existing performance indicators or try to set new targets and performance indicators every three years.

The thesis suggested '*irrelevant ratchet effects*' and '*individual threshold effects*' beyond the strategic behaviour and gaming surrounding performance management systems discussed in academia (Hood, 2006; Terman and Yang, 2016; Taylor, 2021). Whilst ratchet effects occur when an organisation fixes next year's target as an incremental advance over last years' results, the '*irrelevant ratchet effects*' occur when an organisation not only fixes its target but also replaces difficult-to-achieve targets with new manageable targets. The '*individual threshold effects*' extend threshold effects that have been discussed under uniform targets to individual targets set by an organisation.

7.5.2 Methodological contribution

This thesis has examined the effect of performance management systems on public accountability with quantitative and qualitative data. Most research on accountability has used qualitative designs because the meaning of accountability shifts over time, and information about accountability is often sensitive and can be collected from individuals or groups that have a special role in society (Yang, 2014). Thus, quantitative empirical research on accountability is scarce and little mixed methods research has been undertaken in accountability research (Brandsma, 2014; Yang, 2014). However, researchers need to integrate qualitative design with quantitative design in research on accountability

to enhance credibility and richness of research (Yang, 2014). Therefore, this study using mixed methods research design contributes to research on accountability in terms of research methodology.

While the effect of PIMES evaluation on managerial and social accountability of public institutions was analysed with quantitative data for 2014-2019, the effectiveness of incentive schemes, performance indicators and the target-setting method in the PIMES was examined with qualitative data. This mixed methods research enabled this study to examine the impact of PIMES systematically. In addition, quantitative analysis results were supplemented by qualitative analysis results and vice versa. For example, the quantitative finding that *SATIS* has significantly affected *RANK* in the PIMES was explained by the interview data. As performance management systems are complicated and many components interact within a system (Schleicher et al., 2018; Funck and Karlsson, 2019), this study with quantitative and qualitative data provides an exemplar of research methodology for studying performance management systems.

7.5.3 Empirical contribution

This thesis presents the latest and valuable findings on the effect of PIMES evaluation on public accountability and performance. As explained in Chapter 3, the PIMES has a long history over 35 years and has been reformed several times. Therefore, this study with recent quantitative and qualitative data contributes to understanding the effect of significant reforms of PIMES after 2013. For example, the South Korean government reformed the PIMES significantly in 2013 in order to decrease the debt and excessive benefits of public institutions. This thesis analysed the effectiveness of reformed PIMES in decreasing the average welfare benefit of regular employees (*WELFARE*) since

2014 and suggested that the effect of reform has been limited due to the threshold effects.

In the 2000s, the South Korean government attempted systematic efforts to manage performance in the public sector as a strategy of public sector reform in response to the economic crisis of the late 1990s (Yang and Torneo, 2015; Song, 2015). The efforts were influenced by the NPM principles of disaggregation, competition and incentivisation (Song, 2015; Lapuente and Van de Walle, 2020). For example, the subjects of PIMES evaluation were expanded from 24 government-invested institutions to 89 public institutions in 2004 to promote competition and improve performance in the public sector (Song, 2015). However, this thesis has argued that there was a limited effect of the financial incentive scheme in the PIMES due to South Korean workers' moderate values for self-fulfillment and job reputation and emphasis on collective interest such as ideals of loyalty (Lee et al., 2022; Lim et al., 2012). The argument suggests that improving performance via financial incentive schemes may work less well in practice than would be suggested by the NPM principles. According to Funck and Karlsson (2019), the literature examining NPM has been concentrated on NPM in Anglo-Saxon (31%) and NPM in Continental European-Napoleonic (15%) after 1991. Among the articles, only 5% has studied NPM in Asia (Funck and Karlsson, 2019). This thesis has examined NPM in South Korea, studying the PIMES which is to some extent designed according to the principles of the NPM reforms (Park et al., 2019; Lee and Ra, 2015). Therefore, this thesis contributes to expanding research on NPM beyond Western countries.

7.6 Conclusion

This chapter has interpreted and discussed the findings from the analysis of the effect of the PIMES

on the accountability and performance of public institutions. The chapter firstly discussed the impact of PIMES evaluation on the managerial and social accountability of public institutions. The PIMES evaluation process did not influence the managerial accountability of public institutions because the PIMES evaluation did not provide the institutions with the motivation to increase managerial accountability. However, the PIMES evaluation process contributed to increase the social accountability of public institutions through the performance indicator of customer satisfaction scores. The impact of PIMES on the social accountability of public institutions implies that the South Korean citizens play an important role to increase the quality of public service via their views on customer satisfaction.

Subsequently, the chapter discussed the effect of financial and non-financial incentive schemes, the number of performance indicators and the target-setting method in the PIMES on the accountability and performance of public institutions. The financial incentive scheme does not incentivise most employees to improve performance due to their indifference to the financial incentives and work value for self-fulfillment and job reputation and the emphasis on collective interests in the South Korean culture. However, the non-financial incentive scheme influences both management and employees' motivation and is an efficient tool to improve the accountability and performance of public institutions. The increased number of of performance indicators makes the PIMES evaluation process inefficient and causes South Korean public institutions to invest many resources in producing the management performance reports instead of increasing performance. The target-setting method in the PIMES is subject to the '*irrelevant ratchet effects*', which result in the replacement of good targets and performance indicators reflecting the public institutions' main missions.

Thirdly, the chapter discussed how the PIMES evaluation process affects the strategic behaviour and gaming of public institutions. The similar level of targets controlling excessive pay and benefits provides the public institutions having a great capacity to reduce excessive pay and benefit with the perverse incentive to decrease their efforts to the similar level of targets. Moral hazard such as output distortion and exaggeration of performance occurs frequently when employees produce the management performance reports. The similar level of targets causes the public institution performing above the targets to decrease their performance to the similar level of targets. This thesis argued that '*individual threshold effects*' occur when the public institutions set their targets, whereby the public institutions intentionally keep their performance low to achieve the targets easily for next few years. The manipulation of customers was identified in the customer satisfaction survey due to asymmetric information about the customer lists between the public institutions and the customer satisfaction rating agency.

Lastly, the chapter discussed this thesis's theoretical, methodological and empirical contributions, focusing on literature on the relationship between performance management systems and accountability and performance of public organisations.

Chapter 8 Conclusion

8.1 Introduction

Performance management systems have been pervasive in public organizations and many scholars have become increasingly interested in improving public accountability through performance management systems (Pollitt, 2013; Terman and Yang, 2016; Taylor, 2021). The South Korean government set out the Public Institution Management Evaluation System (PIMES) in 1984 and several changes have been made to it since then in an effort to improve the accountability and performance of South Korean public institutions. In particular the PIMES was reformed in the 2000s based on the principles of NPM and hence the PIMES is to some extent designed according to the NPM theory (Lee and Ra, 2015; Park et al., 2019). As the PIMES has been used as a performance management system in the South Korean public sector, the study of the impact of PIMES on public accountability sheds light on how performance management systems work in relation to public accountability.

This thesis has examined whether and how the PIMES has affected the accountability of public institutions (RQ1), the effect of financial and non-financial incentive schemes, performance indicators and target-setting method in the PIMES (RQ2) and has explored the strategic behaviour and gaming of public institutions in the PIMES evaluation process (RQ3). Both quantitative and qualitative approaches were adopted to address the three research questions in this study, analysing data set for 2014-2019 with the Stata program and analysing 42 semi-structured interviews with

NVivo 12.

This chapter summarises the main findings in relation to the three research questions. The chapter then suggests several feasible policy alternatives for improving the current PIMES. In the remainder of this chapter, limitations of this research and future possible research areas are offered.

8.2 Theoretical and conceptual implications

8.2.1 The relationship between performance management systems and public accountability

Much research has been undertaken to examine whether performance management systems strengthen public accountability (Kenk and Haldma, 2016; Jones and Bouckaert, 2017; Mizrahi and Minchuk, 2019a). Advocates for the NPM principles argue that performance management systems are a principal means to strengthen public accountability (Ospina et al., 2004; Christensen and Lægreid, 2015; Lapuente and Van de Walle, 2020). However, many scholars argue that performance management systems become ineffective in improving public accountability over time due to strategic behaviour of public organisation (Goodhart, 1984; Meyer, 1997; Hood, 2012; Van Dooren et al., 2015). In the context of academic debate, this thesis has analysed the impact of the performance management system of PIIMES on the managerial and social accountability of public institutions.

Using a large panel of key data for 2014-2019, this thesis found that the PIMES has not affected the

managerial accountability of public institutions. According to the regression results, the public institutions would not be motivated by the PIMES evaluation to achieve higher level of managerial accountability because it does not result in a good overall grade which in turn would be connected to more pay and other benefits. The finding does not correspond to NPM theory that see a performance management system as the route to improved efficiency for the public sector. This limitation of the PIMES in respect of managerial accountability was found to have resulted from a decrease in the importance of performance indicators regarding financial performance over the years studied in this thesis.

This thesis has found that the PIMES has contributed to improving the social accountability of public institutions by using the customer satisfaction scores as a performance indicator. The regression results showed that the public institutions which receive better overall grades in the PIMES evaluation are more likely to have higher level of social accountability. The regression results were supported by the interview data and literature on analysis of quality of service based on customer satisfaction surveys (Jen et al., 2011; Oña and Oña, 2015). Based on the finding, this thesis argued that the South Korean public institutions are responsive to citizens' demands and endeavour to increase quality of public services in the PIMES evaluation process. Therefore, this thesis suggested that the use of customer satisfaction survey in performance management systems can result in improvement of social accountability of public organisations.

8.2.2 The effectiveness of operational methods in performance management systems

The NPM reforms have goals to improve the efficiency of the public sector by rewarding for specific

performance through a performance-related pay system (Dunleavy et al., 2005; Lapuente and Van de Walle, 2020). As performance-related pay system was pervasive in the public sector, rich research to examine the effects of the system has been undertaken (Perry et al., 2009; Belle and Cantarelli, 2015). In the context of pervasive performance-related pay system, this thesis has discussed the effectiveness of financial incentive scheme in the PIMES. The thesis found that most employees of public institutions are not incentivised to improve performance by the financial incentive scheme. The limited effect of financial incentive scheme in the PIMES is attributed to the employees' indifference to amount and difference in incentive payment, the increase of South Korean employees with moderate values for self-fulfillment and job reputation, and characteristic of collectivistic South Korean society (Lee et al., 2022; Lim et al., 2012). The thesis suggests that the effect of financial incentive scheme in performance management systems is dependent on actors' interest in financial incentives and work value.

Contrary to the limited effect of financial incentives, this thesis argued that the non-financial incentive scheme in the PIMES contributes to improving the accountability and performance of public institutions. The effect of non-financial incentives in the PIMES results from South Korean employees' high job satisfaction that results from a good overall grade and the high value on contribution to society placed by employees in public sector. The thesis argues that expanding non-financial incentives rather than financial incentives is a more effective incentive design in performance management systems in the public sector.

With respect to performance indicators, this thesis found that the increase of performance indicators in the PIMES since 2014 imposes a heavy burden on many employees who are involved in assembling them. This finding implies that the public institutions may invest excessive resources in preparing for

the PIMES evaluation. Therefore, the thesis suggested that the PIMES evaluation process becomes inefficient and employees bemoan the cost of collecting and monitoring these many performance indicators (Carter et al., 1995). The current target-setting method in the PIMES evaluation process was found to cause the strategic behaviour of employees when public institutions are allowed to change difficult-to-achieve targets and performance indicators every three years. This thesis suggested that the strategic behaviour occurs owing to instrumental motivation at an organisational level to avoid public release of poor performance. This thesis named the strategic behaviour *'irrelevant ratchet effect'* because employees tend to set targets somewhat irrelevant to their missions as well as setting them at a manageable level. The *'irrelevant ratchet effect'* suggests that monitoring of public institutions' progress towards targets hinders the improvement of performance under performance management systems and public institutions waste much time by finding new manageable targets.

8.2.3 Strategic behaviour and gaming under performance management systems

This thesis found various instances of strategic behaviour or gaming by the public institutions when they reduce their excessive pay and benefits, produce their management performance reports and conduct their tasks and when the customer satisfaction survey is carried out. First, the threshold effects are identified when the public institutions face the similar level of targets of performance indicators controlling excessive pay and benefits. The public institutions having capacity to reduce excessive pay and benefits more than the similar targets were found to reduce their efforts to the similar level of targets lower than their capacity. Consequently, the level for the public institutions to reduce excessive pay or benefits converged on the similar level of targets (Van Dooren et al., 2015).

253

The thesis found that public institutions engage in gaming in producing their management performance reports through output distortion and exaggeration of performance. This gaming results from employees' concern about the negative impact of poor overall grade on their institutions' reputation and aspirations of employees responsible for producing the report such as career advancement and an intense competition in the PIMES process. The thesis argued that the gaming involved in producing the reports decreases the performance of institutions through increasing monitoring costs for the government, inefficient use of many resources in producing an impressive-looking report.

Some public institutions were found to indulge in gaming by initially keeping their performance low under the targets set by themselves. This thesis named this gaming *'individual threshold effects'* because the threshold effects occur under individual targets unlike under uniform target discussed in literature (Hood, 2006; Terman and Yang, 2016; Taylor, 2021). While the threshold effect has negative impact on performance of organisations above uniform target, the *'individual threshold effects'* have negative impact on performance of organisations regardless of their capacity.

This thesis argued that currently-used customer satisfaction survey method is vulnerable to gaming by the public institutions because they may submit to the customer satisfaction rating agency lists of customers that have a favourable view of their services. The selection of positively-disposed customers distorts the customer satisfaction result and overall grade and hence limits citizens' role in increasing quality of public services via their views on customer satisfaction.

8.3 Implications and Recommendations for the PIMES

This thesis has analysed and discussed the effectiveness of operational methods of PIMES as well as the effect of PIMES on public accountability. In the process of analysis and discussion, significant implications were found in relation to some limitations of PIMES. With respect to the implications, the thesis suggests some important recommendations for practitioners relevant to performance management systems in the public sector as follows.

Firstly, the similar level of targets in the first category of business administration was found to cause the threshold effects. This study recommends that the South Korean government and evaluators replace the similar level of targets with different levels of targets reflecting capability of each public institution to mitigate the threshold effects. Many evaluators suggested in the interviews that some public institutions face excessive targets beyond their capabilities. For example, some public institutions are more vulnerable to accidents due to characteristic of their public services, however, they have to achieve the similar level of targets in the performance indicator of 'safety'. The excessive targets beyond capabilities of public institutions may result in decrease of managerial accountability of them in the process of achieving the excessive targets.

Secondly, customer satisfaction survey was vulnerable to manipulation of customers by the public institutions. This is because the public institutions have more information about their customers than customer satisfaction rating agency. Therefore, it is necessary that the ministries responsible for public institutions verify lists of customers before the public institutions submit the lists to customer

satisfaction rating agency. The responsible ministries have more information about the institutions' customers than the agency and the public institutions liaise with their responsible ministries for management and budget. Therefore, the verification by responsible ministries can reduce the manipulation of customers due to asymmetric information between the public institutions and the agency.

Thirdly, the thesis' research findings suggest expanding non-financial incentives in performance management systems in the public sector to improve performance of public organisations. The literature on work values argues that employees with moderate values for job security, self-fulfillment and job reputation have higher level of job satisfaction than income-oriented employees (Lee et al., 2022). Moreover, employees in the public sector value contributing to society more than employees in the private sector (Lyons et al., 2006). Therefore, expanding non-financial incentives in performance management systems can lead the employees in the public sector to have high job satisfaction and increase performance (Pollitt, 2013; Lee et al., 2022; Raziq and Maulabakhsh, 2015).

Fourth, there is a non-level playing field according to size of a public institution when the public institutions produce the management performance reports. The non-level playing field hinders the improvement of performance of small public institution due to the heavy burden of producing the report. Therefore, it is important to reduce volume of management performance report for small public institutions. As explained in Chapter 6, producing the report has an element of indivisibility and hence the smaller a public institution, the heavier burden of producing the report is placed on the employees who are responsible for it. Hence, diminishing volume of the report for small public institution can alleviate the problem of non-playing field. Some evaluators recommended replacing

many performance indicators in the report with data in the South Korean government official website about South Korean public institutions to reduce volume of the report.

Fifth, the target-setting method in the PIMES evaluation is subject to the '*irrelevant ratchet effects*'. The public institutions do not propose good performance indicators reflecting their main missions and tend to waste much time by replacing difficult-to-achieve targets with manageable targets. Therefore, it is needed to consider employing different target-setting methods in order to prevent the '*irrelevant ratchet effects*' and '*individual threshold effects*'. Many evaluators and employees suggested in the interviews that one of the reasons for '*irrelevant ratchet effects*' and '*individual threshold effects*' is that only the standard deviation target-setting method is enforced in PIMES evaluation process. As targets become higher by about 20% than the targets in previous year, the public institutions are likely to respond to difficult-to-achieve targets by seeking new manageable targets or initially keeping their performance low. Many employees and evaluators argued that different target-setting methods such as target-setting according to an international standard would lead the public institutions to propose good performance indicators that reflect their mission well and induce them to increase performance.

Lastly, it was found that the evaluators lack the time to prepare for the evaluation and evaluate the public institutions. As a result, the problem of moral hazard such as output distortion and exaggeration of performance occurs frequently in the PIMES evaluation process. To mitigate the problem of moral hazard, this study recommends that the evaluators continue to evaluate public institutions in the PIMES evaluation unless they make an error in evaluation. The evaluators said that more than half of the evaluators evaluate the public institutions in the PIMES evaluation process only for one year and

then are replaced with new evaluators in the subsequent year. In addition, some evaluators argued that it is difficult to run an evaluation team as a team leader because most evaluators of the team evaluate public institutions in the PIMES evaluation process for the first time. As analysed in Chapter 6, deliberate omission of poor performance often occurs by the public institutions, however, the evaluators have difficulty in finding adverse selection of the omission due to lack of time and evaluation experience. It can mitigate the adverse selection to give the evaluators several opportunities to evaluate the public institutions.

8.4 Limitations and future research areas

This thesis analysed impact of the PIMES on the accountability and performance of South Korean public institutions in depth and suggested many findings and implications to contribute to knowledge and policy making about performance management systems in the public sector. However, further research with a similar research topic could enrich the findings and implications of this study.

Most countries are operating performance management systems in the public sector to increase accountability and performance of public organisations. A cross-national comparison between Western countries and non-Western countries would provide researchers with an opportunity to compare various impacts of performance management systems on accountability and performance of public organisations with findings of this study. For example, comparative analysis about the effect of financial and non-financial incentive schemes in performance management systems in other countries could enable researchers to compare the effect of incentive schemes with the finding of this study.

This research conducted quantitative analysis with a data set for 2014-2019 because there is difficulty to compare data set for 2014-2019 with data set before 2013. However, the PIMES was reformed in 2000s based on the principles of NPM and hence further research with data set beginning in 2000 could increase the internal validity of this study.

Managerial and social accountability is complex concept and can be measured by many proxies. However, this research measured managerial and social accountability of public institutions only with *PROFIT* and *SATIS* respectively. Further research with other proxies could enrich the finding about effect of performance management systems on managerial and social accountability of public institutions suggested in this study. For example, 'decrease of debt' can be used as a proxy to measure managerial accountability, and 'number of public service users' can be employed as a proxy to social accountability.

This research focused on the effect of currently-used target-setting method in the PIMES evaluation process because it is enforced in the PIMES evaluation process. As different target-setting methods are used in performance management systems in other countries, further research with other target-setting methods is needed to find out effect of target-setting method on the strategic behaviour and gaming of public institutions.

Lastly, future research needs to consider various behaviours of actors under performance management systems. This research concentrated on the strategic behaviour and gaming of public institutions in the PIMES evaluation process from the perspective of principal-agent theory. However, there is a possibility of cooperation between actors under performance management systems. Further research

with various behaviours of actors under other incentive schemes in performance management systems would enable researchers to design better incentive schemes in performance management systems.

BIBLIOGRAPHY

Abrevaya, J. and Donald, S.G. (2017) 'A GMM Approach for Dealing with Missing Data on Regressors', *The Review of Economics and Statistics*, 99(4), pp. 657-662.

Acock, A.C. (2018) *A Gentle Introduction to Stata*, 6th edn. Texas: Stata Press.

Adcroft, A. and Willis, R. (2005) 'The (un)intended outcome of public sector performance measurement', *International Journal of Public Sector Management*, 18(5), pp. 386-400.

Adler, P.A. and Adler, P. (2012) 'Expert voices', in Baker, S.E. and Edwards, R. (ed.) How many qualitative interviews is enough?: Expert voices and early career reflections on sampling and cases in qualitative research. *National Centre For Research Methods Paper*. pp. 8-11. Available at: http://eprints.ncrm.ac.uk/2273/ (Accessed: 8 January 2021).

Agassi, J. (2011) 'The Philosophy of Social Science from Mandeville to Mannheim', in Jarvie, I.C. and Zamora-Bonilla, J. (ed.) *The SAGE Handbook of the Philosophy of Social Sciences*. London: SAGE, pp. 39-80.

Agostino, D. and Arnaboldi, M. (2012) 'Design issues in Balanced Scorecards: The "what' and "how" of control', *European Management Journal*, 30, pp. 327-339.

Agresti, A. and Finlay, B. (2009) *Statistical Methods for the Social Sciences*. 4th edn. New Jersey: Pearson.

Ahmad, R. (2008) 'Governance, Social Accountability and the Civil Society', *Journal of Administration & Governance*, 3(1), pp. 10-21.

ALIO (2020) *statistics*. Available at: http://www.alio.go.kr/statisticsStat7.do (Accessed: 10 April 2020). All Public Information In-One.

An, S. (2014) 'The Performance of Public Institutions and Firm Characteristics', *Korean Journal of Business Administration*, 27(6), pp. 835-859.

APE (2019) *Annual report 2018-2019*. Paris.

Archibald, M.M., Ambagtsheer, R.C., Casey, M.G. and Lawless, M. (2019) 'Using Zoom Videoconferencing for Qualitative Data Collection: Perceptions and Experiences of Researchers and Participants, International Journal of Qualitative Methods, 18(1-8), pp. 1-8.

Arnaboldi, M., Lapsley, I. and Steccolini, I. (2015) 'Performance Management in the Public Sector: The Ultimate Challenge', *Financial Accountability & Management*, 31(1), pp. 1-22.

Aucoin, P. and Heintzman, R. (2000) 'The dialectics of accountability for performance in public management reform', *International Review of Administrative Sciences*, 66, pp. 45-55.

Baek, S. (2016) *Research on Strategic Behavior in Performance Evaluation of Public Organization: Focusing on metric indicators of management evaluation of public institutions.* Ph.D. Thesis. Seoul National University.

BAI (2015) *Audit results report: performance analysis of overseas resources development business.* The Board of Audit and inspection of Korea.

Barberis, P. (1998) 'The New public management and a new accountability', *Public Administration*, 76(3), pp. 451-470.

Behn, R.D. (2001) *Rethinking Democratic Accountability.* Washington, D.C.: Brookings institution press.

Bejerot, E. and Hasselbladh, H. (2013) 'Forms of Intervention in Public Sector Organizations: Generic Traits in Public Sector Reforms', *Organization Studies*, 34(9), pp. 1357-1380.

Belle, N. and Cantarelli, P. (2015) 'Monetary Incentives, Motivation, and Job Effort in the Public Sector: An Experimental Study With Italian Government Executives', *Review of Public Personnel Administration*, 35(2), pp. 99-123.

Benaine, S.L. and Kroll, A. (2020) 'Explaining effort substitution in performance systems: the role of task demands and mission orientation', *Public Management Review*, 22(6), pp. 813-835.

Benton, T. and Craib, I. (2011) *Philosophy of Social Science: The Philosophical Foundations of Social Thought*. 2nd edn. Hampshire: Palgrave Macmillan.

Bertelli, A.M. (2012) *The Political Economy of Public Sector Governance.* New York: Cambridge University Press.

Bessokirnaia, G.P. (2011) 'The Dynamics of Workers' Values and Labor Motives (2003-2007)', *Sociological Research*, 50(6), pp. 17-31.

Bevan, G. (2006) 'Setting Targets for Health Care Performance: Lessons from a Case Study of the English NHS', *National Institute Economic Review*, 197, pp. 67-79.

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), pp. 517-538.

Bird, S.M., Cox, D., Farewell, V.T., Goldstein, H., Holt, T. and Smith, P.C. (2005) 'Performance Indicators: Good, Bad, and Ugly', *Journal of the Royal Statistical Society. Series A, Statistics in Society*, 168(1), pp. 1-27.

Black, S., Briggs, S. and Keogh, W. (2001) 'Service quality performance measurement in public/private sectors', *Managerial Auditing Journal*, 16(7), pp. 400-405.

Blaikie, N. (2010) *Designing Social Science*. 2nd edn. Cambridge: Polity Press

Bohte, J. and Meier, K.J. (2000) 'Goal Displacement: Assessing the Motivation for Organizational Cheating', *Public Administration Review*, 60(2), pp. 173-182.

Bourke, B. (2014) 'Positionality: Reflecting on Research Process', *The Qualitative Report*, 19(33), pp. 1-9.

Boyne, G.A. (2004) 'Explaining Public Service Performance: Does Management Matter?', *Public Policy and Administration*, 19(4), pp. 100-117.

Boyne G.A., Gould-Williams, J., Law, J. and Walker, R. (2002) 'Plans, performance information and accountability: the case of best value', *Public Administration,* 80(4), pp. 691-710.

Bovens, M. (2007) 'Analysing and Assessing Accountability: A Conceptual Framework', *European Law Journal*, 13(4), pp. 447-468.

Bovens, M. (2010) 'Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism', *West European Politics,* 33(5), pp. 946-967.

Bovens, M., Goodin, R.E. and Schillemans, T. (ed.) (2014) *The Oxford handbook of public accountability.* Oxford: Oxford University Press.

Bovens, M., Schillemans, T. and Hart, P. (2008) 'Does public accountability work? An assessment tool', *Public Administration*, 86(1), pp. 225-242.

Bracci, E., Maran, L. and Inglis, R. (2017) 'Examining the process of performance measurement system design and implementation in two Italian public service organizations', *Financial Accountability & Management*, 33(4), pp. 406-421.

Bradley, S., Crouchley, R., Millington, J. and Taylor, J. (2000) 'Testing for quasi-market forces in secondary education', *Oxford Bulletin of Economics and Statistics*, 62(3), pp. 357-391.

Brandsma, G.J. (2014) 'Quantitative analysis', in Bovens, M., Goodin, R.E. and Schillemans, T. (ed.) *The Oxford Handbook of Public Accountability.* Oxford: Oxford University Press, pp. 143-158.

Brandsma, G.J. and Schillemans, T. (2013) 'The Accountability Cube: Measuring Accountability', *Journal of Public Administration Research and Theory*, 23(4), pp. 953-975.

Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in psychology*, 3(2), pp. 77-101.

Bregn, K. (2008) 'Management of the new pay systems in the public sector - some implications of insights gained from experiments', *International Review of Administrative Sciences*, 74(1), pp. 79-93.

Brinkerhoff, D.W. and Wetterberg, A. (2016) 'Gauging the Effects of Social Accountability on Services, Governance, and Citizen Empowerment', *Public Administration Review*, 76(2), pp. 274-286.

Broadbent, J., Dietrich, M. and Laughlin, R. (1996) 'The development of principal-agent, contracting and accountability relationships in the public sector: conceptual and cultural problems', *Critical Perspectives on Accounting,* 7(3), pp. 259-284.

Browne, K. (2005) 'Snowball sampling: using social networks to research non-heterosexual women', *International Journal of Social Research Methodology*, 8(1), pp. 47-60.

Burgess, S. and Ratto, M. (2003) 'The Role of Incentives in the Public Sector: Issues and Evidence', *Oxford review of economic policy*, 19(2), pp. 285-300.

Burgess, S., Metcalfe, R. and Sadoff, S. (2021) 'Understanding the response to financial and non-financial incentives in education: Field experimental evidence using high-stakes assessments', *Economics of Education Review*, 85, pp. 1-15.

Butt, H.A. and Palmer, D.R. (1985) *Value for money in the public sector: the decision maker's guide.* Oxford: Blackwell.

Cabinet Office (2006) *Public Bodies: A Guide for Departments*. London.

Cabinet Office (2019) *Tailored Reviews: Guidance on Reviews on Public Bodies*. London.

Carter, N., Klein, R. and Day, P. (1995) *How Organisations Measure Success: The Use of Performance Indicators in Government.* London: Routledge.

Choi, H. and Hong, W. (2017) *An Empirical Study on the Public Institution Performance Evaluation System.* Korea Institute of Public Finance.

Christensen, T. and Lægreid, P. (2015) 'Performance and Accountability-A Theoretical Discussion and an Empirical Assessment', *Public Organization Review*, 15(2), pp. 207-225.

Christensen, T., Lægreid, P. and Røvik, K.A. (2020) *Organization Theory and the Public Sector: Instrument, Culture and Myth*. 2nd edn. New York: Routledge.

Cohen, N. and Arieli, T. (2011) 'Field research in conflict environments: Methodological challenges and snowball sampling', *Journal of Peace Research*, 48(4), pp. 423-435.

Considine, M. (2002) 'The End of the Line? Accountable Governance in the Age of Networks, Partnerships, and Joined-Up Services', *Governance*, 15(1), pp. 21-40.

Courty, P. and Marschke, G. (2007) 'Making Government Accountable: Lessons from a Federal Job Training Program', *Public Administration Review*, 67(5), pp. 914-916.

Creswell, J.W. (2015) 'Revisiting Mixed Methods and Advancing Scientific Practices', in Hesse-Biber, S.N. and Johnson, R.B. (ed.) *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. New York: Oxford University Press, pp. 57-71.

Creswell, J.W. and Creswell, J.D. (2018) *Research design: qualitative, quantitative and mixed methods approaches*. 5th edn. London: SAGE.

Creswell, J.W. and Plano Clark, V. (2007) *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: SAGE.

Creswell, J.W. and Plano Clark, V.L. (2011) *Designing and Conducting Mixed Methods Research*. 2nd edn. Thousand Oaks, CA: SAGE.

Darcy, S. (2007) *Collective Responsibility and Accountability Under International Law*. Ardsley, N.Y.: Transnational Publishers.

Dawson, E.M. and Chatman, E.A. (2001) 'Reference group theory with implications for information studies: a theoretical essay', *Information research*, 6(3).

Dawson, S. and Dargie, C. (2002) 'New Public Management: a discussion with special reference to UK health', in McLaughlin, K., Osborne, S.P. and Ferlie, E. (ed.) *New Public Management: Current trends and future prospects*. London: Routledge, pp. 34-56.

Day, P. and Klein, R. (1987) *Accountabilities: Five public services*. London: Tavistock Publications.

Dewettinck, K. and van Dijk, H. (2013) 'Linking Belgian employee performance management system characteristics with performance management system effectiveness: exploring the mediating role of fairness', *The International Journal of Human Resource Management*, 24(4), pp. 806-825.

Dionne, G. and Harrington, S.E. (1992) *Foundations of insurance economics: Readings in Economics and Finance*. London: Kluwer Academic.

Dixit, A. (2002) 'Incentives and Organizations in the Public Sector: An Interpretative Review', *The Journal of Human Resources*, 37(4), pp. 696-727.

Downward, P. and Mearman, A. (2007) 'Retroduction as mixed-methods triangulation in economic research: reorienting economics into social science', *Cambridge Journal of Economics*, 31, pp. 77-99.

Dubnick, M.J. (2005) 'Accountability and the Promise of Performance: In Search of Mechanisms', *Public Performance and Management Review*, 28(3), pp. 376-417.

Dubnick, M.J. (2014) 'Accountability as a Cultural Keyword', in Bovens, M., Goodin, R.E. and Schillemans, T. (ed.) *The Oxford Handbook of Public Accountability.* Oxford: Oxford University Press, pp. 23–38.

Dunleavy, P., Margetts, H. and Tinkler, S.B.J (2005) 'New Public Management Is Dead-Long Live Digital-Era Governance', *Journal of public administration research and theory*, 16(3), pp. 467-494.

Dunn, M. (2013) *Inside the Capitalist Firm: An Evolutionary Theory of the Principal Agent-Relation.* Potsdam: Potsdam University Press.

Ferreira, A. and Otley, D. (2009) 'The design and use of performance management systems: An extended framework for analysis', *Management Accounting Research*, 20, pp. 263-282.

Flynn, N. (1986) 'Performance measurement in public sector services', *Policy and Politics*, 14(3), pp. 389-404.

Fox, J.A. (2015) 'Social Accountability: What Does the Evidence Really Say?', *World Development*, 72, pp. 346-361.

Franco, L.M., Bennett, S., Kanfer, R. and Stubblebine, P. (2004) 'Determinants and consequences of health worker motivation in hospitals in Jordan and Georgia', *Social Science & Medicine*, 58(2), pp. 343-355.

Frey, B.S. and Oberholzer-Gee, F. (1997) 'An Empirical Analysis of Motivation Crowding-Out', *The American Economic Review*, 87(4), pp. 746-755.

Frey, B.S. and Gallus, J. (2016) 'Awards as non-monetary incentives', *Evidence-based HRM: a Global Forum for Empirical Scholarship*, 4(1), pp. 81-91.

Fryer, K., Antony, J. and ogden, S. (2009) 'Performance management in the public sector', *International Journal of Public Sector Management*, 22(6), pp. 478-498.

Funck, E.K. and Karlsson, T.S. (2019) 'Twenty-five years of studying new public management in public administration: Accomplishments and limitations', *Financial Accountability & Management*, 36(4), pp. 347-375.

Gailmard, S. (2014) 'Accountability and principal-agent theory', in Bovens, M., Goodin, R.E. and Schillemans, T. (ed.) *The Oxford Handbook of Public Accountability.* Oxford: Oxford University Press, pp. 90–105.

Gauld, R. (2007) 'Principal-agent theory and organisational change: Lessons from New Zealand Health Information Management', *Policy Studies*, 28(1), pp. 17-34.

Gayle, V. and Lambert, P. (2018) *What Is Quantitative Longitudinal Data Analysis?* London: Bloomsbury Publishing Plc.

Gerrish, E. (2016) 'The Impact of Performance Management on Performance in Public Organization: A Meta-Analysis', *Public Administration Review*, 76(1), pp. 48-66.

Gianakis, G.A. (2002) 'The promise of public sector performance measurement: anodyne or placebo?', *Public Administration Quarterly*, 26(1/2), pp. 35-64.

Gillham, B. (2008) *Developing a Questionnaire*. 2nd edn. London: Continuum.

Glaser, B. (1978) *Theoretical sensitivity.* Mill Valley, CA: Sociology Press.

Glynn, J.J. and Murphy, M.P. (1996) 'Public management: Failing accountabilities and failing performance review', *The International Journal of Public Sector Management,* 9(5/6), pp. 125-137.

Goh, S.C. (2012) 'Making performance measurement systems more effective in public sector organizations', *Measuring Business Excellence*, 16(1), pp. 31-42.

Goodhart, C.A.E. (1984) *Monetary theory and practice: the UK experience*. London: Macmillan.

Government Offices of Sweden (2019) *Annual report on state-owned enterprises 2018.* Stockholm.

Grace-Martin, K. (2020) 'Five Ways to Analyze Ordinal Variables (Some Better than Others)', *The Analysis Factor*, 14 April. Available at: https://www.theanalysisfactor.com/ways-analyze-ordinal-variables (Accessed: 3 January 2021).

Grandvoinnet, H., Aslam, G. and Raha, S. (2015) *Opening the Black Box: The Contextual Drivers of Social Accountability*. Washington, DC: World Bank Publications.

Gray, A. and Jenkins, W.I. (1985) *Administrative politics in British central government.* Brighton: Harvester Wheatsheaf.

Gray, A. and Jenkins, W.I. (1986) 'Accountable management in British central government: some reflections on the financial management initiative', *Financial Accountability and Management*, 2(3), pp. 171-186.

Greatbanks, R. and Tapp, D. (2007) 'The impact of balanced scorecards in a public sector environment: Empirical evidence from Dunedin City Council, New Zealand', *International Journal of Operations & Production Management*, 27(8), pp. 846-873.

Greene, W.H. (2012) *Econometric Analysis*. 7th edn. Essex: Pearson Education Limited.

Greener, I. (2019) 'Performance management that works? Improving public services by making use of intrinsic motivation', *Social Policy & Administration*, 53(1), pp. 99-112.

Grix, J. (2010) *The Foundations of Research*. 2nd edn. Hampshire: Palgrave Macmillan.

Guest, G., Bunce, A. and Johnson, L. (2006) 'How Many Interviews Are Enough? An Experiment with Data Saturation and Variability', *Field Methods*, 18(1), pp. 59-82.

Ha, S., Oh, Y. and Ra, Y. (2014) *Policy countermeasure for sustainable debt management of public institutions*. Korea Institute of Public Finance.

Halligan, J. (2007) 'Accountability in Australia: control, paradox, and complexity', *Public Administration Quarterly*, 31(3/4), pp. 453-479.

Hammerschmid, G., Van de Walle, S., Andrews, R. and Mostafa, A.M.S. (2019) 'New Public Management Reforms in Europe and their Effects: Findings from a 20-Country Top Executive Survey', *International Review of Administrative Sciences*, 85(3), pp. 399-418.

Hammersley, M. and Traianou, A. (2012) *Ethics in Qualitative Research: Controversies and Contexts*. London: SAGE.

Han, Y. (2020) 'The impact of accountability deficit on agency performance: performance-accountability regime', *Public Management Review,* 22(6), pp. 927-948.

Haque, M.S. (2020) 'Entrepreneurship-driven public management reforms in Southeast Asia: Critical implications for public accountability', *Public Administration and Development*, 40(4), pp. 220-231.

Hauck, K. and Street, A. (2007) 'Do targets matter? A comparison of English and Welsh National Health priorities', *Health Economics*, 16(3), pp. 275-290.

Heckman, J., Heinrich, C. and Smith, J. (1997), 'Assessing the Performance of Performance-Standards in Public Bureaucracies', *The American Economic Review*, 87(2), pp. 389-395.

Heckman, J.J., Smith, J.A. and Taber, C. (1996) 'What Do Bureaucrats Do? The Effects of Performance Standards and Bureaucratic Preferences on Acceptance into the JTPA Program', *National Bureau of Economic Research Working Paper 5535*.

Heinrich, C.J. (2007) 'Evidence-Based Policy and Performance Management: Challenges and Prospects in Two Parallel Movements', *The American Review of Public Administration*, 37(3), pp. 255-277.

Herd, P., DeLeire, T., Harvey, H. and Moynihan, D.P. (2013) 'Shifting Administrative Burden to the State: The Case of Medicaid Take-Up', *Public Administration Review*, 73(s1), pp. S69-S81.

Herzberg, F., Mausner, B. and Snyderman, B.B. (2011) *The Motivation to Work*, New York: John Wiley & Sons.

Heywood, A. (2013) *Politics*. 4th edn. Hampshire: Palgrave Macmillan.

Hickey, S. and King, S. (2016) 'Understanding Social Accountability: Politics, Power and Building New Social Contracts', *The journal of Development Studies*, 52(8), pp. 1225-1240.

Hoe, K. (2015) *Study on public corporations in main countries.* Sejong: Korea Institute of Public Finance.

Hood, C. (1991) 'A public management for all seasons?', *Public Administration*, 69(1), pp. 3-20.

Hood, C. (2006) 'Gaming in Targetworld: The Targets Approach to Managing British Public Services', *Public Administration Review*, 66(4), pp. 515-521.

Hood, C. (2011) *The Blame Game: Spin, bureaucracy, and self-preservation in government*, Princeton and Oxford: Princeton University Press.

Hood, C. (2012) 'Public Management by Numbers as a Performance-Enhancing Drug: Two Hypotheses', *Public Administration Review*, 72(S1), pp. S85-S92.

Hood, C. and Lodge, M. (2004) 'Competency, Bureaucracy, and Public Management Reform: A Comparative Analysis', *Governance*, 17(3), pp. 313-333.

Hyman, D.N. (2014) *Public Finance: a contemporary application of theory to policy.* 11th edn. Stamford, CT: Cengage Learning.

Hyman, H.H. and Singer, E. (1968) *Readings in reference group theory and research*, New York: Free Press.

Iversen, G.R. and Gergen, M. (1997) *Statistics: The Conceptual Approach*. New York: Springer.

Jacobs, R. and Goddard, M. (2007) 'How Do Performance Indicators Add Up? An Examination of Composite Indicators in Public Services', *Public Money & Management*, 27(2), pp. 103-110.

Jacobs, R., Smith, P.C. and Street, A. (2006) *Measuring Efficiency in Health Care: Analytic Techniques and Health Policy*. Cambridge: Cambridge University Press.

Jalava, N., Joensen, J.S. and Pellas, E. (2015) 'Grades and rank: Impacts of non-financial incentives on test performance', *Journal of Economic Behavior & Organization*, 115, pp. 161-196.

Jang, J., Kwack, C., Shin, W. and Oh, C. (2013) *Changing Phases of the Evaluation Management System for Korea Public Institutions: Part II*. Korea Institute of Public Finance.

Jantz, B., Christensen, T. and Lægreid, P. (2015) 'Performance Management and Accountability: The Welfare Administration Reform in Norway and Germany', *International Journal of Public Administration*, 38, pp. 947-959.

Jen, W., Tu, R. and Lu, T. (2011) 'Managing passenger behavioral intention: an integrated framework for service quality, satisfaction, perceived value, and switching barriers', *Transportation*, 38, pp. 321-342.

Johnsen, A. (2005) 'What Does 25 Years of Experience Tell Us About the State of Performance Measurement in Public Policy and Management?', *Public Money and Management*, 25(1), pp. 9-17.

Jones, S. and Bouckaert, G. (2017) 'Performance for Accountability: The Case of Joint Agreements in a Federal System', *Public Performance & Management Review*, 40(4), pp. 765-796.

Jørgensen, T.B., Hansen, H.F., Antonsen, M. and Melander, P. (1998) 'Public Organizations, Multiple Constituencies, and Governance', *Public Administration*, 76(3), pp. 499-518.

Kahn, C.M., Silva, E.C.D. and Ziliak, J.P. (2001) 'Performance-based wages in tax collection: The Brazilian tax collection reform and its effects', *The Economic Journal*, 111, pp. 188-205.

Kane, T.J. and Staiger, D.O. (2002) 'The Promise and Pitfalls of Using Imprecise School Accountability Measures', *The Journal of Economic Perspectives*, 16(4), pp. 91-114.

Kaya, N., Koc, E. and Topcu, D. (2010) 'An exploratory analysis of the influence of human resource management activities and organizational climate on job satisfaction in Turkish banks', *The International Journal of Human Resource Management*, 21(11), pp. 2031-2051.

Kefay, K. and Kero, C.A. (2019) 'The effect of non-financial incentive scheme on employees' motivation (In case of Commercial Bank of Ethiopia in Jimma Town)', *International Journal of Commerce and Finance*, 5(2), pp. 76-86.

Kenk, K. and Haldma, T. (2016) 'Performance management within the public accountability framework in Estonian local governments', *Journal for Public & Nonprofit Services*, 39(1/2), pp. 102-113.

Kihl, Y. (2015) *Transforming Korean Politics: Democracy, Reform, and Culture*. London and New York: Routledge

Kim, S.S. (2003) *South Korea's Democratization*. Cambridge: Cambridge University Press.

Kim, K. (2014) 'Reflection on the time to introduce Public Corporation Management Evaluation System', Lee, W. and Ra, Y. (ed.) (2015) *SOEs Management Evaluation System 30 Years History and Ahead.* Korea Institute of Public Finance. pp. 13-18.

Kingma, M. (2003) 'Economic incentive in community nursing: attraction, rejection or indifference?', *Human Resources for Health*, 1(1), pp. 1-12.

Kornai, J., Maskin, E. and Roland, G. (2003) 'Understanding the Soft Budget Constraint', *Journal of Economic Literature*, 41(4), pp. 1095-1136.

Lægreid, P. (2014) 'Accountability and New Public Management', in Bovens, M., Goodin, R.E. and Schillemans, T. (ed.) *The Oxford Handbook of Public Accountability.* Oxford: Oxford University Press, pp. 324-338.

Lam, J.T.M. (2009) 'Political Accountability in Hong Kong: Myth or Reality?', *Australian Journal of Public Administration*, 68(S1), pp. S73-S83.

Lapuente, V. and Van de Walle, S. (2020) 'The effects of new public management on the quality of public services', *Governance*, 33(3), pp. 461-476.

Lee, H. (2021) 'Revisiting crowding-out effect of performance management: Its impact on extrinsic and intrinsic motivation', *Transylvanian Review of Administrative Sciences*, 63, pp. 90-109.

Lee, J., Lee, Y., Kim, S. and Song, J. (2022) 'Work values: a latent class analysis of Korean employees', *Higher Education, Skills and Work-based Learning*, 12(5), pp. 834-848.

Lee, S. (2014) *The Impact of the Korean Preliminary Feasibility Study on Budgetary Decisions*. Doctoral Thesis, University of Birmingham.

Lee, W. and Ra, Y. (2015) *SOEs Management Evaluation System 30 Years History and Ahead.* Korea Institute of Public Finance.

Levacic, R. (2009) 'Teacher Incentives and Performance: An Application of Principal-Agent Theory', *Oxford Development Studies*, 37(1), pp. 33-46.

Lim, D., Choi, M. and Song, J. (2012) 'Work-family enrichment in Korea: construct validation and status', *Leadership & Organization Development Journal*, 33(3), pp. 282-299.

Lim, Y. and Roh, J. (2020) 'Political institutions in Korea', in Moon, C. and Moon, M. (ed.) *Routledge Handbook of Korean Politics and Public Administration*. London and New York: Routledge, pp. 21-40.

Lindberg, S.I. (2013) 'Mapping accountability: core concept and subtypes', *International Review of Administrative Sciences*, 79(2), pp. 202-226.

Little, R.J.A. and Rubin, D.B. (2002) *Statistical analysis with missing data*, 2nd edn. New Jersey: Wiley.

Lockwood, B. and Porcelli, F. (2013) 'Incentive Schemes for Local Government: Theory and Evidence from Comprehensive Performance Assessment in England', *American Economic Journal: Economic Policy*, 5(3), pp. 254-286.

Lu, Y., Yang, K. and Thomas, M.B. (2021) 'Designing Performance Systems in Anticipation of Unintended Consequences: Experiences and Lessons From the Compstat-Based Performance Regime in NYPD', *Administration & Society*, 53(6), pp. 907-936.

Lyons, S.T., Duxbury, L.E. and Higgins, C.A. (2006) 'A Comparison of the Values and Commitment of Private Sector, Public Sector and Parapublic Sector Employees', *Public Administration Review*, 66(4), pp. 605-618.

Mankiw, N.G. and Taylor, M.P. (2017) *Economics*, 4th edn. Hampshire: Cengage.

Mankiw, N.G. and Taylor, M.P. (2020) *Economics*. 5th edn. Hampshire: Cengage.

Mathauer, I. and Imhoff, I. (2006) 'Health worker motivation in Africa: the role of non-financial incentives and human resource management tools', *Human Resources for Health*, 4:24, pp. 1-17.

May, T. (2011) *Social research: issues, methods and process.* 4th edn. Maidenhead: Open University Press.

Mayumana, I., Borghi, J., Anselmi, L., Mamdani, M. and Lange, S. (2017) 'Effects of Payment for Performance on accountability mechanisms: Evidence from Pwani, Tanzania', *Social Science & Medicine*, 179, pp. 61-73.

Meyer, R.H. (1997) 'Value-Added Indicators of School Performance: A Primer', *Economics of Education Review*, 16(3), pp. 283-301.

Miles, M.B., Huberman, A.M. and Saldaña, J. (2013) *Qualitative data analysis.* 3rd edn. Thousand Oaks, CA: SAGE.

Miles, M.B., Huberman, A.M. and Saldaña, J. (2020) *Qualitative data analysis: A Methods Sourcebook*. 4th edn. Thousand Oaks, CA: SAGE.

Mizrahi, S. and Minchuk, Y. (2019a) 'The relevance and reliability of performance information for accountability: A survey experiment exploring citizens' views', *Public Administration*, 98(1), pp. 140-158.

Mizrahi, S. and Minchuk, Y. (2019b) 'Performance management, gaming and monitoring in Israeli organisations', *Israel Affairs*, 25(3), pp. 452-466.

Modell, S. (2009) 'In defence of triangulation: A critical realist approach to mixed methods research in management accounting', *Management Accounting Research*, 20, pp. 208-221.

MOEF (2010) News Release *'Dramatic reform of Public Institution Management Evaluation System'*. Korean Ministry of Economy and Finance.

MOEF (2011) *The Public Institutions Management System in Korea*. Korean Ministry of Economy and Finance.

MOEF (2013) *Measures to normalise the Korean public institutions*. Korean Ministry of Economy and Finance.

MOEF (2014) News Release *'Announcement of public institutions management performance evaluation result on FY 2013'*. Korean Ministry of Economy and Finance.

MOEF (2016) News Release *'Confirmation of evaluation result on public institution management performance in 2015'*. Korean Ministry of Economy and Finance.

MOEF (2018) *Guidance on management evaluation of public institutions*. Korean Ministry of Economy and Finance.

MOEF (2019a) *Guidance on management evaluation of public institutions*. Korean Ministry of Economy and Finance.

MOEF (2019b) News Release *'Announcement of public institutions management performance evaluation result'*. Korean Ministry of Economy and Finance.

MOEF (2020) News Release *'Announcement of public institutions management performance evaluation result'.* Korean Ministry of Economy and Finance.

MOEF (2021) News Release *'Announcement of public institutions management performance evaluation result'.* Korean Ministry of Economy and Finance.

Moon, C. (2018) *Analysis of performance of evaluation indicators for business administration in Public Institution Management Performance in 2017,* Korea Institute of Tax and Public Finance.

Moon, C. and Moon, M. (2020) 'Introduction', in Moon, C. and Moon, M. (ed.) *Routledge Handbook of Korean Politics and Public Administration*. London and New York: Routledge, pp. 1-17.

MPB (2007) News Release *'Confirmation of reformation plan about Public Institution Management Performance Evaluation System'.* Korean Ministry of Planning and Budgeting.

Mulgan, R. (2003) *Holding Power to Account: Accountability in Modern Democracies*. Hampshire and New York: Palgrave Macmillan.

Mulgan, R. (2006) 'Government Accountability For Outsourced Services', *Australian Journal of Public Administration*, 65(2), pp. 48-58.

Mulgan, R. (2014) 'Accountability deficits', in Bovens, M., Goodin, R.E. and Schillemans, T. (ed.) *The Oxford Handbook of Public Accountability.* Oxford: Oxford University Press, pp. 545–559.

Murphy, P., Ferry, L., Glennon, R and Greenhalgh, K. (2019) *'Public Service Accountability: Rekindling a Debate',* Cham: Palgrave.

Nayyab, H. H., Hamid, M., Naseer, F. and Iqbal, M. (2011) 'The impact of HRM practices on the organizational performance: The study of banking sector in Okara, Punjab (Pakistan)', *Interdisciplinary Journal of Contemporary Research in Business*, 3, pp. 661-672.

O'Connell, L. (2005) 'Program Accountability as an Emergent Property: The Role of Stakeholders in a Program's Field', *Public Administration Review*, 65(1), pp. 85-93.

OECD (2005) *OECD Guidelines on Corporate Governance of State-Owned Enterprises*. Paris: OECD Publishing.

OECD (2009) *Measuring government Activity*. Paris: OECD Publishing.

OECD (2011) *Government at a glance 2011*. Paris: OECD Publishing.

OECD (2015) *OECD Guidelines on Corporate Governance of State-Owned Enterprises*. 2015 edn. Paris: OECD Publishing.

OECD (2018) *Ownership and governance of State-Owned Enterprises: A Compendium of National Practices*. Paris: OECD Publishing.

OECD (2020) *Implementing the OECD Guidelines on Corporate Governance of State-Owned Enterprises: Review of Recent Developments.* Paris: OECD Publishing.

Office for National Statistics (2019) *Public sector employment, UK: September 2019.*

O'Leary, Z. (2017) *The essential guide to doing your research project.* 3rd edn. London: SAGE.

Oña, J.D. and Oña, R.D. (2015) 'Quality of Service in Public Transport Based on Customer Satisfaction Surveys: A Review and Assessment of Methodological Approaches', *Transportation Science*, 49(3), pp. 605-622.

Osborne, D. and Gaebler, T. (1992) *Reinventing government: how the entrepreneurial spirit is transforming the public sector.* Wokingham: Addison-Wesley.

Osborne, S.P. and McLaughlin, K. (2002) 'The New Public Management in context' in McLaughlin, K., Osborne, S.P. and Ferlie, E. (ed.) *New Public Management: Current trends and future prospects*. London: Routledge, pp. 7-14.

Ospina, S., Grau, N.C. and Zaltsman, A. (2004) 'Performance evaluation, public management improvement and democratic accountability: Some lessons from Latin America', *Public Management Review,* 6(2), pp. 229-251.

Pampino, R.N., MacDonald, J.E., Mullin, J.E. and Wilder, D.A. (2004) 'Weekly Feedback vs. Daily Feedback: An Application in Retail', *Journal of Organizational Behaviour Management*, 2(3), pp. 21-43.

Panagiotou, G. (2007) 'Reference theory: strategic groups and competitive benchmarking', *Management decision*, 45(10), pp. 1595-1621.

Park, H. (2013) 'Management method of public institution in major foreign countries: focusing on performance management system', *The journal of tax and finance brief*, 13(5).

Park, J., Kim, J. and Chung, S. (2020) *Evaluation and management methods of public institutions in major foreign countries*. Seoul: National Assembly Budget Office.

Park, J., Kim, J. and Kim, C. (2019) *Is the management evaluation system of state-owned enterprises in the Republic of Korea a good tool for better performance?* Tokyo: Asian Development Bank Institute.

Parkhurst, J. (2017) *The politics of evidence: from evidence-based policy to the good governance of evidence*. Oxon: Routledge.

Perloff, J.M. (2018) *Microeconomics*. 8th edn. Harlow: Pearson.

Perri 6 and Bellamy, C. (2012) *Principles of Methodology: Research Design in Social Science*. London: SAGE.

Perry, J.L., Engbers, T.A. and Jun, S. (2009) 'Back to the Future? Performance-Related Pay, Empirical Research, and the Perils of Persistence', *Public Administration Review*, 69(1), pp. 39-51.

Peters, B.G. (2014) 'Accountability in public administration', in Bovens, M., Goodin, R.E. and Schillemans, T. (ed.) *The Oxford Handbook of Public Accountability.* Oxford: Oxford University Press, pp. 211-225.

Pidd, M. (2008) 'A critical assessment of performance measurement for policy making', in Hartley, J., Donaldson, C., Skelcher, C. and Wallace, M. (ed.) *Managing to improve public services.* Cambridge: Cambridge University Press, pp. 65-87.

Pindyck, R. and Rubinfeld, D. (2018) *Microeconomics.* 9th edn. Essex: Pearson.

Pollitt, C. (2003) *The Essential Public Manager.* Maidenhead: McGraw-Hill Education.

Pollitt, C. (2011) 'Performance blight and the tyranny of light? Accountability in Advanced Performance Measurement Regimes' in Dubnick, M.J and Frederickson, H.G. (ed.) *Accountable governance: problems and promises.* New York: M.E. Sharpe, pp. 81-97.

Pollitt, C. (2013) 'The logics of performance management', *Evaluation*, 19(4), pp. 346-363.

Pollitt, C. and Bouckaert, G. (2017) *Public management reform: a comparative analysis-into the age of austerity*, 4th edn. Oxford: Oxford University Press.

Pollitt, C and Dan, S. (2013) 'Searching for impacts in performance-oriented management reform: A Review of the European Literature', *Public Performance & Management Review*, 37(1), pp. 7-32.

Pollitt, C., Talbot, C., Caulfield, J. and Smullen, A. (2004) *Agencies: How Governments do Things Through Semi-Autonomous Organizations*. Hampshire and New York: Palgrave Macmillan.

Power. M. (1997) *The Audit Society: Rituals of Verification.* Oxford: Oxford university press.

Propper, C. and Wilson, D. (2003) The Use and Usefulness of Performance Measures in the Public Sector, *Oxford review of economic policy*, 19(2). pp. 250-267.

Pullin, L. and Haidar, A. (2003) 'Performance contract management in regional local government – Victoria', *Asia Pacific Journal of Human Resources*, 41(3), pp. 279-297.

Punch, K.F. (2014) *Introduction to Social research: Quantitative & qualitative approaches.* 3rd edn. London: SAGE.

Rajavuori, M. (2018) 'Governing the Good State Shareholder: The Case of the OECD Guidelines on Corporate Governance of State-Owned Enterprises', *European Business Law Review*, 29(1), pp. 103-142.

Rana, T. and Hoque, Z. (2020) 'Institutionalising multiple accountability logics in public services: Insights from Australia', *The British Accounting Review*, 52(4), pp. 1-14.

Raziq, A. and Maulabakhsh, R. (2015) 'Impact of Working Environment on Job Satisfaction', *Procedia Economics and Finance*, 23, pp. 717-725.

Ringen, S., Kwon, H., Yi, I., Kim, T. and Lee, J. (2011) *The Korean State and Social Policy: How South Korea Lifted Itself from Poverty and Dictatorship to Affluence and Democracy*. Oxford: Oxford University Press.

Risjord, M. (2014) *Philosophy of Social Science: A Contemporary Introduction*. Oxfordshire: Routledge.

Rock, E. (2020) *Measuring Accountability in Public Governance Regimes*. Cambridge: Cambridge University Press.

Roh, J. (2018) 'Improving the government performance management system in South Korea: Focusing on central government agencies', *Asian Education and Development Studies*, 7(3), pp. 266-278.

Rossi, P.H., Lipsey, M.W. and Freeman, H.E. (2004) *Evaluation: A Systematic Approach.* 7th edn. Thousand Oaks, California: Sage Publications.

Ryan, R.M. and Deci, E.L. (2017) *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. New York and London: The Guilford Press.

Ryu, N. (2014) *Recognition survey of trust in public sector.* Seoul: Korea Institute of Public Administration.

Saldaña, J. (2013) *The Coding Manual for Qualitative Researchers*. 2nd edn. Los Angeles: SAGE.

Samaratunge, R., Alam, Q. and Teicher, J. (2008) 'The New Public Management reforms in Asia: a comparison of South and Southeast Asian countries', *International Review of Administrative Sciences*, 74(1), pp. 25-46.

Sansone, C. and Harackiewicz, J.M. (ed.) (2000) *Intrinsic and Extrinsic Motivation: The Search for Optimal Motivation and Performance*. San Diego and London: Academic Press.

Sarikaya, A. and Gleicher, M. (2018) 'Scatterplots: Tasks, Data and Designs', *IEEE transactions on visualization and computer graphics*, 24(1), pp. 402-412.

Saunders, M.N.K. and Townsend, K. (2016) 'Reporting and justifying the number of participants in organisation and workplace research', *British Journal of Management,* 27, pp. 837-852.

Saunders, M.N.K., Thornhill, A. and Lewis, P. (2019) *Research Methods for Business students.* 8th edn. Harlow: Pearson Education Limited.

Saunders, M.N.K and Townsend, K. (2018) 'Choosing Participants', in Cassell, C., Cunliffe, A.L. and Grandy, G. (ed.) *The Sage Handbook of Qualitative Business and Management Research Method: History and Traditions*. London: SAGE, pp. 480-494.

Sayer, A. (1992) *Method in Social Science: A realist approach*. 2nd edn. London: Routledge.

Sayer, A. (2000) *Realism and Social Science.* London: SAGE.

Schillemans, T. (2011) 'Does Horizontal Accountability Work? Evaluating Potential Remedies for the Accountability Deficit of Agencies', *Administration & society*, 43(4), pp. 387-416.

Schillemans, T. (2016) 'Calibrating public sector accountability: Translating experimental findings to public sector accountability', *Public Management Review*, 18(9), pp. 1400-1420.

Schlechter, A., Thompson, N.C. and Bussin, M. (2015) 'Attractiveness of non-financial rewards for prospective knowledge workers: An experimental investigation', *Employee Relations*, 37(3), pp. 274-295.

Schleicher, D.J., Baumann, H.M., Sullivan, D.W., Levy, P.E., Hargrove, D.C. and Barros-Rivera, B.A. (2018) 'Putting the System Into Performance Management Systems: A Review and Agenda for Performance Management Research', *Journal of Management*, 44(6), pp. 2209-2245.

Schultze, C.L. (1968) *The politics and economics of public spending.* Washington: The Brookings Institution.

Seawright, J. (2016) *Multi-Method Social Science: Combining Qualitative and Quantitative Tools*. Cambridge: Cambridge University Press.

Shin, S. (2013) *Dysfunctional consequences of the Korean performance budgeting system and their policy implications*. Doctoral Thesis, University of Birmingham.

Simon, H.A., Simthburg, D.W., and Thompson, V.A. (1950) *Public administration.* New Brunswick, N.J.: Knopf.

Singh, G. and Slack, N.J. (2022) 'New Public Management and Customer Perceptions of Service Quality – A Mixed-Methods Study', *International Journal of Public Administration*, 45(3), pp. 242-256.

Skelcher, C. (2008) 'Does governance perform? Concepts, evidence, causalities and research strategies', in Hartley, J., Donaldson, C., Skelcher, C. and Wallace, M. (ed.) *Managing to improve public services*. Cambridge: Cambridge University Press, pp. 27-45.

Sloman, J., Garratt, D. and Guest, J. (2020) *Economics*, 10th edn. Harlow: Pearson.

Smith, E. (2008) *Using Secondary Data in Educational and Social Research*. Berkshire: Open University Press.

Song, H. (2015) 'The reflection and outlook of management evaluation of government affiliated institutions', Lee, W. and Ra, Y. (ed.) (2015) *SOEs Management Evaluation System 30 Years History and Ahead.* Korea Institute of Public Finance. pp. 143-151.

Song, H., Yu, S. and Sun, T. (2020) 'Reducing the quality risk of elderly care services in government procurement from market-oriented private providers through *ex-ante* policy design: lessons from the principal-agent theory analysis', *BMC Health Services Research*, 20(1), pp. 1-11.

Speklé, R.F. and Verbeeten, F.H.M. (2014) 'The use of performance measurement systems in the public sector: Effects on performance', *Management Accounting Research*, 25, pp. 131-146.

Stewart, J.D. (1984) 'The role of information in public accountability', *Issues in public sector accounting.* Hopwood, A.G. and Tomkins, C. (ed.) Oxford: Phillip Alan. pp. 13-34.

Stone, B. (1995) 'Administrative accountability in the "Westminster Democracies": Towards a new conceptual framework', *Governance,* 8(4), pp. 505-526.

Stringer, C. (2007) 'Empirical performance management research: observations from AOS and MAR', *Qualitative Research in Accounting & Management*, 4(2), pp. 92-114.

Stringer, C. and Shantapriyan, P. (2012) *Setting Performance Targets*. New York: Business Expert Press.

Sung, S. (2014) 'What Determines the Tenure of Cabinet Members? A Comparative Study of Korea and the United States, 1948-2013', *The Korean Journal of Policy Studies*, 29(3), pp. 31-51.

Talbot, C. (2010) *Theories of performance: organizational and service improvement in the public domain*. Oxford: Oxford University Press.

Taylor, J. (2021) 'Public Officials' Gaming of Performance Measures and Targets: The Nexus between Motivation and Opportunity', *Public Performance & Management Review*, 44(2), pp. 272-293.

Taylor, S.J., Bogdan, R. and DeVault, M.L. (2016) *Introduction to Qualitative Research Methods: A Guidebook and Resource*. 4th edn. New Jersey: Wiley.

Terman, J.N. and Yang, K. (2016) 'Reconsidering Gaming in an Accountability Relationship: The Case of Minority Purchasing in Florida', *Public Performance & Management Review*, 40(2), pp. 281-309.

Tesch, R. (1990) *Qualitative research: analysis types and software tools*. London: Falmer.

Townley, B. (2008) 'Performance measurement systems and the criminal justice system: rationales and rationalities', in Hartley, J., Donaldson, C., Skelcher, C. and Wallace, M. (ed.) *Managing to improve public services*. Cambridge: Cambridge University Press, pp.134-152.

Treasury (2004) *The UK Government's Public Service Agreement Framework*. London: Stationary Office.

UK Government Investments (2018) *UK Government Investments Framework Document April 2018*.

Van De Walle, S. and Cornelissen, F. (2014) 'Performance reporting', in Bovens, M., Goodin, R.E. and Schillemans, T. (ed.) *The Oxford Handbook of Public Accountability.* Oxford: Oxford University Press, pp. 441-455.

Van der Geer, E., van Tuijl, H.F.J.M. and Rutte, C.G. (2009) 'Performance management in healthcare: Performance indicator development, task uncertainty, and types of performance indicators', *Social Science & Medicine*, 69, pp. 1523-1530.

Van Dooren, W., Bouckaert, G. and Halligan, J. (2015) *Performance management in the public sector.* 2nd edn. London: Routledge.

Van Osch, D., De Ruiter, R. and Yesilkagit, K. (2023) 'An accountability deficit? Holding transgovernmental networks to account', *Journal of European Public Policy*, 30(2), pp. 315-333.

Van Thiel, S. and Leeuw, F.L. (2002) 'The Performance Paradox in the Public Sector', *Public Performance & Management Review*, 25(3), pp. 267-281.

Verbeeten, F.H.M. (2008) 'Performance management practices in public sector organizations: Impact on performance', *Accounting, Auditing & Accountability Journal*, 21(3), pp. 427-454.

Vriens, D., Vosselman, E. and Groβ, C. (2018) 'Public Professional Accountability: A Conditional Approach', *Journal of Business Ethics*, 153(4), pp. 1179-1196.

Walker, R.M., Andrews, R., George, B. and Tu, X. (2023) 'Organizational size and public service performance: a meta-analysis and an agenda for future research', *Asia Pacific Journal of Public Administration*, pp. 1-34.

Watt, P.A. (1998) 'White-collar services in local government: competition and trust', in Coulson, A. (ed.) *Trust and contracts: Relationships in local government, health and public services.* Bristol: The Policy Press, pp. 183-205.

Weibel, A., Rost, K. and Osterloh, M. (2010) 'Pay for Performance in the Public Sector-Benefits and (Hidden) Costs', *Journal of Public Administration Research and Theory*, 20(2), pp. 387-412.

Willems, T. and Van Dooren, W. (2012) 'Coming to Terms with Accountability', *Public Management Review*, 14(7). pp. 1011-1036.

Williams, M. and May, T. (1996) *Introduction to the philosophy of social research*. London: Routledge.

Willis, E., Carryer, J., Harvey, C., Pearson, M. and Henderson, J. (2017) 'Austerity, new public management and missed nursing care in Australia and New Zealand, *Journal of Advanced Nursing*, 73(12), pp. 3102-3110.

Woo, Y. (2017) 'Exploratory Study of Evaluability assessment on Evaluation Indicators for Main programs of Public Organization Group 1', *Journal of Social Science,* 19, pp. 125-149.

Woodhouse, D. (1994) *Ministers and Parliament: Accountability in Theory and Practice*. Oxford: Oxford University Press.

Wooldridge, J.M. (2020), *Introductory Econometrics: A Modern Approach*, 7th edn. Boston: Cengage.

Wouters, M. and Wilderom, C. (2008) 'Developing performance-measurement systems as enabling formalization: A longitudinal field study of a logistics department', *Accounting, Organizations and Society*, 33(4/5), pp. 488-516.

World Bank (2012) *Strengthening Governance, Tackling Corruption: The World Bank's Updated Strategy and Implementation Plan*. Washington, DC. Available at: http://siteresources.worldbank.org/PUBLIC_SECTORANDGOVERNANCE / Resources/285741-13268161182754/GAC StrategyImplementationPlan.pdf. (Accessed 4 April 2023).

Yang, K. (2014) 'Qualitative analysis', in Bovens, M., Goodin, R.E. and Schillemans, T. (ed.) *The Oxford Handbook of Public Accountability.* Oxford: Oxford University Press, pp. 159-176.

Yang, S. and Torneo, A.R. (2015) 'Government performance management and evaluation in South Korea: History and current practices', *Public Performance & Management Review*, 39(2). pp. 279-296.

Yin, R.K. (2014) *Case Study Research: Design and Methods.* 5th edn. London: SAGE.

Yin, R.K. (2018) *Case Study Research and Applications: Design and Methods*. 6th edn. Thousand Oaks, CA: SAGE.

Yu, S. (2010) 'Analysis of reliability and feasibility of management evaluation of public institutions: Focusing on FY 2008 evaluation result', *Korean Journal of Policy Analysis and Evaluation,* 20 (4). pp. 171-202.

Zachariadis, M., Scott, S. and Barrett, M. (2013) 'Methodological Implications of Critical Realism for Mixed-Methods Research', *MIS Quarterly*, 37(3), pp. 855-879.

Zwalf, S. (2022) 'Managing goal conflict. The case of agency theory in the policy settings for public-private partnerships; A perspective on citizen and government interests', *Annals of Public and Cooperative Economics*, 93(4), pp. 913-930.

Appendix: Subjects of the management evaluation on performance in 2018

Category		Name of public institution
Public corporation (35)	Public corporation I (10)	Korea Gas Corporation; Korea National Oil Corporation; Korea Electric Power Corporation; Korea District Heating Corporation; Incheon International Airport Corporation; Korea Airports Corporation; Korea Expressway Corporation; Korea Water Resources Corporation; Korea Land & Housing Corporation; Korea Railroad Corporation
	Public corporation II (25)	Korea Minting; Korea Resources Corporation; Korea South-East Power Company; Korea Midland Power Company; Korea Southern Power Company; Korea Western Power Company; Korea East-West Power Company; Korea Hydro & Nuclear Power Company; Gangwon Land Company; Security Printing & ID Card Operating Corporation; Grand Korea Leisure Company; Korea Racing Authority; Busan Port Authority; Korea Gas Technology Corporation; Korea Coal Corporation; Korea Electric Power Corporation Engineering & Construction Company; Korea Electric Power KDN Company; Jeju Free International City Development Center; Korea Housing & Urban Guarantee Corporation; Korea Appraisal Board; Yeosu Gwangyang Port Authority; Ulsan Port Authority; Incheon Port Authority; Korea Marine Environment Management Corporation; Korea Broadcast Advertising Corporation

	Fund-management-type (16)	Teacher's pension; Korea Sports Promotion Foundation; Korean Film Council; Arts Council Korea; Korea Press Foundation; Korea Trade Insurance Corporation; Korea Radioactive Waste Agency; National Pension Service; Korea Workers' Compensation & Welfare Service; Korea Technology Finance Corporation; Korea SMEs and startups Agency; Korea Credit Guarantee Fund; Korea Deposit Insurance Corporation; Korea Asset Management Corporation; Korea Housing Finance Corporation; Government Employees Pension Service
Quasi-governmental Institution (93)	Commissioned-service-type (34)	Korea Student Aid Foundation; National IT industry Promotion Agency; National Research Foundation of Korea; Korea Internet & Security Agency; National Information Society Agency; Korea International Cooperation Agency; Korea Tourism Organization; Korea Institute for Animal Products Quality Evaluation; Korea Agro-Fisheries & Food Trade Corporation; Korea Trade-Investment Promotion Agency; Korea Gas Safety Corporation; Korea Institute for Advancement of Technology; Korea Evaluation Institute of Industrial Technology; Korea Industrial Complex Corporation; Korea Petroleum Quality & Distribution Authority; Korea Energy Agency; Korea Electrical Safety Corporation; Korea Power Exchange; National Health Insurance Service; Social Security Information Service; Korea National Park Service; Korea Environment Corporation; Korea Occupational Safety & Health Agency; Human Resources Development Service of Korea; Korea Employment Agency for the Disabled; Korea Transportation Safety Authority; Korea Land and Geospatial Informatix Corporation; Korea Ship Safety Technology Authority; Korea Elevator Safety Agency; Small

		Enterprise & Market Service; Korea Consumer Agency; Korea Institute of Nuclear Safety; Korea Veterans Health Service; Korea Road Traffic Authority; Korea Fire Institute; Korea Intellectual property Strategy Agency
	Strong and small Institutions (43)	Korea Fisheries Resources Agency; Korea Rural Community Corporation; Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry; Korea Communications Agency; Korea Public Finance Information Service; Korea Education and Research Information Service; Postal Savings & Insurance Development Institute; Korea Postal Service Agency; Korea Postal Logistics Agency; Korea Foundation for the Advancement of Science & Creativity; INNOPOLIS Foundation; Korea International Broadcasting Foundation; Korea Creative Content Agency; Asia Culture Institute; Korea Agency of education, Promotion, and Information Service in Food, Agriculture, Forestry and Fisheries; Mine Reclamation Corporation; Korea Institute of Design Promotion; Korea Institute of Ceramic Engineering and Technology; Korea Institute of Energy Technology Evaluation and Planning ; Health Insurance Review & Assessment Service; Korea Labor Force Development Institute for the aged; Korea Human Resource Development Institute for Health & Welfare; Korea Health Industry Development Institute; National Institute of Ecology; Korea Environmental Industry & Technology Institute; Korea Employment Information Service; Korea Youth Counselling & Welfare Institute; Korea Youth Work Agency; Korea Agency for infrastructure Technology Advancement; Korea Infrastructure Safety Corporation; Korea Rail Network Authority; Korea Institute of Marine Science & Technology Promotion; Korea Institute

		of Maritime and Fisheries Technology; Korea Technology & Information Promotion Agency for SMEs; Community Media Foundation; The Independence Hall of Korea; Korea Agency of HACCP Accreditation and Services; Korea Forest Welfare Institute; Korea Forestry Promotion Institute; Foundation of Agri. Tech. Commercialization & Transfer; Korea Meteorological Institute