Research Paper

An Empirical Study on the Public Institution Performance Evaluation System

December 2017

Han-soo Choi
Woo-hyeong Hong

KIPF KOREA INSTITUTE OF PUBLIC FINANCE

Korea Institute of Public Finance
336, Sicheong-daero, Sejong-si, Korea
Tel: 82-44-414-2114 Fax: 82-44-414-2179
URL: www.kipf.re.kr

An Empirical Study on the Public Institution Performance Evaluation System

December 2017

Han-soo Choi · Woo-hyeong Hong

KIPF KOREA INSTITUTE OF PUBLIC FINANCE

Contents

List of Tables

List of Figures

I

Introduction

1 Background

The public institution management performance evaluation system (hereinafter the PIMPES) is an institutional measure to objectively assess the management performance of public institutions and tie the outcomes to the tenure of the head of a given public institution and to performance-based pay for employees (Ministry of Strategy and Finance 2016A). The objective of the system is to improve the management efficiency of public institutions and promote autonomous and responsible management practices among those institutions (Ministry of Strategy and Finance 2016A).

The origin of the PIMPES can be traced back to the performance evaluation system on government-invested institutions pursuant to the「Framework Act on the Management of Government-Invested Institutions」enacted in 1983 (Won-hee Lee and Young-jae La, 2015). However, the evaluation system of that time was different from the PIMPES in many respects. Back then, there were only a few dozen public institutions that were subject to performance evaluation, and the number of evaluation items was limited as well. In this respect, it is more reasonable to say that the current performance evaluation system originated from the「Act on the Management of Public Institutions」(hereinafter the AMPI) introduced in 2007, considering the methods of analysis and evaluation frameworks currently employed by the PIMPES.

There are conflicting views on the effectiveness of the PIMPES and the

need to maintain it. Some point out that the current performance evaluation system is a very advanced one introducing contemporary management techniques to public institutions (Ji-in Jang, et al., 2013) and has been a benchmark for other developing countries. Others argue that the system is causing inefficiency in the operation of public institutions and actually functions as a means of bureaucratic control by applying uniform evaluation criteria with no consideration of differences among institutions (Young-jae La and Tae-beom Yoon 2013).

The following anecdote about the Incheon International Airport Corporation illustrates a delicate situation faced by the PIMPES. In 2016, a baggage crisis delayed the departure of about 160 airplanes at Incheon International Airport. To make things worse, illegal entries ensued. This gave rise to a series of criticism, from both domestic and international sources, about the public corporation's security and safety management issues. In response to this, Il-young Jeong, president of Incheon International Airport Corporation, said in a press interview that he would now spend enough money on security- and safety-related matters even if this would mean losing points on management performance review.[1] This statement might sound strange to the general public. Security and safety management is one of the core business areas of the Incheon International Airport Corporation as an international airport operator. Should investment in its major business focus be disadvantageous to itself according to the criteria for the evaluation of public institutions, it is quite natural to question the raison d'être of the PIMPES.

The problem confronted by the Incheon International Airport Corporation seems to demonstrate that the rigidness of performance evaluation criteria fails to take into account a given public institution's responsiveness to its urgent task required by changed business environments. This may also be interpreted as an example of the fact that the PIMPES is actually serving as a shackle that hampers the autonomy and responsiveness of public institutions.

This study is part of an effort to find answers to these questions. In this study, we aim to understand the PIMPES based on the concepts of information

1) "I would rather give up scores on the public corporation performance evaluation to spend enough money on security for Incheon International Airport." 『*Chosunilbo*』 February 3, 2016. Accessed October, 2016. http://news.chosun.com/site/data/html_dir/2016/02/03/2016020300349.html

and incentives—two basic analytical frameworks of economics—and analyze its functions and dynamics accordingly. How can we justify the raison d'être of the PIMPES from an economic point of view? From the perspective of an incentive contract, what kind of institutional design would lead the PIMPES to produce socially optimal outcomes? Would tying performance-based pay for executives and employees in public institutions to the outcomes of performance evaluation bring enhanced efficiency and publicness of public institutions? In designing the PIMPES, how should we factor in the double mandate of public institutions that have to pursue efficiency and publicness at the same time?

The primary purpose of this study consists in finding an adequate theoretical framework of analysis that will help reach a correct answer by re-interpreting these questions with a new perspective and deriving proper institutional solutions accordingly. The detailed methodology and analytical framework of this study are as follows. We will consider the PIMPES as an incentive contract concluded between the principal and the agent in the sense that remuneration of public institutions' executives and employees is influenced by the outcomes of performance evaluation. In particular, this study will draw on the theory of optimal contract between the principal and the agent under the multitask conditions—first presented by Holmström and Milgrom (1991) in the paper, "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design published in the Journal of Law, Economics, & Organization—as our key theoretical framework that can best analyze the characteristics of the PIMPES. We will also borrow general insights from contract theory and deduce several normative propositions that will help us ultimately assess how optimally the current PIMPES has been designed. In addition, we will look at how efficiently the system has been designed and operated in light of these normative propositions by using actual data collected from the current PIMPES.

The detailed structure of this study is as follows. First, we will examine whether the present PIMPES is functioning properly as a measure of control over moral hazard within public institutions. Why moral hazard? It is because the PIMPES is structurally more similar to the terms and conditions commonly used in the insurance market (e.g. fire insurance) than stock option contracts signed with CEOs, as commonly observed in the labor market, in the following two aspects.

Above all, the performance of a private corporation is usually measured by its outcome—i.e. the value of their stocks. As for public institutions, performance is evaluated not only by outcome, but also by *input* and *process*. As a typical incentive contract, a stock option contract is rarely designed this way.

Moreover, the performance of a private corporation is evaluated by stock price as well. Stock price fluctuates in real time according to whether the management is adequately responding to constant contingencies and uncertainties arising in the course of running their company. In this respect, stock price is a highly flexible indicator for performance evaluation. On the contrary, the performance of public institutions is measured by a handful of indicators set by the government. These indicators are notified in advance to public institutions, and they are, in principle, not alterated. For this reason, the outcomes of performance evaluation on public institutions are, in essence, lack information that can show how they respond properly to contingencies and uncertainties occurring after evaluation indicators have been announced. Rather, the government prefers these indicators to be rigid for the purpose of ensuring fairness and consistency in the evaluation process.

In other words, the current PIMPES has been structured to evaluate particular behaviors of a given public institution that correspond to evaluation indicators notified in advance. Again, this type of contract structure is often observed in insurance contracts, but not in private stock option contracts. In insurance contracts, terms and conditions enumerate, in advance, the types of behavior that are closely related to moral hazard on the part of the policyholder, and insurance premium gets discounted–or surcharged–based on whether or not a given behavior is compliant with the aforementioned terms and conditions. This type of contract has been recognized as necessary in that it monitors the moral hazard of policyholders and, thereby, creates more socially desirable results. The same logic applies if the system can be regarded as a watchdog for curbing moral hazard among public institutions in the sense that it does notify public institutions of evaluation indicators beforehand and give monetary gains (or losses) in the form of performance-based pay, depending on whether or not they comply with those indicators. If this is the case, then the current PIMPES can be acknowledged to have some degree of rationality. In this respect, we will verify with data whether the current PIMPES has served as a watchdog

to control morally hazardous behaviors among public institutions.

In addition to its role as a control mechanism concerning moral hazard, the PIMPES is also required to play the role of enforcing an incentive contract by encouraging the members of a given public institution to fulfill their expected roles that suit the purpose of its establishment. In order for the PIMPES to be effective as an incentive contract, objective measurability of management performance is a prerequisite. If the agent has made sufficient efforts to achieve a good performance and, yet, the measurement of performance fails to reflect this, a correspondence between effort and performance falls apart. Instead, non-essential factors would prevail in the evaluation process, which means the failure of the PIMPES in playing its role as an effective incentive contract.

This leads to our second proposition to be tested: to which extent does the outcome of management evaluation get influenced by factors that are not so much related to the degree of a given public institution's efforts to improve its efficiency or publicness? In this study, we will consider the size and age of a given public institution to be the case of non-essential factors that are not related to efforts made by the institution and examine how these factors affect evaluation outcomes.

Another way to infer how much non-essential factors—i.e. factors unrelated to the efforts of a given public institution—are affecting the evaluation outcome of its management performance is to look at variability in the ranking order of evaluation criteria employed by the PIMPES. Intuitively speaking, short-term variability is not likely to be significant—considering the institutional characteristic of the PIMPES being conducted annually—if we assume that evaluation outcomes are determined by essential factors, such as a given public institution's efficiency. If the outcomes are, on the contrary, determined by a sheer chance or non-essential factors—for instance, in the case in which a public institution receives high scores just because of well-prepared documents for evaluation or because a certain person in favor of the institution happens to be included in the evaluation group—, the ranking order of evaluation criteria would change frequently year by year and the extent of such changes would be quite large.

The last proposition to be tested is related to the most important prediction made by the Holmström and Milgrom model (1991). That is, how is an incentive

contract to be designed in order to produce optimal outcomes in terms of social welfare when performance evaluation has to be conducted on tasks with discrete natures and varying degrees of measurability, such as management efficiency (or business value) and public interest?

If the agent has multiple types of actions to take, it may be easy to observe the extent of efforts put in for some of the actions, while not so easy for others. According to Holmström and Milgrom (1991), the optimal contract, in this case, is to tie the agent's wage system very weakly to two different performance measures (Bitmaro Kim, 2016). This means that a performance-based pay system should not be designed in such a way to reflect what is difficult to measure objectively. This is because the allocation of efforts made by the agent would be distorted if a performance-based pay system is greatly swayed by a particular type of performance that is difficult to assess in terms of efforts put into it—e.g. leadership and responsible management. This paper will examine the extent to which differentials in performance-based pay across public institutions are affected by indicators used for evaluating leadership and responsibility management, with its primary interest in how to optimally design a performance-based pay contract under the PIMPS.

All in all, this study presents the following four propositions in order to look into whether the current PIMPES has been designed and operated to properly function as a means to enhance management efficiency and public interest among public institutions.

1. *Public Institutions with a serious problem of moral hazard (for example, institutions with a high debt ratio or a high level of spending on welfare benefits) should be disadvantaged on management performance evaluation.*
2. *In the case of public institutions whose scores are poor on such evaluation criteria this period, more efforts should be made to solve the moral hazard in the next period (for example, to address the debt ratio or the high level of spending on welfare benefits.)*
3. *The results of management performance evaluation should not be determined by factors that are independent of the institution's efforts to improve efficiency or publicness (for instance, the size or age of given public institution.)*

4. *The results of management performance evaluation should not be heavily influenced by items that do not seem to contain accurate information on the level of effort made to improve efficiency or publicness (for example, leadership or responsible management categories.)*

2 Main Findings

The results of the analysis of management performance evaluation conducted on over 100 public institutions annually from 2008 to 2013 are as follows. First, the existing PIMPES has not functioned properly in reducing the degree of moral hazard among public institutions. We found no evidence that the evaluation grade of a public institution with a higher debt ratio is lower than that of an institution with a lower debt ratio. Neither did we find evidence that public corporations that had received poor grades in terms of debt ratio showed an improvement in the following period. The same applies to the case of welfare benefit expenditure, another indicator to measure the level of moral hazard. The level of spending on welfare benefits did not reveal a significant difference in the results of management performance evaluation. In other words, public institutions with a high level of welfare benefit expenditure did not receive disadvantages on management performance evaluation.

Secondly, management performance evaluation results were influenced by factors that were not related to efforts made to improve the efficiency or publicness of a public institution in question. Instead, the size of assets—at the significance level of 10%—affected the evaluation outcomes. In other words, it has been confirmed that public institutions with large assets receive more favorable evaluation results than those with small- or medium-sized assets. The age of a given public institution was also influential. That is, the longer a given public institution has remained in business, the better the evaluation outcomes are. Considering that the implementation of the PIMPES—with the current framework—has been over 10 years at least, older public institutions are more likely to have internally accumulated their experience with the existing PIMPES. This may have resulted in favorable outcomes on their management performance

evaluation.

As for the question of whether evaluation outcomes are swayed—regardless of the level of effort—by variability in the ranking order of evaluation criteria, our main findings are as follows. By using a transition probability matrix, we looked into whether this phenomenon is actually observed. The result shows that when a public institution is categorized into one of the three—i.e. upper, middle or lower—groups based on evaluation grades in a given period, the probability of its remaining in the same group in the following period was only 45%. To put it differently, public institutions have a 55% chance of moving from one group to another, which is quite high.

In particular, the probability that an institution previously categorized as the upper group fall down to the lower group in the following year—and vice versa—was 13%. In other words, 13 out of 100 public institutions experienced a large change in their rankings within the short period of one year. From the fact that rankings fluctuate year by year and the extent of such fluctuations is pretty large, we can infer that the current PIMPES is considerably influenced by non-essential factors or short-term efforts made by individual public institutions to improve their evaluation outcomes.

Finally, we examined the extent to which the outcomes of management performance evaluation are influenced by items that have large measurement error due to their intrinsic nature and do not provide reliable information regarding the level of effort made by a given public institution—for instance, variables like leadership and responsible management. Our analysis indicates that the overall rankings of public institutions are most influenced by rankings based on leadership and responsible management indicators. In other words, leadership and responsible management indicators take up a very large portion, compared to other indicators (e.g. achievements in major projects or efforts to improve management efficiency), in the process of ranking public institutions.

The results of analysis did not change much by year. Based on data from 2008 to 2013—excluding 2011—, we found that leadership and responsible management indicators, which are assessed in a non-metric manner, were relatively more important, compared to other indicators, in determining the total rankings of public institutions evaluated under the PIMPES. Considering that a significant change has been introduced in terms of evaluation items since 2011,

we even divided the time horizon of our analysis into the periods before and after 2011, the result of which did not make much difference. This implies that the size of a given public institution's performance-based pay is greatly influenced by scores on leadership and responsible management.

3 Policy Implications

Based on the empirical analysis presented above, we propose the following measures to improve the PIMPES. First, the function (or purpose) of the PIMPES needs to be clarified. The current PIMPES has been designed to make it difficult for public institutions to respond swiftly to changing business environments. This is because no weights are given to such efforts. Due to inefficiency stemming from the rigidity of existing evaluation indicators, it is inevitable that the PIMPES' role as an incentive contract gets weak.

If the government wants to maintain the evaluation framework of the current PIMPES, it needs to operate the system as a measure of control over the moral hazard of public institutions, not as an incentive contract. The evaluation target of the PIMPES should be limited to to several objective financial indicators, such as debt ratio and welfare benefits, and give incentives—or penalties in some cases—to public institutions based on those evaluation indicators. Debt ratio and welfare benefits are indicators that quite accurately show the actual level of moral hazard within public institutions. In addition, there exists a sufficiently rational ground to regulate operating expenses like debt ratio or welfare benefits since neglecting a rise in such expenses incurred by moral hazard among public institutions can amount to a heavy financial burden on national finance.

The second proposal is an issue related to the design of evaluation indicators. If policymakers aim to enhance public institutions' performance and their efforts for improvement through evaluation measures, the current PIMPES needs an extensive institutional redesign. According to contract theory, one of the most important prerequisites for an incentive contract to work effectively is that performance indicators must reflect the level of effort made by the agent as

accurately as possible. In other words, if the results of performance evaluation are determined by factors irrelevant to the efforts of the public institutions, the role of the PIMPES as an incentive contract is greatly undermined.

To this end, it is necessary to simplify evaluation criteria so that the burdens of being evaluated can be alleviated for public institutions. As the findings of this study suggest, evaluation outcomes being affected by the size or age of a given institution mean that the current PIMPES imposes an unnecessary burden on public institutions subject to its performance evaluation. There is a widespread perception among public institutions that getting higher scores on non-metric evaluation indicators require sleek-looking reports prepared with an extensive amount of financial and human resources (Joo-chan Kim and Min-chang Lee, 2014). In order to eliminate the particular inefficiency caused by this distorted perception, it is necessary to drastically cut down on evaluation indicators and to simplify their contents as much as possible.

Also, performance evaluation based on "adjusted" financial information or "target amounts achieved"—rather than actual financial information—should be urgently redressed in that this ends up producing inaccurate information about the long-term financial soundness of a given public institution. Some argue that this type of evaluation method is inevitable in order to take into account risks that are beyond public institutions' control. However, debates about the objectivity and fairness of performance evaluation arising from risks that public institutions cannot control should be resolved by adjusting the amount of performance-based pay and sensitivity thereof, not by arbitrarily adjusting evaluation indicators in question.

Moreover, the government needs to make a decisive move by excluding indicators that do not mirror the actual level of efforts made by public institutions or that are not directly related to the goal of publicness to be achieved through business activities the objectives of which are unique to a given public institution. Such indicators include: customer satisfaction, leadership (2013 standard indicators), strategic planning and organizational innovation (2015) and government-recommended policy indicators. The more these indicators sway performance evaluation outcomes, the less effective the PIMPES' function becomes as an incentive contract.

Finally, we need a more thorough review of the appropriateness of

institutional design concerning an existing performance-based pay system. The current PIMPES requires public institutions to take a balanced approach to efficiency and public interest. To this end, the government has tied performance-based pay to evaluation outcomes based on indicators targeting these two goals. Yet, a problem arises from the fact that the degree of accuracy differs greatly between the two. What Holmström and Milgrom (1991) suggest with respect to this particular case is that we need to have public institutions pursue efficiency and publicness in a balanced manner by avoiding compensation for tasks whose performance is difficult to measure and, at the same time, by lowering performance-based pay for the remaining tasks that are relatively easy to measure. This means that what is socially optimal may lie in gradually reducing the proportion of performance-based pay or, in some cases, abolishing it altogether under the current PIMPES.

Many studies have been conducted on the PIMPES, mainly by scholars of business administration and public administration. To the best of our knowledge, however, this is the very first study that conceptualizes the nature of the PIMPES as an incentive contract between the principal and the agent and analyzes the appropriateness of its institutional design. One of the greatest advantages of this approach is that we can tap into the framework and policy implications of contract theory developed over the last forty years by both theoretical and empirical economists and apply them to the analysis and improvement of the PIMPES. Through this process, we have been able to look at the problem of institutional design concerning the current PIMPES more scientifically and accurately recognize internal trade-offs among various evaluation indicators that make up the system, thereby, drawing several policy implications required for improving the PIMPES. In particular, this study can contribute to expanding the horizon of possible performance evaluation measures regarding the public sector, considering the reality in which the PIMPES has been a reference point for evaluating the performance of local public corporations and financial public institutions.

The major components and contents of this study are organized in the following order. Chapter II briefly introduces the characteristics of the PIMPES and previous studies thereof. In Chapter III, we examine these characteristics from a theoretical perspective. In particular, we employ the Holmström and

Milgrom model (1991) in order to take stock of prerequisites and components required for having performance-based pay serve properly as an incentive contract. Chapter Ⅳ presents evaluation indicators necessary for assessing the appropriateness of the current PIMPES from a normative perspective in the form of propositions. In Chapter V, we test these propositions by using actual data collected from the PIMPES. Finally, Chapter VI concludes with policy suggestions to improve the current PIMPES.

Ⅱ

Institutional Background of the PIMPES and Previous Studies

1 Main Characteristics of the PIMPES

It is said that South Korea's management performance evaluation of public institutions was first attempted in 1968. At the time, however, most public institutions were under strong government control. Naturally, the extent of their autonomy in the planning of budgets and business and personnel affairs were very much limited. In such a situation, the evaluating of public institutions' management performance had various limitations.

In this regard, it is more accurate to say that the introduction and implementation of a full-fledged performance evaluation system began with the Framework Act on the Management of Government-Invested Institution (hereinafter the FAMGII) enacted in 1983. From 1983 to 2003, the performance evaluation system was operated pursuant to the FAMGII. The number of public institutions evaluated was not constant year by year since certain institutions were often excluded or added as the situation dictated. Roughly speaking, the number ranged from 12 to 25 in the years from 1983 to 2003. In 1999, the target of performance evaluation was extended to include the heads of public institutions as the evaluation of management contracts with presidents of public institutions and presidents themselves was put into action.

From 2004 to 2006, 88 government-affiliated institutions were added as the target of performance evaluation. With the enactment of the Framework Act

on the Management of Government-Affiliated Institutions (hereinafter the FAMGAI), the performance evaluation of public institutions was conducted pursuant to the FAMGAI as well as to the preexisting FAMGII. Evaluation types became more specified and the number of types also increased to 3 government-invested institutions and 8 government-affiliated institutions from 2 to 5 preexisting categories (Korea Institute of Public Finance 2010).

As the AMPI was introduced in in 2007, the management of public institutions became more systematized. The scope of institutions recognized as public corporations or quasi-governmental institutions became more clarified as well. Starting from 2008, public institutions and their heads were separately evaluated based on different criteria. In 2010, there was a significant improvement in the PIMPES. The main aspects of the revision are as in Table II-1 below.

〈Table II-1〉 **Revision of the PIMPES in 2010**

	Purpose	Main Features
1	Strengthening the competitiveness of public corporations	Compare the evaluation outcome of a given public institution with that of global corporations, not with the previous year's outcome of the public institution concerned
2	Strengthening social responsibility	Create new indicators to evaluate social responsibility (e.g. job creation, shared growth, etc.)
3	Creating jobs	Strengthen evaluation measures to reflect actual performance concerning job creation
4	Alleviating the burden of being evaluated among public institutions	Cut down on and simplify evaluation indicators
5	Customizing evaluation by public institution types	Empower public institutions to select evaluation indicators
6	Promoting fair evaluation	Diversify evaluation group members (Various figures including the CEOs of private corporations)
7	Preventing reckless management	Impose a stricter penalty on reckless management
8	Reflecting current situations	Strengthen evaluation measures on tasks concerning current social issues
9	Improving the satisfaction of the public	Reflect evaluation by the public

Source: Press release by the Ministry of Strategy and Finance, 「Major Revision of the Public Institution Management Performance Evaluation System」 December 29, 2010

Since 2011, public corporations and quasi-governmental institutions have been classified in more detail: type I public corporation; type II public corporation II; fund-management type; commissioned-service type; and small-but-strong type. Evaluation categories can largely be divided into general management and major projects. The subcategories of general management include: management strategy and social responsibility; work efficiency; the management of organization, human resources and performance; financial budget management and performance thereof, and pay and welfare management. As for major projects, subcategories concern the planning, activities and outcomes of the major projects unique to a given public institution (Ministry of Strategy and Finance 2016A).

There are six grades—from an S to an E—as per performance evaluation outcomes. Based on the outcomes, follow-up measures are carried out, which includes differential performance-based pays, personnel actions and budgetary changes in the following year. The nature of the PIMPES as an incentive contract is clearly revealed by the fact that performance-based pays for the head of a given institution and its employees fluctuate according to evaluation outcomes. According to the 2016 fiscal guideline on public corporations and quasi-governmental institutions, public corporations receive performance-based pay within 250% of the basic monthly wage; quasi-governmental institutions receive performance-based pay within 100% of the wage; and preexisting government-invested institutions receive performance-based pay within 300% of the basic monthly wage (Ministry of Strategy and Finance 2016B).

[Figure II-1] and [Figure II-2] show evaluation indicators and weight references regarding public corporations and quasi-governmental institutions, which have been gathered from the 2012 Public Institution Management Evaluation Manual.

[Figure II-1] Evaluation Indicators and Weight References Concerning Public Corporations and Quasi-Governmental Institutions (Small- and Medium-Sized Institutions Excluded)

Evaluation Category	Evaluation Indicator	Public Corporation		Quasi-Governmental Institutions							
				Commissioned-Service Type		Fund-Management Type					
						Pension Type		Policy-Finance Type		Project-Fund Type	
		Non-Metric	Metric	Non-Metric	Metric	Non-Metric	Metric	Non-Metric	Metric	Non-Metric	Metric
Leadership and Responsible Management	1. Leadership	5		5		5		5		5	
	2. Responsible Management	3		3		3		3		3	
	3. Public Evaluation		5		8		8		8		8
	4. Social Contribution										
	- Social Responsibility	2		2		2		2		2	
	- Government-Recommended Policy		5		5		5		5		5
	Subtotal	10	10	10	13	10	13	10	13	10	13
Management Efficiency2」	1. Work Efficiency		8		6		6		6		6
	2. Organization and Human Resources Management	4		4		4		4		4	
	3. Financial Budget Management and Performance Thereof 1」										
	- Financial Budget Management	4		4		6		5		5	
	- Financial Budget Performance		6		2		4		2		1
	- Metric Management Costs		2		5		4		4		4

[Figure II-1] **Continue**

Evaluation Category	Evaluation Indicator	Public Corporation		Quasi-Governmental Institutions							
				Commissioned-Service Type		Fund-Management Type					
						Pension Type		Policy-Finance Type		Project-Fund Type	
		Non-Metric	Metric	Non-Metric	Metric	Non-Metric	Metric	Non-Metric	Metric	Non-Metric	Metric
Management Efficiency2」	4. Management of Wage and Performance-Based Pay - Management of Wage and Performance- Based Pay	4		4		4		4		4	
	- Increase Rates of Total Labor Costs		4		4		4		4		4
	5. Labor Relations Management	3		3		3		3		3	
	Subtotal	15	20	15	17	17	18	16	16	16	15
Major Projects 2	Subtotal	15	30	25	20	22	20	25	20	26	20
Total		40	60	50	50	49	51	51	49	52	48

1」 As for the fund-management type, the 'financial budget management and performance thereof (financial budget management, financial budget performance and metric management operational costs)' indicator has been changed to the 'fund operation management and performance thereof (fund operation management, fund operation performance and metric management operational costs)' indicator.

2」 Weights have been adjusted for each evaluation indicator, as per particular characteristics to be considered (e.g. current tasks unique to individual public institutions and limitations in deriving appropriate performance indicators).

Source: Ministry of Strategy and Finance, 2012 Public Institution Management Evaluation Manual, p.13, December 2011

[Figure II-2] Evaluation Indicators and Weight References Concerning Small- and Medium-Sized Quasi-Governmental Institutions

Evaluation Category	Evaluation Indicator	Small– and Medium–Sized	
		Non–Metric	Metric
Leadership, Responsible management	1. Public Evaluation		8
	2. Government–Recommended Policy		5
	Subtotal		13
Management Efficiency	1. Work Efficiency		6
	2. Financial Budget Performance1」 – Financial Budget Performance1」 – Metric Management Costs		2 5
	3. Increase Rates of Total Labor Costs		4
	Subtotal		17
Major Projects 2」	1. Appropriateness of Performance Management Concerning Major Projects	10	
	2. Major Project Performance		20
	Subtotal	10	20
Total		10	50

1」 In the case of small–and medium–sized fund–management–type quasi–governmental institutions, the 'financial budget performance' indicator has been changed to the 'fund operation management and performance thereof' indicator.

2」 The weights of metric and non–metric indicators have been adjusted for each evaluation indicator, as per particular characteristics to be considered (e.g. limitations in deriving appropriate performance indicators for individual public institutions).

Source: Ministry of Strategy and Finance, 2012 Public Institution Management Evaluation Manual, p.14, December 2011

2 Previous Studies

Studies on the PIMPES have been performed concerning the following aspects: the significance and necessity of the PIMPES; the appropriateness of evaluation indicators and the analysis of factors influencing evaluation outcomes; and policy suggestions for systemic improvements.

The Korea Institute of Public Finance (2010) and Ji-in Jang et al. (2013) set out the history of the introduction of the PIMPES and a series of institutional

changes made in the process. Won-hee Lee and Young-jae La (2015) provide a variety of information, such as the introduction of the PIMPES, the background of its institutional design and the purpose of the evaluation system, by tapping into recollections by officers of competent authorities and the heads and administrative secretaries of the evaluation group who participated in the design and operation processes of the PIMPES.

As for the design of the PIMPES, Seok-won Lee (2005) points to the problem of unfair comparison arising from the application of uniform evaluation criteria to various public institutions with discrete objectives and varying natures of their respective projects. Chang-gil Lee and Sung-rak Choi (2009) argue that non-metric evaluation—which is nominally regarded as absolute evaluation—is practically conducted in relative terms. Won-hee Kim (2010) points out that it is excessive to make an exception of the risk factor concerning changes in policy environments in the evaluation manual, and argues that this has the effect of distorting information about the long-term financial soundness of public institutions.

In relation to systemic improvement, Mi-jeong Park (2010) maintains that evaluation indicators and weights should be applied differentially by type and that rankings should be excluded from evaluation outcomes. She also raises a concern about the halo effect of subjective evaluation that could affect the public image of a given public institution. Oh Lee and Seung-hyun Yu (2010) point out the heavy burden of being evaluated to be borne by public institutions, and assert that evaluation should be conducted based only on metric indicators concerning select key performance areas in the case of public institutions that have been evaluated to be high-performing. In addition, they point out that small- and medium-sized institutions are likely to receive higher scores than other types because only metric indicators concerning management performance are employed in their evaluation.

Based on a survey of public institution employees, Tae-beom Yoon and Young-jae La (2013) raise a question about whether the amount of human resources, efforts and time invested in preparations for performance evaluation is reasonable, through which they suggest various ways for improvement so as to reduce the burden of performance evaluation on the part of public institutions.

Ⅲ

Theoretical View of the PIMPES

1 Economic Rationale for the PIMPES

A. Comparison of Stock Option Contract for CEOs and the PIMPES

In order to analyze the PIMPES from an economic perspective, the first thing that needs to be done is to ask about the rationale for the existence of the PIMPES. There is one thing to be cautioned here, however. The intention of asking this question is not to suggest that attempts to measure the performance of public institutions are unnecessary. Rather, our question is premised on the necessity of such measurement and aimed at examining advantages and disadvantages of a particular method of evaluation used in the current PIMPES. As mentioned above, what we mean by a "particular method" here concerns the following three components that cannot be found, specifically, in the performance evaluation of private corporations.

First, in the case of public institutions, many people try to measure and evaluate their performance in accordance with a distinct criterion called public interest, in addition to efficiency. As pointed out by James Wilson (1989) about organizational differences between the U.S. federal government and private corporations, what sets the federal government apart from private corporations lies in—among others—the fuzziness of the former's organizational mission and the rationale behind its establishment.

Second, public institutions' efforts to achieve the multiple tasks of efficiency

and public interest are not evaluated by the outcome only. The actual process of the PIMPES begins by breaking down the efforts of public institutions to attain their organizational goals into three phases: i.e. plan, performance and result. Next, evaluation categories are set up, corresponding to each of the three phases (leadership, management system and management performance). Then, several sub-evaluation indicators are assigned to each evaluation category, and, for each evaluation indicator, a weight is arbitrarily applied. With all the scores of these evaluation indicators tallied come a given public institution's final score (and rank).

As will be discussed later in more detail, there are many components in this evaluation process that are rarely found among private corporations. For example, there exist a number of evaluation items included in the PIMPES indicators—such as public evaluation, social responsibility and work efficiency—that are inherently difficult to quantify and measure and, therefore, inevitably cause a dispute about the objectivity of such evaluation.

Moreover, when assessing an institution's performance, the PIMPES independently assesses the appropriateness of planning and implementation phases, in addition to the outcome of efforts made. Then, a question arises: if the PIMPES can evaluate as appropriate a given public institution's plan made for the purpose of conducting a certain business and its implementation process, what is the standard? In other words, in evaluating the propriety of a business plan and its execution method, the question comes down to whether there can be independent evaluation criteria, other than how much such business has brought good outcomes.

Let's assume that we have already evaluated the business outcome of the public institution and kept that information in hand and that the evaluation results provide accurate information about the level of effort made by the institution (or the degree of efficiency or publicness achieved). In this case, there is no reason for the evaluator to use additional information concerning business planning and implementation in order to assess the performance of the institution in question. This is because the amount of new information, from which the evaluator intends to infer about the institution's effort level, is very small—that is, the marginal value of such information is very low. In some cases, using such information may end up drawing incorrect conclusions. Suppose that the

final evaluation results are not good while the planning or its implement process was actually appropriate. This may be the result of inappropriate indicators used for evaluating them.

Especially, assessing the appropriateness of business planning or its implementation process is far more difficult than evaluating final results. Among others, this is because it is not so easy to find appropriate, objective indicators. In this sense, the evaluation of planning and its execution process often ends up relying on qualitative judgments, which can lead to an arbitrary evaluation. Even if it were not for the arbitrariness of the evaluation, problems arising from measurement errors always exist. Therefore, it is difficult to say that the evaluation results of planning or its execution process provide accurate information about a given public institution's efforts to achieve its goal, compared to the evaluation results of its final performance.

The problem is that distortion and inefficiency arise when a wage system is designed in this way—that is, based on inaccurate evaluation indicators that fail to adequately reflect a public institution's effort level. For this very reason, wages for executives is usually linked to stock prices in the case of private corporations. In contrast with public institutions, there is hardly any case in which a wage system relies on the evaluation of business plans or execution processes.

Finally, a unique element of the PIMPES, compared to the performance evaluation system of private corporations, is that all evaluation indicators are announced before public institutions perform their respective tasks and that those indicators are rarely modified once they have been announced. Efficiency-wise, this has the following problems.

First, public institutions' performance is assessed only on the basis of indicators proposed in advance, which significantly reduces the incentive for them to respond to contingencies. The performance of each institution is basically determined by the degree of effort that the institution puts in for a particular task, but at the same time, the uncertainty and contingency factor—which cannot be predicted in advance (e.g. external factors, such as a foreign exchange crisis and falling oil prices, or internal factors, such as sudden death of the CEO or a fire in a production facility)—influences its performance. In some cases, a performance variation among public institutions

is determined by how much they respond to such contingencies in a flexible and decisive manner.

The problem is that by determining evaluation indicators in advance, the existing PIMPES significantly reduces the incentive to actively respond to contingencies that can have a decisive impact on an institution's performance. Assume that an institution is aware that responsiveness to contingencies corresponds to the rationale for its establishment. However, as long as the the contingency factor is excluded from evaluation indicators, there is no incentive for the institution to actively deal with it. The dilemma of the Incheon International Airport Corporation mentioned earlier in the introduction is representative of such situation. Of course, one might argue that there is no big problem since contingencies occurring after evaluation indicators have been set are taken into account by modifying the indicators in the evaluation process.

However, if an institution is risk-averse, it will not take initiative in this until a request for the improvement of preexisting indicators is accepted. Let us go back to the case of the Incheon International Airport Corporation. Suppose that a large-scale investment in security increases its debt ratio and, thereby, results in short-term deterioration in financial performance due to lower profitability. From the viewpoint of the institution's president, he will not invest in security unless the management evaluation group recognizes the necessity of such investment and, accordingly, corrects relevant indicators—for instance, by not counting investment in security and safety matters as a cost.

If so, how much are requests for the improvement of the PIMPES evaluation indicators being accepted in reality? Actual data shows that the acceptance rate—and, therefore, attendant revisions of evaluation indicators— was only 30% on average in 2015.

〈Table III-1〉 **Request for Indicator Improvement and Acceptance Result in 2015**

Category	Type I Public Corporation	Type II Public Corporation	Fund-Management Type	Commissioned-Service Type	Small-But-Strong Type	Total
Number of Requests	47	66	81	63	145	402
First-Round Review	26	22	17	21	43	129
Second-Round Review	15	28	20	15	44	122
Final Review	15	32	14	22	36	119
Final Ratio	31.9%	48.5%	17.3%	34.9%	24.8%	29.6%

Source: Internal documents from the PIMPES evaluation group in 2016

In conclusion, the existing PIMPES has not been elaborately designed to meet its original purpose of serving as an incentive contract to raise public institutions' efforts to a socially required level. Intuitively speaking, the fact that no private corporation enters into an incentive contract in the same way as the PIMPES is the strongest evidence that unequivocally demonstrates the problem of institutional design concerning the system. Assume that someone plans to design the same kind of incentive contract as in the current PIMPES and apply it to a private corporation.

In this case, we should be able to find the following detail of an annual-salary contract in the market. In the general meeting of stockholders in March, the stockholders of the private corporation decide to divide performance evaluation into three parts composed of planning, implementation and outcome and to set evaluation indicators and weights for each of the three components, and announce that they will evaluate the management based only on these criteria and give out performance-based pays accordingly. As everyone knows, there is not a single private corporation that enters into such contract and determines the annual salaries of the management in this way.

B. Comparison between Terms and Conditions and the PIMPES

In the foregoing, we have argued that it is almost impossible to find a case

in which the members of the top management sign their annual salary contracts that are similar to what is currently done under the PIMPES in the market. However, this does not mean that there are no such contracts structured in the same way as the PIMPES. Terms and conditions provided by insurance companies to their insurants takes such forms in the sense that these contracts detail what the parties concerned have to do for each type or condition before signing and compliance is monitored according to what the contracts stipulate.

For example, life insurance companies include a provision in their terms and conditions that they may refuse to pay insurance money if the insured person commits suicide within a certain period of time after the purchase of such insurance. There is a similar provision in fire insurance as well. One representative case is the discounting of premiums provided when the insured install fire detectors or fire-fighting equipment in their buildings (Gersby et al. 2014, p. 670).

The existence of this type of contract is attributable to the problem of moral hazard. To alleviate the problem of a so-called "information asymmetry," in which the behaviors of the insured cannot be observed after contract signing, insurance companies taxonomize typical cases of moral hazard and include them in their terms and conditions. As the precondition for the insured to be covered by insurance, they define the type of behaviors that may constitute moral hazard beforehand. After signing, insurance companies monitor and evaluate the compliance of the insured and pay insurance money differentially based on such evaluation, which is the very point on which terms and conditions are structurally similar to the current PIMPES. From the case of insurance contracts illustrated above can we deduce an implication that the introduction of the PIMPES may lead to a more desirable outcome in terms of social welfare. That is, the PIMPES acquires its significance when it functions as an institutional measure of control over moral hazard among public institutions. Let us explain this in more detail.

There is an asymmetry of information between the government and public institutions as between insurance companies and insurants. It is difficult for government officials to oversee whether a public institution concerned is making necessary efforts to carry out its delegated tasks prescribed by the purpose of its establishment and relevant laws. It may be because government officials lack

the high level of expertise needed to evaluate the institution's work relative to the operators of the public institution or because the public institution itself does not provide accurate information to the government officials. For whatever reason, the problem of moral hazard arises from this information asymmetry.

In addition to the general information-asymmetry issue, there are additional factors that can worsen the problem of moral hazard among public institutions. Unlike private corporations, public institutions face virtually no risk of being weeded out from the market, i.e., bankruptcy. This is because there exist strong expectations about the government's guarantees of public institutions' debts (Han-soo Choi and Chang-min Lee 2015). In other words, various stakeholders of public institutions (creditors, workers and the heads public institutions) expect that public institutions will be fully protected from the impact of such incidents as defaults due to the government's (implicit) guarantee. In short, they have a belief that the government is providing them with full insurance.

The belief in having insurance against risks incites the insured to be risk-taking. Han-soo Choi and Chang-min Lee (2015) demonstrate that public institutions' excessive debts can be attributed to moral hazard resulting from expectations about the government's implicit guarantee of public institutions' debts. Moral hazard often accounts for many other inefficient management behaviors among public institutions than excessive debts. Thus, the need for the PIMPES can be more readily acknowledged if its key evaluation items and indicators therein concern typical signs of moral hazard occurring in public institutions.

C. Multidimensional Nature of the Founding Purpose of Public Institutions

In general, public institutions have an additional purpose called publicness, unlike private corporations whose sole purpose is maximizing profits. For this reason, the performance of public institutions cannot be measured with a single objective indicator, such as the amount of profits or stock prices. That is, it is unavoidable to use non-metric indicators in evaluating the extent to which a particular public institution is making efforts to achieve one of its founding purposes—i.e. promoting publicness. Besides publicness, the peculiarities of public institutions—compared to private corporations—should be considered in

evaluating efficiency.

Unlike private corporations, public institutions have little autonomy or discretion over their personnel and budgetary matters. As for goods or services produced by public institutions, pricing is often controlled for public purposes, rather than freely determined by market principles of supply and demand. In this regard, it is unreasonable to evaluate the degree of efforts made by public institutions based only on indicators concerning business outcomes, as in private corporations. Instead, it is necessary to comprehensively evaluate the whole process: i.e. mission and vision establishment; strategy and planning; and implementation. In evaluating the whole management process, there may be some gaps in terms of what is actually intended to evaluate for each item. However, performance is measured with objective indicators (e.g. a customer satisfaction survey used to evaluate leadership and responsible management items) and subjective indicators (e.g. subjective assessment of leadership by the evaluation group) that are available.

This is a valid argument in some sense. For the argument to be more persuasive, however, two additional issues need to be examined. Above all, how accurate is the information conveyed by existing evaluation indicators with respect to a public institution's efforts to improve its leadership or responsible management aspects? Another issue pertains to the extent to which evaluation items whose measurement outcomes are not highly reliable (e.g. leadership and responsible management) actually affect overall evaluation results and the determination of performance-based pay. In Sections 2 and 3 of this chapter, we will use a simple theoretical model in order to discuss the practical implications of these issues in more detail.

2 The PIMPES as a Contract

According to contract theory, an incentive contract is needed to solve the problem of moral hazard arising from a situation in which one party cannot observe the other's behavior. Specifically, conditions that require an incentive contract can be described as follows.

(1) There are two economic entities, the principal (or, in some cases, stockholders or country) and the agent (or, in some cases, the management or public institution), and their respective interests are not in perfect agreement. (2) In terms of risk attitude, the agent is risk-averse while the principal is risk-neutral. (3) If the principal is not able to fully observe the agent's behavior (or the level of its efforts), the principal needs an incentive contract to raise the effort level of the agent. In particular, there should be no one-to-one correspondence between the agent's behavior (or the effort level) and its outcome. In other words, since the outcome is affected by factors other than the agent's efforts, the principal must infer the level of effort made by the agent after observing the outcome of the latter's behavior. In this case, the principal can design the agent's wage system in two ways. First, it is possible to pay a fixed salary to the agent regardless of the observed outcome. In this case, the principal does not quibble over why the outcome was not good. An unsatisfactory outcome may have resulted from the agent's negligence in its task or a sheer bad luck struck at the agent despite its best efforts (for example, a foreign exchange crisis as in the IMF bailout that South Korean went through in the past). Providing fixed wage means that the principal bears all the risks associated with the given contract. This corresponds to the case where even if a private corporation did not perform well in terms of operating profit, it still pays the promised salary to its employees.

Here, the problem is that if all the risks associated with the outcome are transferred to the principal, the agent will make efforts needed for avoiding a bad outcome (or for producing a desired outcome) below the socially required level. One of the ways to solve this problem is to tie the agent's wage to performance. That is, the principal can raise the level of the agent's efforts by signing an incentive contract such as the "performance-based pay" with the agent. However, the task of properly designing a performance-based pay system is not easy. One of the important factors to consider when designing a reward system is to properly measure the outcome, which is the very basis to which the agent's wage is tied. What problems will arise when the performance of the agent is measured by objective yet imprecise evaluation indicators?

Bolton and Dewatripont (2005, pp. 137-139) demonstrates that the level of effort the agent chooses to make decreases as the accuracy of indicators

measuring such level gets lower under relatively general assumptions[2] Below is a simple formula for explaining this.

Wage *(w)* received by the agent has the characteristics of performance-based pay tied to a specific performance indicator *(q)*. In this case, another component variable of wage *(s)* is the parameter value that shows how much the wage is affected by the performance indicator. If s is 0, then the wage *(w)* is determined irrespective of performance and the agent receives only the amount of t. In other words, wage is pegged at a fixed rate. On the other hand, if s is not 0, the greater its size, the more sensitive the wage gets to performance.

Assume that performance *(q)* consists of the agent's effort *(a)* and error term *(ε)*. Here, error term can be interpreted in two ways. One way is to interpret this as an external factor (e.g. external shocks like a foreign exchange crisis) that cannot be controlled by both the principal and the agent as contracting parties. Another interpretation is to understand this as a measure of the accuracy of the performance indicator used for assessing the level of effort made by the agent (i.e. the extent of measurement error). In other words, the large value of ε means that the indicator measuring performance provides inaccurate information about the level of effort put in by the agent.

The utility function (U_A) of the agent is defined as follows: UA $(w, a) = - e^{\eta [w-c (a)]}$. That is to say, the amount of utility felt by the agent goes up as the amount of wage it receives increases and as the level of effort required to carry out its task rises. The wage function of the principal (U_p) is defined as $U_p = q\text{-}w$.

The discussion so far is mathematically expressed as follows.

$$U_p = q\text{-}w$$
$$U_A(w, a) = -e^{\eta[w-c(a)]}$$
$$w = sq + t \ (t>0)$$
$$q = a + \varepsilon, \quad \varepsilon \sim N\ (0, \sigma^2)$$
$$c(a) = a^2$$

2) The assumptions are as follows: the contract is designed in such way that wage is linear to performance; the noise is normally distributed; and the utility function exhibits constant absolute risk aversion (CARA).

To solve the optimization problem under these conditions, the optimal level of effort made by the agent *(a*)* can be expressed as a function of σ^2 as in $a^w = a(\sigma^2)$. In other words, the optimal level of the agent's effort depends on the accuracy of the evaluation indicator that measures performance. If the relationship between the accuracy of the indicator and the optimal level of effort is examined again through comparative statics, what we reach is $\frac{da}{d\sigma^2} < 0$. That is, as σ^2 increases, the optimal level of the agent's effort a^* becomes smaller. In addition, the relationship between the accuracy of the evaluation indicator and the sensitivity of wage to performance can be confirmed as well: $\frac{ds}{d\sigma^2} < 0$. In other words, as σ^2 increases, the performance sensitivity of wage must also decrease.

In the above, σ^2 has been defined as measurement error that occurs in the process of converting the level of the agent's effort—which is not easy to observe from the outside—into an observable scale. Suppose σ^2 is infinite. This means that the indicator we observe to measure the level of the agent's effort does not actually provide any valuable information about how much effort the agent puts in to achieve the task.

Meanwhile, it can be explained in another way. Let us assume that the performance indicator is effectively determined by external factors that cannot be controlled by the agent's efforts. Here, "external factors that cannot be controlled" often refer to macroeconomic factors, such as the foreign exchange crisis followed by the IMF bailout program as mentioned above. However, there is no reasonable ground on which those external factors should be limited to macroeconomic aspects only. For example, government intervention is another kind of external factors that are beyond the control of public institutions.

Under these conditions, how is the level of an institution's efforts determined, and at what level should the principal set the performance sensitivity of the agent's wage? The conclusion is the same in both cases. The fact that we have been able to identify through the comparative statics process is that if the indicator used to measure the agent's performance does not provide accurate information about the level of effort that the agent actually has made, the agent always makes less effort than the socially desirable level. That is, the level of effort made

by the agent gets lower as the informativeness of the indicator declines. Therefore, in the extreme case in which information concerning the level of the agent's effort has practically no value—as when σ^2 is infinite—, the agent will not make any effort. This is because the agent tries to increase its utility level by lowering the cost of effort to 0 when effort does not affect performance at all.

Furthermore, since the principal is also aware of the fact that the agent does not make any effort when σ^2 is infinite, it will not try to infer the agent's level of effort through the indicator *(q)* that it observes.

This means that identifying the indicator *(q)* that would provide accurate information on the level of the agent's effort—i.e. that which has a small measurement error—is a very critical for the principal to raise the agent's effort level through the incentive contract. Also, identifying an objective indicator that can estimate the level of the agent's effort is a precondition for introducing and operating a performance-based pay system.

Our discussion so far can be rearranged to fit the context of the PIMPES as below. First, the institutional design and components of the current PIMPES are more similar to the terms and conditions of fire insurance than to the performance-based pay contract signed with a CEO. This legitimizes the policy authorities' use of the PIMPES as a tool for curbing moral hazard among public institutions. From this perspective, we can derive a proposition to assess the effectiveness of the PIMPES in this study: evaluation indicators should be improved to catch signs of moral hazard by introducing the PIMPES. Otherwise, it means that the PIMPES is not working properly.

Now, let's assume that the PIMPES is not a measure to control moral hazard, but a performance-based pay system as in stock options for CEOs. Like what is intended by adopting stock options, this means that the PIMPES must function as an incentive contract to raise the level of effort made by public institutions. The most important prerequisite for this is that evaluation indicators measuring the performance of public institutions should provide as accurate information as possible about the level of effort they put in. If the results of management evaluation are determined by factors irrelevant to a given institution's efforts to improve its efficiency or publicness, the role of the PIMPES as an incentive contract may result in socially undesirable outcomes. Also, in such cases, we need to reduce the performance sensitivity of wage.

3 Optimal Contract Theory concerning the Case of Multiple Tasks Delegated to the Agent

The theoretical review of incentive contract presented in the previous section deals with the problem of an organization trying to achieve a single goal delegated to it. For public institutions, however, the goal itself is complex and multidimensional. Various stakeholders—such as the government, the public, and customers—require them to make a "balanced" effort to attain the double goals of efficiency and public interest (Young-jae La and Tae-beom Yoon 2013).

The Holmström and Milgrom (1991) model to be introduced in this section concerns the problem of moral hazard that arises when multiple tasks are delegated to a single agent and the optimal contract design. The basic starting point for this model is whether it is desirable to adopt a performance-based pay system when an agent is required to carry out multiple tasks (or a single task with multidimensional characteristics).

The difficulties faced by the agent when the principal requires it to simultaneously perform multiple tasks are not as simple as a trade-off between risk sharing and incentives as in the case of general contracts.

For example, let's say there is an agent that needs to carry out tasks A and B at the same time. Suppose that wage for the agent depends on the outcome of task A. In other words, the outcome of task B does not affect the amount of wage that the agent would receive. In this case, the kind of problem faced by the agent is no longer about who will bear the risk of damage that goes beyond the control of both parties as we have discussed in the preceding example. Instead, what this multi-task situation poses to the agent is about how to allocate its effort and time between the two tasks—that is, task A that affects the amount of wage it will receive and task B that has no relevance to such matter. Simply put, the nature of the problem confronted by the agent is changed to the allocation of effort among several tasks (Bolton and Dewatripont 2005, pp. 227-219).

Let us explain this problem with a more realistic case. The government wants students to attain both numerical abilities and creativity through education, so it decides to pay a bonus to teachers so that they can perform their tasks more effectively. The level of numerical abilities will be evaluated by mathematics

scores on nationwide mock examinations. On the other hand, creativity will be excluded from performance evaluation since it is difficult to objectively measure changes in the level of creativity. When the bonus scheme is designed in this way—i.e. to be tied only to students' mathematics scores—, how much time and effort would teachers spend to foster their students' creativity?

To answer this question, Holmström and Milgrom (1991) formulate the following theoretical model. For convenience, let us explain this model based on the same assumptions and model specifications as in the principal-agent model discussed in the previous section. What is different here is that that the number of tasks *(i)* is now plural (i = 1, 2) and the agent's wage structure (w) depends on the outcomes of tasks 1 and 2 (q_i). Again, the performance of individual tasks relies on the amount of effort (a_i = 1, 2) and measurement error (ε_i) for each task.

In this case, s_i plays a crucial role in the design of the wage structure. Here, s_i shows how much the wage of the agent depends on the respective evaluation indicators for the two tasks (q_i)—that is, a variable that captures the performance sensitivity of wage. For example, if s_1 is 0, the wage of the agent is not affected by the performance of task 1 at all. If s_1 is 1, the agent's wage is entirely dependent on the performance of task 1. Both s_1 and s_2 being zero means that the agent's wage has a fixed structure not relying on performance at all. The explanation so far can be expressed with a formula as below.

$$w = t + s_1 q_1 + s_2 q_2$$
$$q_i = a_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma_i^2), \quad I = 1,2$$
$$c(a_1, a_2) = (c_1 a_1^2 + c_2 a_2^2) + 2\delta a_1 a_2$$

What is important here is the relationship between task 1 and task 2. It should first be noted that there is no externality between task 1 and task 2. In other words, the effort a_1 spent on task 1 by the agent only affects q_1, but it has no effect on q_2 (and vice versa). It should also be noted that the efforts to carry out task 1 and task 2 respectively substitute each other. This means that the level of effort for task 1 affects the marginal cost of task 2. The parameter to capture this relationship is δ. If $\delta > 0$, the two tasks are substitutes for each other and, thus, an increase in the level of effort for one task means a decrease

in the level of effort for the other (Bitmaro Kim, 1986).

The principal expects the agent to allocate some time to both task 1 and task 2 and to achieve the intended outcomes of both tasks. In other words, the utility function of the principal depends on the performance of both tasks 1 and 2. (In the case of teachers above, the education authorities want students to think creatively as much as they want the students to be good at mathematics well.) However, the problem is that the measurement scale for task 2—i.e. the goal of creativity development in the above case—is very incomplete. In other words, σ_2^2 is very large. Meanwhile, the performance evaluation of task 1 has an objective standard to compare achievements. Suppose that σ_1^2 is not so large. Under these conditions, Holmström and Milgrom (1991) present a solution to how the principal should determine s_1 and s_2 so that the agent makes a "well-balanced" effort for both tasks.

Their conclusion is as follows. The more incomplete the indicator measuring the performance of task 2 *(q_2)* is as a measure of the effort required for the agent to perform task 2 *(a_2)*—i.e., the larger σ_2^2 is—, the smaller s_2 should be. In other words, the wage of the agent should be designed to be less dependent on the performance indicator for task 2. In the extreme, if the indicator *(q_2)* does not provide any meaningful information about a_2—i.e., if σ_2^2 goes to infinity —s_2 should be zero. That is, the wage of the agent should be designed not to be affected by the performance indicator for task 2.

The insight of Holmström and Milgrom (1991) goes a step further here: if this is the case, then should wage be tied to the performance of task 1? If the effort *(a_1)* for task 1 and the effort for task 2 *(a_2)* substitute each other and the principal wants the agent to put effort into both task 1 and task 2, they argue, s_1 should also be set to zero. This means that the wage structure for the agent should be fixed wage with no rewards for improved performance. Let us explain why this is so by going back to the case of a performance-based pay scheme for of teachers mentioned above.

Suppose that it is virtually impossible to determine whether teachers have made efforts to increase students' creativity by using a given indicator because its design is too imprecise to measure the level of students' creativity. That is, s_2 should be 0 if σ_2^2 goes to infinity. But what if s_1 is not zero in this case? Teachers will spend all of their time in improving students' mathematics scores

while not even trying to foster the creativity of the students. Therefore, if the state wants teachers to allocate their time to both goals—i.e. numerical abilities and creativity— adequately, a fixed-wage scheme should be adopted.

This theoretical insight provides an important implication for the design of the PIMPES. The existing PIMPES requires public institutions to attain the double goals of efficiency and public interest in a "balanced" way. Efficiency (let us call it task 1) is measured in a relatively objective way by using such financial indicators as debt ratio and productivity concerning labor and capital. Public interest (task 2) is measured primarily by such evaluation items as leadership and responsible management.[3] Currently, the PIMPES has two different performance-based pay schemes as per the respective evaluation outcomes of tasks 1 and 2.

Here, the problem comes down to whether the evaluation indicators for task 2 (q_2) is indeed an accurate measure of a given institution's publicness. If the indicators associated with task 2 do not provide accurate information on the level of efforts made to improve the institution's contribution to public interest, the proportion of s_2 should be small. In other words, the evaluation outcome of task 2 should not be a major factor in determining differentials in performance-based pay among public institutions. We will discuss this in more detail in Section 3 of Chapter IV.

3) To put it differently, efficiency is usually measured by metric indicators while publicness–which is hard to quantify–is measured primarily by non–metric indicators.

Ⅳ

Propositions concerning the Appropriateness of the PIMPES Design

Based on contract theory discussed in the foregoing chapter, we will assess whether the current PIMPES is properly designed and operated. To test this, a set of propositions will be put forth in this chapter.

1 The PIMPES as a Moral-Hazard Control Device

The first issue to be examined is whether the current PIMPES fulfills its role and function as an institutional device to control the moral hazard of public institutions. To this end, we must first identify effective evaluation indicators to represent the degree of moral hazard among public institutions.

Debt ratio is a suitable choice for the following reasons. According to Han-soo Choi and Chang-min Lee (2015), the problem of moral hazard within public institutions is fundamentally caused by the fact that—unlike for private corporations—the management discipline mechanism of bankruptcy does not work under the government's (implicit) guarantee. As a result, this increases the tendency of public institutions to be risk-taking, which can be best captured by the size of a given institution's debts.

Of course, there are diverse factors that lead public institutions to accumulate debts. The high debt ratios among public institutions result from inefficient management, unreasonable promotion of national projects or excessive

government control over fees and fares (see Han-soo Choi and Chang-min Lee 2015, pp. 23-24). However, the most significant difference between public and private corporations in terms of how funds are raised for investment purposes is that public institutions often procure the capital through debts, rather than through equity. On the one hand, this is due to institutional constraints, yet on the other hand, the problem arises from the fact that public institutions can raise the capital at a significantly lower rate, compared to private corporations. This low cost of capital procurement is the result of the government's (implicit) guarantee of public institutions. If this was not the case, many of the projects carried out by public institutions in the name of national projects—for instance, the Four-River project and resource diplomacy—would have been stranded or severely scaled down by the mechanism of market discipline. Thus, debt ratio can be an apt indicator to measure the level of moral hazard incurred by the government's (implicit) guarantees.

In this study, we will use information on debt ratios of public institutions to examine whether the PIMPES is functioning properly as a mechanism to curb moral hazard among them. To this end, the following two propositions will be tested.

Propositions

1. *A public institution with a high debt ratio will be disadvantaged on management evaluation.*
2. *In the case of public corporations whose evaluation scores are poor in this period, their debt ratios will be improved in the following period.*

Proposition 1 is presented to examine how much management evaluation results provide meaningful information about the debt ratio of a public institution. Proposition 2 concerns the disciplinary effect of the PIMPES. If evaluation results are operating properly as a control mechanism to redress management practices related to moral hazard, they should affect the level of debt ratio in the next evaluation period. To be more specific, public institutions whose evaluation outcomes are not good in this period should show an improvement in their debt ratios in the next period. If not, this suggests that the PIMPES is not functioning

well as an institutional device of control over moral hazard among public institutions.

Welfare benefits are another good evaluation indicator to reveal of the degree of moral hazard among public institutions. In general, in the case of a public institution whose risk of bankruptcy has disappeared, they will be incentivized to raise the amount of expenses associated with private benefits for their employees higher than the optimal level. Of course, the most representative expense items related to moral hazard would be labor cost items, such as wages for the management (or workers).

In the case of labor costs in public institutions, however, the increase rates are controlled in a fairly thorough manner by the guidelines of the budget authorities, such as the Ministry of Strategy and Finance. Therefore, it is not easy to find a variation in the increase rate of labor costs among public institutions. For this reason, we will look at welfare benefits—which are similar to labor costs in character, but are less likely to be controlled by the budget authorities—as an alternative indicator to show a possible variation in operation practices among public institutions and examine how they affect evaluation results, which leads us to Proposition 2-1.

Proposition 2-1

A public institution that spends excessively on welfare benefits will be disadvantaged on management evaluation.

2 Appropriateness of Evaluation Indicator Design

As mentioned earlier, it is inevitable that the level of the optimum effort made by the agent goes down as the indicator for measuring it becomes more inaccurate. Assume that the goal of the PIMPES is only to improve the management efficiency of a given public institution—that is, let us set aside, for the moment, the goal of public interest to be achieved through business projects unique to individual institutions. In order to attain the goal of promoting

management efficiency, evaluation results should be determined by financial information, which can best represent the efficiency of a given institution. In other words, if financial performance is good, the ranking of the institution should go up as per its performance evaluation outcome. Suppose, however, that the outcome is influenced, for whatever reason, merely by the amount of human and material resources that an institution can mobilize to prepare for performance evaluation. If the PIMPES is designed in this way, public institutions will devote their effort and time to such matters by finding expedient measures to achieve good performance in the short term, rather than making a fundamental move to enhance their management efficiency.

Proposition 3

Management evaluation outcomes should not be affected by factors that are irrelevant to efforts to promote the publicness or efficiency of a given public institution.

What would be such factors? Previous studies have pointed out that those subject to management evaluation spend a considerable amount of human and material resources on evaluation preparation in order to receive higher scores. For example, Wan-hee Kim (2010) points out that it takes a significant amount of time just to understand how metric evaluation is conducted since its calculation formulae and evaluation methods are complicated. Meanwhile, Young-jae La and Tae-beom Yoon (2013) argue that the evaluation cycle of one year is too short for public institutions to get ready for performance evaluation, which increasingly burdens personnel in charge of evaluation preparation within public institutions. In particular, this has been a steady complaint raised by small-sized institutions.

An excessive evaluation burden accompanying the PIMPES process leads to a side effect that the evaluation results are swayed by a public institution's ability to mobilize resources to prepare for performance evaluation. After all, an institution's resource mobilization capacity depends on the size of the institution. Especially the larger the workforce, it is more likely that a given institution can have personnel dedicated to evaluation preparation, which can

affect evaluation outcomes. Therefore, this study will use information related to public institutions' human resources and assets to look into how these factors affect management evaluation results.

The age of a public institution may also affect the outcomes. Institutions with a longer business history may have accumulated much more know-how on the PIMPES, compared to newly established ones. If this intangible knowledge is handed down within the organization and passed on to personnel in charge of evaluation preparation, the institution is likely to receive higher scores on performance evaluation.

Another way to test Proposition 3 is to track a year-to-year variation in evaluation results. The reason why this can be used to verify the proposition is as follows. Suppose that management evaluation results are determined by an indicator that shows the level of efficiency. In many cases, the efficiency of a public institution is often determined by its operational capability, which does not change easily in the short term. In fact, this can only be improved through continuous investments and efforts with a long-term plan. Therefore, if such factor determines the outcomes of management evaluation, it is very unlikely that an institution belonging to a higher-ranking group would fall into a lower-ranking group just in one or two years (and vice versa.)

On the other hand, what if the results of management evaluation are determined by non-essential factors, such as a given institution’s ability to respond to evaluation or by temporary or exogenous factors like a sheer luck? In this case, it is possible that the evaluation ranking of a specific institution would fluctuate considerably. In the long term, such phenomena as mean reversion may occur.

3 Appropriateness of a Performance-Based Pay Scheme Design under the PIMPES

As mentioned above, the PIMPES includes several items that are difficult to objectively evaluate management performance within its evaluation indicators. Leadership and responsible management are two typical ones. According to the

2012 Public Institution Management Evaluation Manual, the category of leadership and responsible management consists of five evaluation indicators: leadership; responsible management; public evaluation; social contribution; and government-recommended policy (Ministry of Strategy and Finance 2011).

As for public corporations, 20 points are assigned in this category. Ten points are allotted to non-metric indicators—5 points for leadership, 3 points for responsible management and 2 points for social contribution. The remaining 10 points are set aside for metric indicators: 5 points for public evaluation and 5 points for government-recommended policy.[4] Leadership is defined as below: "The efforts and achievements of an institution's head concerning such matters of establishing and executing management strategies to fulfill the institution's vision and goals and of solving key current issues." Responsible management mainly focuses on official notice, governance structure and ethics management system while social contribution includes such items as job creation, shared growth with small- and medium-sized enterprises (SMEs) and volunteering activities.

Two main problems arise in the evaluation of public institutions' efforts based on the above items. The first is already mentioned above—that is, a question about where the rational basis lies in evaluating the leadership or responsible management items apart from the financial or business performance of public institutions. Historically, this is the product of having applied the Malcolm Baldrige model to management evaluation. In other words, Evaluation areas have been classified into three components of management *process*—i.e. management method, implementation and learning stages (Wan-hee Kim 2010). The problem is, as briefly stated above, that evaluating plan and implementation process—apart from the evaluation of final performance—is only meaningful in the following two cases—i.e. the planning and implementation phases have been poorly executed, yet their outcomes have been measured to be satisfactory (and vice versa). For consistency's sake, we should determine which evaluation results are more reliable than others. After all, this boils down to the question

4) In the case of quasi-governmental institutions, 8 points are assigned to public evaluation; as a whole, metric scores amount to 13 points and non-metric scores make up 10 points.

of which evaluation indicators are more reliable. In this case, the outcome of performance evaluation is generally more reliable than that of planning or implementation process evaluation.

In fact, the evaluation of leadership or responsible management as part of a public institution's planning or process is meaningful when the evaluation of the institution's financial soundness or major business performance produces a negative outcome while the planning or process parts themselves are good.[5] But even in this case, there still remains a question as to whether the evaluation of leadership provides accurate information about the level of public institutions' effort to improve their management performance. This is because, in essence, it remains to be answered whether the evaluation indicators of leadership or responsible management indeed reflect public institutions' efforts and achievements.

In fact, this is a universal problem occurring in the evaluation of non-metric (or qualitative) items. Let us assume, for instances, that a public institution (for example, the Korea Gas Corporation) was negligent in cooperation with SMEs or volunteering activity in a given year and that this resulted in low scores on the social responsibility item. In this case, do the low social responsibility scores provide accurate information on the institution's efforts to promote publicness and efficiency?

According to a survey of the PIMPES committee members, the percentages of those who responded positively to the question of whether public evaluation and government-recommended policy indicators within the leadership and responsible management evaluation category reflect institutions' efforts and performance were about 32 % and 34% respectively. However, these indicators should be regarded as ones whose measurement error is considerably large (Young-jae Lee and Tae-beom Yoon 2013). In this respect, the government itself announced that it would abolish the leadership and responsible management

5) In the last case, we can consider a case in which performance evaluation outcome is poor due to external shocks (e.g. financial crisis). Assume that leadership was good, but it has not been reflected in management efficiency or in the performance of major project owing to those external shocks. Even in this case, however, the PIMPES adopts a relative performance evaluation method–i.e. evaluation of public institutions by institution type. As such, there is no reason to solve this problem by introducing evaluation items, such as leadership.

category and transfer them to evaluation components concerning the management performance agreement system[6] (Young-jae La and Tae-beom Yoon 2013).

In the current PIMPES, the main reason for questioning the assessment of an institution's efforts with regard to the leadership and responsible management category is that such evaluation results have a significant impact on the institution's management evaluation rankings and, thus, may determine the size of performance-based pay. We have already discussed this problematic situation in which outcomes of management evaluation *(w)* are swayed by such inaccurate indicators. Here, let us explain this by using a simple formula as below.

The current wages of public institution employees are basically dependent on the sum of scores on three evaluation indicators: management efficiency; leadership and responsible management; and major projects. Let us call them q_1, q_2, and q_3 respectively. Here, the wage is composed of parts not affected by evaluation outcomes *(t)* and performance-based wage tied to such outcomes $(sf(q_1 + q_2 + q_3))$. s is the proportion of performance-based wage in the total amount of wage.

$$w = t + sf(q_1 + q_2 + q_3)$$

f( ) refers to the management evaluation grade function

If there exists little or no variation among public institutions in terms of scores on the management efficiency or major project categories, however, the management evaluation grade will be practically determined by evaluation results concerning the leadership and responsible management indicators *(q2)*. In this case, the wage function *(w)* changes, in effect, as follows.

$$w = t + s_2 f(q_2)$$
$$q_2 = a_2 + \varepsilon_2,\ \varepsilon \sim N(0, \sigma^2)$$

The leadership indicator is influenced by intrinsic factors—e.g., an institution head's efforts to improve the publicness and efficiency of an institution—that

6) However, this has yet been realized.

are intended to measure leadership and measurement error (ε_2). In the current situation, the leadership indicator does not provide accurate information about efforts made by the head of a public institution—i.e. ε_2 is large—because the leadership indicator *(q_2)* has not been set properly. As a result, efforts to enhance leadership *(a_2)* do not get properly evaluated. If this is the case, then the wages of public institution employees are, in fact, determined by external factors that are irrelevant to their efforts. From the analysis so far, we derive the final proposition of this study:

Proposition 4

Management evaluation results should not be greatly affected by items that are not considered to contain accurate information about the publicness of a given public institution—e.g. leadership or responsible management indicators.

V

Main Findings of Empirical Analysis

1 Basic Statistical Analysis of Public Institution Management Evaluation Results

A. Basic Statistical Analysis of Public Institution Management Evaluation Rankings

We have surveyed the entire public institutions—i.e. 138 public institutions—that were subject to management evaluation from 2008 to 2013 and conducted an empirical analysis to examine whether the existing PIMPES serves as an incentive contract mechanism.

As introduced in Chapter II, the PIMPES evaluates three main aspects: 1) leadership and responsible management; 2) management efficiency; and 3) major projects. According to sub-items for each evaluation indicator, management performance is assessed, the results of which are converted into scores. Then, appropriate weights are assigned to respective evaluation scores so as to produce the total score on a scale of 100 points. Based on the total scores, public institutions are grouped by type and ranked, after which final grades are determined. There are six grades from an S to an E—yet, grade S was rarely observed. Based on the final grades, the results of public institution management evaluation are published annually.

In this way, the PIMPES can be regarded as a system that relatively evaluates the performance of public institutions by type on the basis of the evaluation scores. Therefore, the management evaluation rankings in the type of a public

institution can be regarded as the final result of evaluation of the public institution management evaluation. In this study, the yearly management evaluation rankings will be considered a key variable that represents a public institution's management performance. We will use it as a main dependent variable when an empirical analysis of the PIMPES is attempted.

In this section, we will analyze what the PIMPES rankings reveal and their implications.

Since rank is a relative concept, it is not advisable to use it as a metric indicator. For instance, a rise by one rank among 100 institutions and a rise by one rank among 1,000 institutions do not have the same significance. In this study, therefore, we take into account the fact that the number of public institutions subject to management evaluation changes every year and present the standardized rank by year and institution type as follows.

$$\text{Standardized Rank: } zRank_{ijt} = \frac{Rank_{ijt}}{(\#\ of\ firms)_{jt}} \times 100$$

Specifically, public institution's *i*'s standardized rank in the type *j* in a given evaluation term *t* ($zRank_{ijt}$) is defined as the value obtained by dividing the public institution *i*'s rank in the type *j* in the term *t* ($Rank_{ijt}$) by the number of institutions belonging to the type *j* in term *t* $(\#\ of\ firms)_{jt}$ and then multiplying it by 100.

The standardized rank as presented above can be considered a weight-adjusted rank, and this can be easily understood by using the concept of distance. In other words, the standardized rank can be regarded as a state in which *n* number of public institutions are arranged in the order of evaluation scores in the same interval of (100/*n*) between 0 and 100 with the lowest rank (or the maximum value) being always 100. Thus, a one-rank rise in the standardized ranking means a one- rank rise among 100 institutions and a ten-rank rise among 1,000 institutions. In this way, the standardized rank has been conceptualized in relative terms, which has the merit of comparing public institutions by type and year.

Using this standardized rank, we can find the answers to the following two

questions related to the effectiveness of the PIMPES. The first question concerns the way in which year-to-year variability is figured in evaluation outcomes and the extent of such variation. In other words, it is related to whether it is possible the rankings change according to the level of efforts made to improve management performance and how much such change is possible.

The second question is about stability—or a lack thereof—in management evaluation rankings. That is, it asks whether high-ranking public institutions continue to be so while low-ranking ones tend to remain where they are. On the flip side, this is also related to the magnitude of transition probability in evaluation rankings: i.e. high-ranking institutions fall to low ranks and low-ranking ones rise to high ranks. At the same time, the question concerns the extent of the probability that individual institutions' efforts to improve their management performance are reflected in actual results.

Private corporations continuously receive incentives as well as pressures to improve management efficiency for the market. On the contrary, public institutions have basically few incentives to enhance their management efficiency. Since the purpose of the PIMPES lies in improving efficiency of management activities by providing the incentives, significant changes in the final evaluation ranking can be a very crucial yardstick for assessing the effectiveness of the PIMPES. In this respect, both questions above can be understood to ask if significant changes are observed in management evaluation rankings.

To test the magnitude of the variability in evaluation rankings, let us look at the basic statistics of the extent of variation by employing the 1st difference of the standardized rank as defined above. To this end, the standardized rank variation is formulated as follows: $\Delta nRank_{iit} = nRank_{iit} - nRank_{iit-1}$. Here, it should be noted that the rank would rise as the value of the difference becomes more negative since a higher rank means a smaller value of the standardized rank variation.

Above all, <Table V-1> shows the basic statistics of the standardized rank variation. Here, we can confirm several facts. First, the average value of the standardized rank variation is close to zero. The average value of the standardized rank variation goes negative when there are many cases in which public institutions see a sharp rise in their rankings; and the average value goes positive

when there are many cases of a dramatic drop in ranking among public institutions. However, when the average value is close to zero, it means that rank variation is not skewed in a certain direction.

Second, considering that the maximum value of the standardized rank variation is 100, the fact that standard deviations are 30.2 and 32.7 means year-to-year variability in evaluation rankings is considerably large. In addition, the average of standardized rank variation's absolute value is 25.5, which indicates that individual public institutions can expect their rankings to decrease or increase by an average of 25.5% per year. This reveals a large extent of variability in PIMPES rankings.

Finally, he minimum and maximum values of the standardized rank variation are close to -100 and 100 respectively. This suggests that there are indeed changes in evaluation rankings among public institutions: from the highest to the lowest and vice versa.

〈Table V-1〉 **Basic Statistics of Standardized Rank Variation (2007-2013)**

Variable	Observed Value	Mean	Standard Deviation	Minimum Value	Maximum Value
Standardized Rank Variation	596	−0.6023	32.6696	−92.1569	93.3333
Absolute Value of Standardized Rank Variation	596	25.5017	20.4019	0.0000	93.3333

Source: Data compiled by the authors based on internal documents of the Korea Institute of Public Finance

<Table V-2> presents data by calculating the basic statistics of <Table V-1> by year. Main findings are not much different from the previous ones. That is, the same pattern can be observed: the average of the standardized rank variation is close to zero; and variability of ranking is considerably large.

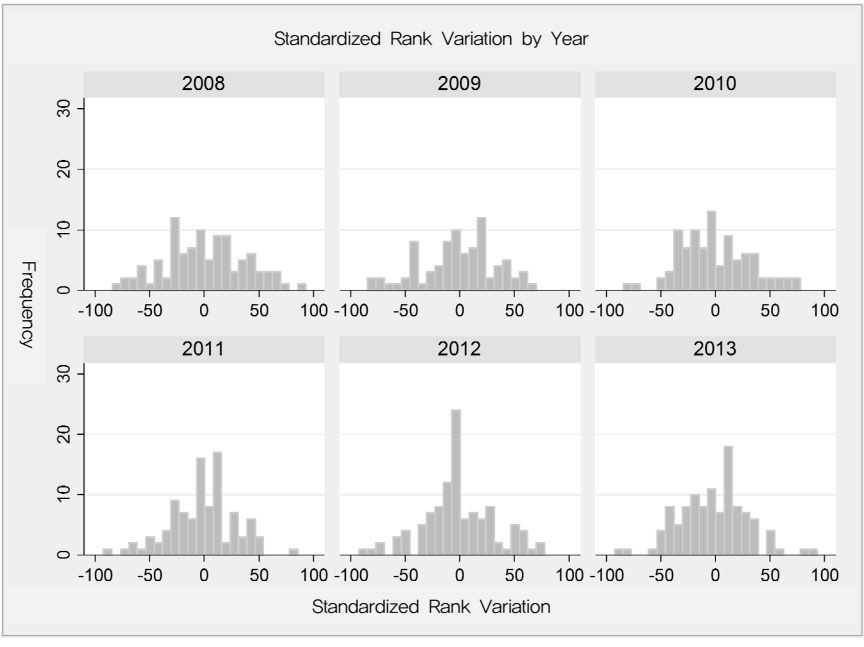
〈Table V-2〉 **Basic Statistics of Standardized Rank Variation by Year**

Year	Change in Standardized Rank		
	Observed Value	Mean	Standard Deviation
2008	100	1.0000	36.6093
2009	84	−0.9133	34.0343
2010	94	−0.0163	31.9876
2011	99	−1.2206	30.2410
2012	109	−1.0164	32.8802
2013	110	−1.3552	30.9282
Total	596	−0.6023	32.6696
Year	Absolute Value of Standardized Rank Variation		
	Observed Value	Mean	Standard Deviation
2008	100	29.5329	21.4533
2009	84	26.7548	20.8504
2010	94	25.4783	19.1595
2011	99	22.8428	19.7207
2012	109	24.3321	22.0140
2013	110	24.4521	18.8419
Total	596	25.5017	20.4019

Source: Data compiled by the authors based on internal documents of the Korea Institute of Public Finance

[Figure V-1] shows the year-to-year distributions of the standardized rank variation. As can be seen below, distributions before 2012 slightly lean to one side and the shapes of the distributions are relatively uneven. On the other hand, recent trends (2012 and 2013) show that the distribution of rank variations has become more even and widespread like a normal distribution.

[Figure V-1] **Distribution of Standardized Rank Variation by Year**



Source: Data compiled by the authors based on internal documents of the Korea Institute of Public Finance

It is natural that rank variation converges to 0 on average in the results above. However, the distribution of rankings is fairly even and widespread, and the extent of variation is above 20% on average. This leaves us to wonder whether the currently observed rank variation means a significant change. Of course, one may interpret that an active change in rankings manifests intense competition among public institutions to improve their management performance in reality. Nevertheless, the relatively large variation in rankings still poses the question of whether the PIMPES properly evaluates the innovation of management activities, management capability or management performance of public institutions.

The management capability and performance of public institutions can be improved by short-term efforts. Basically, however, it is more likely that such improvement is determined by individual public institutions' unique and

inherent abilities, which would require long-term investment and enhancement efforts. For this reason, frequent year-to-year changes in rankings and the large extent of such variation may also indicate that the current PIMPES is designed to encourage public institutions to focus more on making short-term management improvement efforts and having them reflected in evaluation outcomes, rather than attending to the fundamental aspects of management activities.

This implies that evaluation results are likely to be easily influenced by non-essential factors that are not related to the level of efficiency or publicness of public institutions, as set out in ***Proposition 3***. In this respect, it suggests that the existing PIMPES needs an improvement in developing proper evaluation indicators.

One consistent complaint raised in every evaluation cycle is that the current system is designed in such way that particular public institutions are rated high every year. This suggests that high-ranking institutions remain to be so while low-ranking ones continue to stay where they are in terms of ranking under the existing PIMPES and that such tendency is quite strong as well. Of course, a certain degree of stability in rankings is a natural phenomenon. Yet, a significantly high level of such stability can be problematic since this can undermine public institutions' commitment to improving their management performance. For this reason, we will examine the degree of stability in rankings by calculating the transition probability matrix of the standardized ranks.

The transition probability matrix can be defined as the probability of transition from the state space of the previous term to the current state space. If there are an n number of states, the transition probability matrix becomes the $n \times n$ matrix. And the (i, j) element represents the probability of transition from the state i of the previous term to the state of the current term j.

To obtain the transition probability matrix of standardized ranks, we have first formulated a state space by defining the low-ranking group (LOW) as below 30%, the middle-ranking group (MIDDLE) as above 30% and below 70% and the high-ranking group (HIGH) as above 70%, based on the final scores of management evaluation. Then, the transition probability matrix of the standardized ranks has been calculated by tracing trajectories on which the individual ranking groups take from term t-1 to term t.

<Table V-3> shows the probability of transition of the standardized ranks by using the entire data. Each value means represents the probability of transition from a ranking group in the previous term to another ranking group in the next term. For example, the element of (LOW, MIDDLE) in <Table V-3> indicates that the probability of transitioning to the MIDDLE in term t is 9% when an institution belongs to the LOW in term t-1. Therefore, the diagonal elements of the transition probability matrix—i.e., LOW-> LOW; MIDDLE-> MIDDLE; and HIGH-> HIGH—represent the probabilities that the past state will continue as it is in the present. For this reason, they can serve as a benchmark for examining stability in rankings. Meanwhile, off-diagonal elements refer to the probabilities that the current state deviates from where it used to be in the past.

Based on this, it is observed that the probability of individual public institutions currently staying in the same ranking group as in the past is about 45%. This, in turn, means that the probability of individual institutions deviating from their past ranking groups is about 55%. Therefore, this suggests that even though we can observe some degree of stability in rankings among public institutions—i.e. high-ranking ones continue to be highly ranked while low-ranking ones continue to be stuck in their low-rankings—, the level of its intensity is not high because the deviation rate is very high. Based on objective data, therefore, it proves to be not so persuasive to argue that the PIMPES has been structured to give preferential treatment to certain institutions. An institution's efforts (or intrinsic competence)—which determine its management performance—get accumulated as its unique resources and, thus, take on the characteristics of continuity and stability. On the other hand, enhanced management performance owing to short-term efforts of improving management activities would be discontinuous and temporary, from which we can expect short-term rank variations. Therefore, a high rate of deviation from the past ranking groups can provide a reasonable ground to argue that the existing PIMPES indicators have been designed to assign a relatively high weight on short-term management activities. And this is far from the standards of desirable evaluation indicators suggested by ***Proposition 3***.

〈Table V-3〉 **Transition Probability Matrix of Standardized Ranks**

Term t Term (t−1)	LOW	MIDDLE	HIGH
LOW	0.16	0.11	0.06
MIDDLE	0.09	0.13	0.12
HIGH	0.07	0.11	0.16

Note: The total number of observations is 596.
Source: Data compiled by the authors based on internal documents of the Korea Institute of Public Finance

B. Basic Statistical Analysis of Key PIMPES Indicators

Since the total score and ranking, which determine the final grade of a public institution, are relative concepts, comparison is still possible through standardization despite year-to-year changes. As for key PIMPES indicators or evaluation items, however, it is difficult to conduct a meaningful comparative analysis; when evaluation criteria or sub-items are changed, they become entirely different indicators.

As a matter of fact, the standards of management evaluation went through two comprehensive reformation during the period of analysis for this study—from the years 2008 to 2013. Such changes are the main obstacle to testing the effectiveness of management evaluation by empirical analysis. In this regard, we will first examine the consistency of the PIMPES indicators.

As of the years from 2011 to 2013, key PIMPES indicators can largely be divided into metric and non-metric indicators. There are three main categories —leadership and responsible management (hereinafter leadership); management efficiency; and major projects—and these three are composed of six indicators in total. Main evaluation items under these categories are listed in <Table V-4>. Besides, there exist sub-items for each of the main evaluation indicators. As a whole, the PIMPES is composed of very diverse and complicated evaluation indicators in its overall structure.[7)]

7) For details on these sub-items for performance evaluation, see Ministry of Strategy and Finance, 『Public institutions Administration Manual』, 2013.

〈Table V-4〉 **Management Evaluation Category and Indicator Structure**

Category	Key Evaluation Indicators
Leadership and Responsible Management	– Leadership – Responsible Management – Public Evaluation – Social Contribution (Social Responsibility and Government–Recommended Policy)
Management Efficiency	• Work Efficiency • Organization and Human Resources Management • Financial Budget Management and Performance (Financial Budget Management, Financial Budget Performance, Metric Management Operational Costs) * As for Pension–Fund–Type Institutions: Fund Operation Management and Performance • Wage and Performance Management (Wage and Performance • Labor Relations Management
Major Projects	• Comprehensive evaluation of plans, activities and performances by major businesses of public institutions

Source: Young–jae La and Tae–beom Yoon (2013), p. 61

In addition, these sub-items lack consistency in the sense that they go through vicissitudes whenever there is a change in the structure of the PIMPS and that weights and score formation are frequently changed as well. Therefore, this study will utilize for analysis six large evaluation categories that can be used as consistently as possible: metric leadership; metric management efficiency; metric main project; non-metric leadership; non-metric management efficiency; and non-metric main project indicators.

[Figure V-2] and [Figure V-3] are year-to-year box plots of metric and non-metric evaluation indicators respectively, which allow us to identify the distribution patterns of key evaluation indicators by year and check consistency.[8] The fact that we can distinctively see with these figures is that the main indicators show a very different shape from 2011.

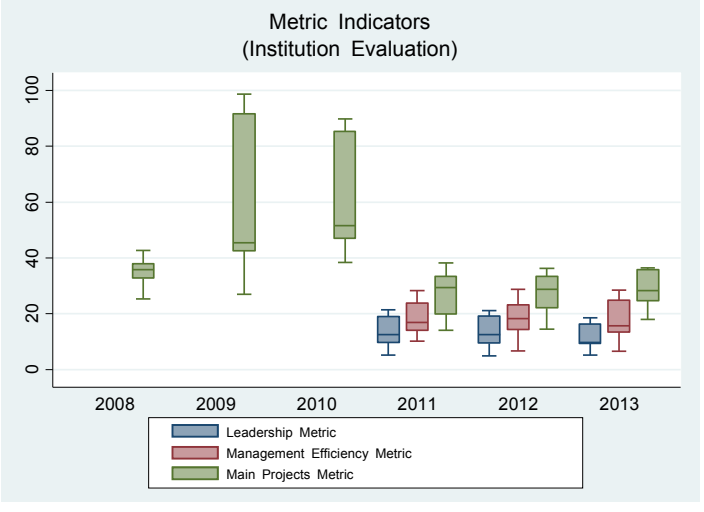
8) A box plot is useful for examining the year–to–year consistency of each evaluation indicator since it shows how measured values of data are distributed by using maximum value, minimum value, median value and quartile deviation within a single frame.

Prior to 2011, to begin with, metric leadership, metric management efficiency, and non-metric main project indicators were classified into other indicators, not into the six large categories mentioned above. Second, the distributions of individual indicators for the two periods (i.e. years from 2008 to 2010 and years from 2011 to 2013) are quite similar to one another. In particular, the distributions of individual indicators in the period of 2011-2013 are very similar in terms of scope and level, and seem to be very consistent year by year. In the 2008-2010 period, all evaluation indicators seem to show a fairly similar distribution pattern except for the metric main project indicator in 2008.

Since individual indicators are distributed differently year by year, it is not reasonable to forcefully apply continuity to individual indicators in order to utilize them for empirical analysis. To be more specific, there are two aspects to consider. First, as mentioned above, the management evaluation indicators have been segmented in terms of weight assignment and score formation since 2011, and sub-items are not perfectly matched. Therefore, if they are forcefully connected, the quality and composition of information contained in individual indicators become heterogeneous.

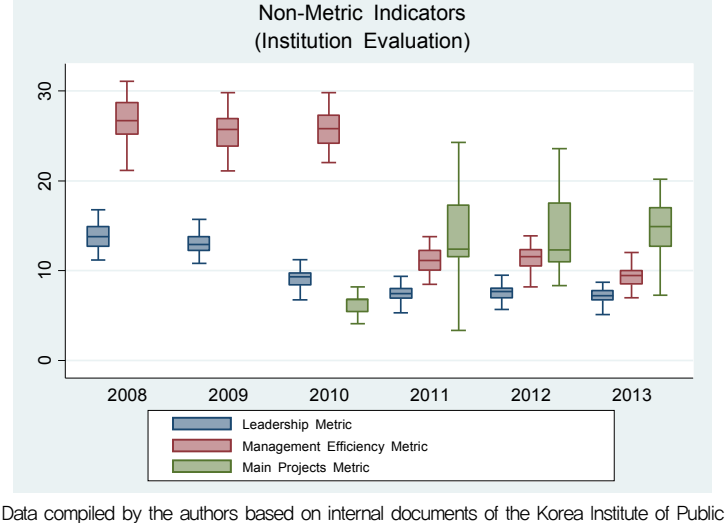
Second, even if individual evaluation indicators can be linked while maintaining the quality of respective information therein, applying continuity leads to some degree of inevitable loss of information contained in the indicators since the number of evaluation indicators is not the same between the two periods as shown above. In order to develop a consistent formation of evaluation indicators with the least amount of information loss, we define the two periods of analysis—i.e. period 1 (2008-2010) and period 2 (2011-2013)—to conduct an empirical analysis of the PIMPES' effectiveness.

[Figure V-2] **Distribution of Metric Management Evaluation Indicators by Year**



Source: Data compiled by the authors based on internal documents of the Korea Institute of Public Finance

[Figure V-3] **Distribution of Non-Metric Management Evaluation Indicators by Year**



Source: Data compiled by the authors based on internal documents of the Korea Institute of Public Finance

2 Empirical Analysis for Proposition Testing

A. Analysis of Determinants of Management Evaluation (Verification of Proposition 4)

First of all, regression analysis has been conducted with regard to the effect of individual evaluation indicators on the standardized rank so as to identify important determinants that affect evaluation outcomes. Of course, it is easily expected that the influence of the evaluation indicators on rankings among public institutions would be considerably high in general since the rankings are determined by total scores tallied after weights are assigned to indicators. Nevertheless, it is necessary to examine the extent to which the effect of evaluation indicators on the ranking varies indicator by indicator.

In this process, we can test a hypothesis about ***Proposition 4*** by checking the different levels of individual indicators' importance in affecting the results of the management evaluation. In other words, the appropriateness of individual indicators can be assessed by comparing differentials in such importance between indicators that are difficult to measure quantitatively—i.e. the leadership and responsible management items or non-metric indicators—and the rest.

<Table V-8> shows the results of regression analysis concerning the effect of individual evaluation indicators on the standardized rank. Each column represents the result of regression analysis for the periods 1 and 2 respectively, which uses the standardized reverse ranking within each public institution type as a dependent variable. In addition, our regression analysis has included the type of public institution and time-fix effect by year in order to control for the heterogeneous effect of public institution types and the specificity of each year's management evaluation. It should be noted For the sake of convenience of interpretation should it be noted that we have employed the standardized reverse ranking, instead of the standardized rank used above.[9] Therefore, a

9) The value of the standardized reverse ranking ranges from 0 to 100, as in the case of the standardized rank. However, the value increases as the ranking goes up; for instance, the value of the first rank is 100. Therefore, the results of regression analysis of the standardized reverse ranking only show changes in a coefficient's sign.

positive coefficient for each evaluation indicator means that the indicator contributes to a rise in ranking.

〈Table V-5〉 **Analysis of Determinants of Management Evaluation Ranking**

Variable	Dependent Variable: Standardized Reverse Ranking (within Each Public Institution Type)	
	Period 1: 2008–2010	Period 2: 2011–2013
Leadership (Metric)		4.0567*** (0.9937)
Management Efficiency (Metric)		4.6201*** (0.5384)
Major Projects (Metric)	4.9797*** (0.2552)	3.1667*** (0.5627)
Leadership (Non–Metric)	7.3128*** (1.0788)	4.7511** (2.1322)
Management Efficiency (Non–Metric)	5.3011*** (0.6889)	7.3045*** (1.5143)
Major Projects (Non–Metric)		5.5456*** (0.9085)
Dummy for Public Institution Type	YES	YES
Time–Fix Effect by Year	YES	YES
R^2	0.71	0.85
Observed Value	184	186

Notes: 1. Values in parentheses refer to standard errors.
2. *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.
Source: Data compiled by the authors based on internal documents of the Korea Institute of Public Finance

<Table V-5> allows us to derive two results as below.

First, in the period 1, the leadership indicator has the greatest effect on ranking; its importance is relatively high in the period 2, too. This means that evaluation outcomes under the current PIMPES tend to be swayed by leadership and responsible management indicators, rather than achievements in major projects carried out by public institutions or improved management.

Secondly, we find that the coefficient of non-metric indicators is larger than the coefficient of metric indicators, which means that the non-metric indicators

have a relatively large effect on the results of management evaluation compared to metric indicators. Under the current PIMPES, therefore, evaluation outcomes seem to be more dependent on qualitative data than quantitative data.

These results consistently indicate that evaluation indicators that are difficult to measure quantitatively and, thus, have large measurement error—i.e. leadership and responsible management indicators or non-metric indicators—are functioning as more crucial factors in evaluation results. All in all, they support a conclusion that is contrary to ***Proposition 4***.

The large impact of indicators with large measurement error implies that there is a possibility that evaluation outcomes are highly dependent on a sheer luck. In this sense, it is necessary to overhaul the design of evaluation indicators under the current PIMPES.

〈Table V-6〉 **Analysis of Determinants of Management Evaluation Ranking by Year**

Variable	Dependent Variable: Standardized Reverse Ranking (within Type)					
	Period 1: 2008–2010			Period 2: 2011–2013		
	2008년	2009년	2010년	2011년	2012년	2013년
Leadership (Metric)				4.5485** (1.7612)	2.9464 (1.9738)	5.7552*** (1.6518)
Management Efficiency (Metric)				3.9812*** (1.2396)	3.2702*** (1.0441)	6.0098*** (0.8834)
Major Projects (Metric)	5.3070*** (0.3784)	3.9450*** (0.4219)	6.9745*** (0.5368)	1.6888 (1.1961)	1.9694 (1.1947)	6.0892*** (0.8037)
Leadership (Non-Metric)	6.8426*** (1.4723)	9.3429*** (2.6518)	7.2693*** (1.6831)	2.5510 (4.2064)	11.5049* (5.8138)	6.4592** (2.6284)
Management Efficiency (Non-Metric)	4.0082*** (1.0198)	5.9066*** (1.6274)	6.3863*** (0.9646)	9.3750*** (2.8362)	3.0693 (3.0619)	5.9627** (2.2482)
Major Projects (Non-Metric)				7.4014*** (2.1865)	4.0939 (3.2652)	5.6428*** (1.3035)
Dummy for Public Institution Type	YES	YES	YES	YES	YES	YES
R^2	0.87	0.86	0.91	0.69	0.72	0.86
Observed Value	75	56	55	62	60	62

Notes: 1. Values in parentheses refer to standard errors.
2. *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.
Source: Data compiled by the authors based on internal documents of the Korea Institute of Public Finance

To test this possibility, we have estimated whether the importance of an individual indicator on evaluation outcomes changes by year. Specifically, this has been be done by conducting the same kind of analysis as in Table V-5. The results of the analysis are presented in <Table V-6>. Our estimation reveals that the leadership (non-metric) indicator is consistently the most important determinant of evaluation outcomes, which is similar to what the foregoing findings have demonstrated. On the other hand, we can observe ups and downs in the relative importance of the other evaluation indicators year by year, which indicates that these indicators lack year-to-year consistency.

B. Empirical Analysis of Proposition based on Financial Information

As discussed earlier, this study considers the PIMPES to be an incentive contract, and locates, from an economic perspective, the important purpose of the PIMPES and its legitimacy in the role of controlling moral hazard within public institutions. Unlike the private sector, there is a high possibility of moral hazard because public institutions are not subject to the regulatory mechanism of of bankruptcy owing to the government's (implicit) guarantee. In this economic environment, the PIMPES can improve management efficiency of public institutions and, thus, play the role of overseeing moral hazard among them.

We have already presented various propositions on the role of the PIMPES as an institutional measure of control over moral hazard in Chapter IV. In this section, we conduct an empirical testing of each proposition based on financial information about public institutions and other sources related to evaluation outcomes.

Key financial information about public institutions was collected through Alio (www.alio.go.kr) so as to link them to management performance evaluation outcomes. As a result, a total of 112 public institutions have been included in our analysis, as listed in <Table V-7>. Since Alio discloses financial information for the last five years only, we have set the analysis period at the period 2 (2011-2013) within the available range. Using the data thus constituted, the basic statistics of the major financial variables used in the empirical analysis are presented in <Table V-8>

〈Table V-7〉 **Public Institutions subject to Analysis by Type**

Institution Type	Name
Market Type	Busan Port Authority, Incheon International Airport Corporation, Korea Gas Corporation, Korea Airports Corporation, Korea Resources Corporation, Korea South-East Power Co., Ltd., Korea Southern Power Co., Ltd., Korea East-West Power Co., Ltd., Korea Western Power Co., Ltd., Korea National Oil Corporation, Korea Hydro and Nuclear Power Co., Ltd., Korea Electric Power Corporation, Korea Midland Power Co., Ltd, Korea District Heating Corporation
Quasi-Market Type	Korea Coal Corporation, Korea Housing and Urban Guarantee Corporation ., Yeosu Gwangyang Port Authority, Ulsan Port Authority, Incheon Port Authority, Jeju Free International City Development Center, Korea Appraisal Board, Korea Tourism Organization, Korea Expressway Corporation, Korea Racing Authority, Korea Broadcast Advertising Corporation, K-Water, Korea Minting, Security Printing and ID Card Operating Corporation, Korea Railroad Corporation, Korea Land and Housing Corporation (LH Corporation), Korea Marine Environment Management Corporation
Commissioned-Service Type	Health Insurance Review and Assessment Service, Korea Transportation Safety Authority, Korea National Park, National Health Insurance Service, Korea International Broadcasting Foundation, Korea Agency for Infrastructure Technology Advancement, Agriculture, Korea Agency of Education, Promotion and Information Service in Food, Agriculture, Forestry and Fisheries, Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries, Foundation of Agricultural Technology Commercialization and Transfer, Agricultural Technology Application Foundation, Korea Trade-Investment Promotion Agency, Korea Land and Geospatial InformatiX Corporation, Korea Road Traffic Authority, Independence Hall of Korea, Korea Ship Safety Technology Authority, Small Enterprise and Market Service, Korea Energy Agency, Postal Savings and Insurance Development, Korea Postal Logistics Agency, National IT Industry Promotion Agency, Korea Technology and Information Promotion Agency for Small- and Medium-Sized Enterprises, Korea Livestock Products HACCP Accreditation Service, Korea Institute for Animal Products Quality Evaluation, Korea Gas Safety Corporation, Korea Employment Information Service, Korea Foundation for the Advancement of Science and Creativity, Mine Reclamation Corporation, Korea Education and Research Information Service, Korea Meteorological Industry Promotion Agency, Korea Labor Force Development Institute for the Aged, Korea Agro-Fisheries and Food Trade Corporation, Korea Rural Community Corporation, Korea Institute of Design Promotion, Korea Communications Agency, Korea Human Resources Development Institute for Health and Welfare, Korea Health and Welfare Information Service, Korea Health Industry Development Institute, Korea Veterans Welfare and Healthcare Corporation, Korea Institute for Advancement of Technology, Korea Evaluation Institute of Industrial Technology, Korea Industrial Complex Corporation, Korea Occupational Safety and Health Agency, Human Resources Development Service of Korea, Korea Petroleum Quality and Distribution Authority, Korea Institute of Ceramic Engineering and Technology, Korea Fire Institute, Korea Consumer Agency, Korea Fisheries Resources Agency, Korea Infrastructure Safety and Technology Corporation, Korea Institute of Energy Technology Evaluation and Planning, National Research Foundation of Korea, Korea Securities Depository, Korea Postal Service Agency, Korea Institute of Nuclear Safety, Korea Internet and Security Agency, Korea Forestry Promotion Institute, Korea Employment Agency for the Disabled, Korea Student Aid Foundation, Korea Electrical Safety Corporation, Korea Power Exchange, National Information Society Agency, Korea Rail Network Authority, Korea Youth Counseling and Welfare Institute, Korea Youth Work Agency, Korea Creative Content Agency, Korea Institute of Maritime and Fisheries Technology, Korea Environment Corporation, Korea Environmental Industry and Technology Institute
Fund-Management Type	Government Employees Pension Service, National Pension Service, Korea Workers' Compensation and Welfare Service, Korea Technology Finance Corporation, Teachers' Pension, Korea Credit Guarantee Fund, Korean Film Council, Korea Deposit Insurance Corporation, Small and Medium Business Corporation, Korea Trade Insurance Corporation, Arts Council Korea, Korea Press Foundation, Korea Radioactive Waste Agency, Korea Asset Management Corporation, Korea Housing-Finance Corporation
Total	12 Public Institutions

Source: Data compiled by the authors based on internal documents of the Korea Institute of Public Finance

〈Table V-8〉 **Basic Statistics of Major Financial Information**

Variables	Observed Value	Mean	Standard Deviation	Minimum Value	Maximum Value
Debt-to-Equity Ratio	297	14.8676	162.3401	-530.559	2421.092
Return on Assets (ROA)	289	0.021806	0.109946	-0.90125	0.761841
Assets (KRW 100 million)	297	242.8653	4059.902	0.001204	69973.19
Average Wage per Regular Employee (KRW million)	318	64.91797	12.37033	35.894	101.021
Welfare Benefits per Regular Employee (KRW million)	318	1.666072	1.939957	0.015	15.442
Number of Full-Time, Regular Employees (thousand)	318	1.563577	3.720349	0.0498	29.732
Average Length of Continuous Service among Regular Employees (years)	318	12.97264	4.553116	0.81	22.79
Age of Public Institution (years)	320	23.38125	20.05349	0	105
Dummy for the Past Career of a Public Institution's Head					
Politician	320	0.04375	0.204859	0	1
Government Official	320	0.528125	0.49999	0	1
Dummy for Public Institution Type					
Fund-Management Type	320	0.134375	0.341589	0	1
Market Type	320	0.13125	0.338202	0	1
Commissioned-Service Type	320	0.596875	0.4912937	0	1
Quasi-Market Type	320	0.1375	0.344914	0	1

Source: Financial data was compiled by the authors using the Alio data; Data on the past career of each public institution's head was compiled by the authors based on internal documents of the Korea Institute of Public Finance.

Using financial information we aim to zoom in on, the empirical analysis presented in this study is centered on two main themes: 1) the effect of public institutions' key financial variables on management evaluation results (i.e. the standardized reverse ranking and evaluation grades); and 2) the effect of evaluation results of the previous term on public institutions' major financial variables.

Above all, <Table V-9> estimates the effect of major financial variables on management evaluation results. In the columns (1) and (3), the standardized reverse ranking is used as a dependent variable; and evaluation grade is used as a dependent variable in the columns (2) and (4). As for the latter, a five-grade scale has been used: from an A (including an S) to an E.[10] Given that our dependent variables take on the character of a latent variable, an ordered logit model has been applied. To demonstrate the robustness of the results of the empirical analysis, we have excluded the fund-management type from our samples in the last two columns and executed the same analysis as in the first two columns.

Financial information about fund-management type public institutions is likely to be overestimated due to the amount of fund they manage. There is another concern about the possibility of distortion in their financial information since this type of public institutions performs practically no production activities.

For this reason, fund-management type public institutions have been excluded from empirical analysis in previous studies as well (e.g. Se-jeong Ha 2014). In this study, we expect that the distortion of financial information would be minimized because only financial information about those institutions' own accounts have been taken into consideration—that is, we have excluded fund accounts of the fund-management type institutions. Given the possible problem of information distortion to be incurred by including these institutions, however, we present an additional empirical analysis using samples from which fund-management-type public institutions have been left out.

<Table V-10> and <Table V-11> show that the estimated results concerning the effect of evaluation outcomes of the previous term (i.e. the standardized ranking of the previous evaluation term or a dummy for evaluation grades) on public institutions' key financial variables (i.e. debt-to-equity ratio, ROA, welfare benefits for regular employees and the average wage of regular employees) by using the entire samples on the one hand and by excluding fund-management-type

10) During the period of analysis for this study, there is only one observation in which grade S was given. In this respect, it seems reasonable to include grade S in grade A.

institutions on the other. It should be noted that a simultaneity bias might occur due to the problem of endogeneity since individual components of financial information affect one another and are generated at the same time. Considering the problem of endogeneity, only the variables considered to be exogenous enough have been included as independent variables in the analysis presented here.

Now let us examine the results of this empirical analysis based on propositions presented in Chapter IV. For a more efficient discussion, we present those propositions again in the below. Through an empirical analysis based on financial information, the propositions will be used to assess the appropriateness of the PIMPES design (***Proposition 4*** excluded.)

Propositions

1. *A public institution with a high debt ratio will be disadvantaged on management evaluation.*
2. *In the case of public corporations whose evaluation scores are poor in this period, their debt ratios will be improved in the following period.*

2-1. *public institution that spends excessively on welfare benefits will be disadvantaged on management evaluation.*

3. *Management evaluation outcomes should not be affected by factors that are irrelevant to efforts to promote the publicness or efficiency of a given public institution .*

〈Table V-9〉 **Effect of Public Institution's Financial Information on Management Evaluation Results**

	Entire Samples		Samples with Fund–Management Type Excluded	
	Standardized Rank	Evaluation Grade	Standardized Rank	Evaluation Grade
	(1)	(2)	(3)	(4)
Debt–to–Equity Ratio	0.0019	–0.0001	0.0084	0.0002
	(0.0102)	(0.0006)	(0.0107)	(0.0006)
Return on Assets (ROA)	31.6870	1.6931	46.7246	3.2808
	(15.7616)**	(1.2237)	(18.4782)**	(1.3632)**
Assets (KRW trillion)	0.6858	0.0099	0.6996	0.0105
	(0.4075)*	(0.0237)	(0.4079)*	(0.0238)
Age of Public Institution	0.1246	0.0100	0.1328	0.0107
	(0.0932)	(0.0059)*	(0.0961)	(0.0061)*
Number of Full–Time, Regular Employees (thousand)	–0.4721	0.0079	–0.4569	0.0087
	(0.4946)	(0.0309)	(0.4978)	(0.0309)
Welfare Benefits per Regular Employee (KRW million)	0.0426	–0.0315	–0.0963	–0.0501
	(0.8603)	(0.0564)	(0.8689)	(0.0570)
Average Wage per Regular Employee (KRW million)	0.4823	0.0423	0.4311	0.0377
	(0.1608)***	(0.0116)***	(0.1691)**	(0.0121)***
Average Length of Continuous Service among Regular Employees	1.9999	0.1891	2.3176	0.2071
	(1.5088)	(0.1086)*	(1.5147)	(0.1088)*
(Average Length of Continuous Service among Regular Employees)2	–0.1201	–0.0112	–0.1318	–0.0119
	(0.0650)*	(0.0046)**	(0.0654)**	(0.0046)**
Past Career of a Public Institution's Head: Politician	–3.7461	–0.3269	–3.9216	–0.3396
	(9.1066)	(0.5678)	(9.1282)	(0.5685)
Past Career of a Public Institution's Head: Government Official	9.2688	0.4619	8.1155	0.3927
	(3.6205)**	(0.2371)*	(3.7197)**	(0.2435)
Dummy for Public Institution Type	YES	YES	YES	YES
Time–Fix Effect by Year	YES	YES	YES	YES
R^2 / LR χ^2	0.11	54.25	0.11	49.28
N	289	289	269	269

Notes: 1. Values in parentheses refer to standard errors.
2. *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.
Source: Data compiled by the authors

〈Table V-10〉 **Effects of Management Evaluation Results on Major Financial Information (Entire Samples)**

	Debt-to-Equity Ratio		ROA		Welfare Benefits per Regular Employee		Average Wage per Regular Employee	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Standardized Rank $_{t-1}$	0.0131		0.0005**		0.0014		0.0973***	
	(0.2115)		(0.0003)		(0.0044)		(0.0259)	
Dummy for Grade A_{t-1}		−7.2873		0.0446		−0.0875		15.4277***
		(33.5689)		(0.0398)		(0.7206)		(4.2058)
Dummy for Grade B_{t-1}		2.2850		0.0194		−0.3233		9.7024**
		(31.6113)		(0.0373)		(0.6790)		(3.9630)
Dummy for Grade C_{t-1}		−5.1297		0.0060		−0.0440		9.2905**
		(32.0488)		(0.0379)		(0.6876)		(4.0132)
Dummy for Grade D_{t-1}		−9.6596		−0.0028		−0.5352		3.5441
		(36.5716)		(0.0434)		(0.7810)		(4.5582)
Age of Public Institution	−0.3464	−0.3601	0.0002	0.0002	0.0107*	0.0112*	−0.0151	−0.0072
	(0.3059)	(0.3094)	(0.0004)	(0.0004)	(0.0065)	(0.0065)	(0.0385)	(0.0382)
Past Career of a Public Institution's Head: Politician	−6.6384	−3.9366	−0.0323	−0.0348	−0.0291	−0.0473	−1.9269	−2.5420
	(35.8361)	(36.2797)	(0.0422)	(0.0428)	(0.6686)	(0.6727)	(3.9717)	(3.9262)
Past Career of a Public Institution's Head: Government Official	11.1788	9.6762	−0.0035	−0.0007	0.0962	0.1121	−3.9099**	−3.3445**
	(12.6913)	(13.0376)	(0.0151)	(0.0156)	(0.2636)	(0.2697)	(1.5656)	(1.5742)
Dummy for Public Institution Type	YES	YES	YES	YES	YES	YES	YES	YES
Time-Fix Effect by Year	YES	YES	YES	YES	YES	YES	YES	YES
R^2	0.01	0.02	0.06	0.06	0.03	0.04	0.29	0.32
N	191	191	186	186	207	207	207	207

Notes: 1. Values in parentheses refer to standard errors.
2. *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.
3. The base grade of dummy for management evaluation grade is grade E
Source: Data compiled by the authors

〈Table V-11〉 **Effects of Management Evaluation Results on Major Financial Information (Samples Excluding Fund Management Type)**

	Debt-to-Equity Ratio		ROA		Welfare Benefits per Regular Employee		Average Wage per Regular Employee	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Standardized Rank $_{t-1}$	0.0015		0.0005		0.0014		0.0881	
	(0.2226)		(0.0003)*		(0.0048)		(0.0281)***	
Dummy for Grade A$_{t-1}$		−9.1824		0.0500		−0.1746		14.3889
		(34.9607)		(0.0407)		(0.7614)		(4.3217)***
Dummy for Grade B$_{t-1}$		2.3762		0.0150		−0.2437		9.9379
		(32.7117)		(0.0379)		(0.7124)		(4.0437)**
Dummy for Grade C$_{t-1}$		−4.9988		0.0084		−0.0856		9.6619
		(33.1267)		(0.0384)		(0.7215)		(4.0950)**
Dummy for Grade D$_{t-1}$		−10.0149		−0.0024		−0.4948		3.9387
		(37.7553)		(0.0440)		(0.8223)		(4.6672)
Age of Public Institution	−0.3475	−0.3567	0.0002	0.0002	0.0113	0.0114	−0.0213	−0.0153
	(0.3202)	(0.3239)	(0.0004)	(0.0004)	(0.0070)	(0.0071)	(0.0404)	(0.0400)
Past Career of a Public Institution's Head: Politician	−6.6321	−3.9262	−0.0325	−0.0355	−0.3174	−0.3367	−4.7126	−5.4401
	(36.9936)	(37.4575)	(0.0428)	(0.0434)	(0.8059)	(0.8158)	(4.6622)	(4.6304)
Past Career of a Public Institution's Head: Government Official	12.0777	10.0516	−0.0011	0.0036	0.2092	0.2070	−3.7497	−3.2248
	(13.3794)	(13.8267)	(0.0157)	(0.0163)	(0.2915)	(0.3011)	(1.6862)**	(1.7092)*
Dummy for Public Institution Type	YES	YES	YES	YES	YES	YES	YES	YES
Time-Fix Effect by Year	YES	YES	YES	YES	YES	YES	YES	YES
R^2	0.01	0.02	0.04	0.05	0.04	0.04	0.29	0.32
N	179	179	174	174	179	179	179	179

Notes: 1. Values in parentheses refer to standard errors.
2. *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.
3. The base grade of dummy for management evaluation grade is grade E
Source: Data compiled by the authors

The first three propositions evaluate whether the PIMPES can function as a control device to address moral hazard, and the fourth proposition assesses whether its evaluation indicators are appropriate.

First of all, we can test ***Proposition 1*** and ***Proposition 2*** by looking at the relationship between management performance evaluation results and debt ratio. <Table V-9> shows that none of the coefficients of the debt-to-equity ratio are statistically significant, indicating that management performance rankings or evaluation grades are not affected by debt ratio. In addition, the columns (1) and (2) of <Table V-10> and <Table V-11> suggest that the standardized ranking or evaluation grade of the previous year does not affect a change in the current debt-to-equity ratio.

Based on these findings, we can confirm that the design of the PIMPES has nothing to do with hypotheses proposed in ***Proposition 1*** and ***Proposition 2***. That is, the problem of moral hazard arising from public institutions' risk-taking behaviors is not adequately reflected in management evaluation, which suggests that the existing PIMPES are not properly functioning as a control mechanism to regulate possible moral hazard, either.

There are several reasons why the debt ratio is not reflected in the PIMPES as a discipline mechanism concerning moral hazard. Among others, the most important reason is that changes in the amount of public institutions' debts tend to be influenced by external factors.

In 2011, the Ministry of Strategy and Finance announced that the LH Corporation's borrowings from the National Housing Fund and the K-Water's debts incurred by the Four-River project and the Ara Waterway Development project constituted 'liabilities arising from reasons that public institutions cannot control in the process of conducting their business, such as government-policy implementation.' As such, these debts were left out from the evaluation of the two public institution's debt ratios (Board of Audit and Inspection of Korea 2013). The adjustment in evaluation standards illustrated above suggests that debt ratio does not provide meaningful information to represent the financial soundness of public institutions for management evaluation.

To rephrase this in the context of the model described above, debt ratio reflects much more noise caused by measurement error than other indicators while failing to provide sufficient information about management performance.

When debt ratio is determined by external factors as shown above, it becomes very difficult to identify signs of moral hazard through the debt ratio. In this respect, it seems very natural that our analysis reveals no relevance between debt ratio and evaluation outcomes.

On the other hand, ROA (net income/assets)—which is a representative variable for profitability—shows a statistically significant positive effect (<Table V-9>) on management evaluation rankings. Also, the average effect of management evaluation results of the previous year on ROA (<Table V-10> and the column (3) of <Table V-11>) is statistically significant to some extent. These results imply that the current PIMPES reflects some degree of information about the profitability of public institutions. For this reason, it is possible to say that public institutions are trying to improve their evaluation results by increasing profitability.

Now, let us turn to ***Proposition 2-1*** by testing the relationship between welfare benefits—which are considered to be another indicator of moral hazard in this study—and management evaluation results. As mentioned earlier in Chapter IV, public institutions have incentives to increase all components of their expenditure on consumption above the optimal level if the risk of bankruptcy disappears as a result of the government's (implicit) guarantee. If moral hazard of public institutions is widespread, therefore, we can expect that spending on labor costs—especially on welfare benefits—will be increased. However, if management evaluation is designed in such a way that it can regulate such moral hazard, it is expected that expenditure on welfare- benefits will show a negative relationship with evaluation results.

<Table V-9> to <Table V-11> show that there is no relevance between welfare benefits expenditure and management evaluation, as in the case of debt ratio. In other words, welfare benefits do not affect evaluation results, and the the evaluation results of the previous year do not influence the current spending on welfare benefits. To sum up, evaluation concerning welfare benefits expenditure does not contain any significant information about the moral hazard of public institutions under the current PIMPES.

In this study, we have maintained that the average wage (or annual salary) per regular employee is not an adequate indicator for assessing the level of moral hazard. Wage is similar to welfare benefits in the sense that both of them

are benefits to employees. However, wage takes on a different character since the average wage reflects a given employee's work performance or productivity while welfare benefits do not. For this reason—yet, unlike our previous discussion—, it seems reasonable to examine the effect of average wage per employee on management evaluation outcomes through the lens of employee's productivity.

The results of <Table V-9> indicate that the average amount of wage per employee has a highly significant positive effect on evaluation outcomes, which seems to be very natural if we assume that an employee's abilities are reflected in the average wage. In addition, the last two columns of <Table V-10> and <Table V-11> show that the PIMPES ranking or grade of the previous year has a statistically very significant positive effect on the amount of the current wage. This is another natural phenomenon because management evaluation results get reflected in an employee's performance-based pay.

Finally, let us examine the results of empirical analysis to test ***Proposition 3***. The first two propositions have been analyzed from a perspective that the PIMPES is to serve as a measure of control over moral hazard. With the ***Proposition 3***, we will discuss the adequacy of PIMPES indicators themselves.

Since the goal of the PIMPES lies in promoting the management performance of public institutions by evaluating their management efficiency, evaluation outcomes have to sufficiently mirror the intrinsic value or efforts of a given institution. Therefore, it is more desirable that financial information determines evaluation results since the information represents an institution's management efficiency. If the results of management evaluation are determined by factors irrelevant to efforts or intrinsic values—especially human and material resources invested for evaluation preparation—, however, public institutions will lower the level of their efforts required for improving management efficiency.

There are many external factors that are independent of the internal value of a public institution. Nevertheless, they affect the results of management evaluation. For example, as noted in previous studies, public institutions use significant amounts of human and material resources to prepare for management evaluation; so the larger the institution, the more likely it is to receive better results in the sense that they are more capable of mobilizing personnel dedicated to handling evaluation preparation than smaller ones. Also, there is a possibility

that older institutions can better cope with management evaluation because it must have accumulated knowledge about how to deal with the PIMPS, compared to newly created institutions. In this respect, we have conducted an empirical analysis by taking into account institution size—represented by assets or the number of full-time employees—and age that are not directly related to management efficiency but can affect management evaluation results.[11)]

Let us first discuss what <Table V-9> reveals. Looking at the effect of assets and the number of regular employees on management evaluation results, we find that the size of assets positively affects management evaluation rankings although its significance level is not high (the columns (1) and (3)). On the other hand, the coefficient of the number of regular employees is not statistically significant, which leads us to reject the hypothesis that institutions are likely to receive better evaluation outcomes as the number of regular employees increases. Another external factor (i.e. the age of an institution) shows a positive relationship with evaluation grade although the level of significance is not high (the column (2) and the column (4)). In other words, older public institutions are likely to get better evaluation results.

The results of our empirical analysis are not straightforward. Yet, they still supports the fact that management evaluation results are affected to some extent by an institution's intrinsic values or factors that are not related to effort—which is contrary to ***Proposition 3***. In other words, the larger and older a public institution is, the better their management evaluation results are because such institution is better positioned to cope well with frequently changing PIMPES standards.

In addition, <Table V-9> considers other various external variables or factors

11) Non-essential factors-i.e. ones that independent of the internal value of an institution-affecting evaluation outcomes tend to be temporary and discontinuous because they are irrelevant to the institution's capabilities. For this reason, it is expected that non-essential factors will greatly affect the variability of of evaluation grade. On the other hand, the age and size of an institution are expected to have less influence on the extent of a rise in grade variability. This is because their effects are continuous and stable even though they are non-essential factors. In this respect, it seems reasonable to conclude that the age and size factors and the rest of the non-essential factors have discrete effects while the the former's effect on management evaluation and the latter's effect on the variability of evaluation grade trade off each other in appearance. We find it valid to classify the former as external factors that have a continuous effect on evaluation outcomes; and the latter as external factors that have a short-term and discontinuous effect on grade variability.

that may affect management evaluation results, the results of which are quite interesting.

As for the average length of continuous service by employees, we can see that its squared term has a negative coefficient at a considerably significant level. This indicates that the results of the management evaluation have an inverted U-shaped relationship with the average length of continuous service by employees. Based on the estimated coefficients, the relationship between the average length of continuous service and evaluation results transitions from positive (+) to negative (-) at around 8.3-8.7 years. This suggests that evaluation results tend to worsen when the employees of a given institution are either older or younger than other institutions.

In the case of the relationship between the past career of an institution's head and evaluation outcomes, the dummy variable representing the past career of an institution's head being a government official has a statistically significant positive effect on evaluation outcomes. In other words, management evaluation results tend to be better when the head of a given public institution used to be a government official. But we are very cautious about interpreting these results. On a positive note, evaluation results would be better because the past career of an institution's head being a government official can be related to relatively better management ability thanks to his or her past experience with the organizational culture. The opposite interpretation would be that former government officials are likely to land on public institutions that are expected to perform better since they are more versed in public institutions. It is not clear where the cause lies. However, the past career of an institution' head seems to be clearly related with management evaluation.

VI

Conclusion and Policy Implications

1 Summary

In this study, we have examined whether the current PIMPES is functioning well in accordance with the rationale for its existence by testing the following propositions.

1. An institution with a serious problem of moral hazard (e.g., an institution with a high debt ratio or a high level of welfare expenditure) should be disadvantaged on management evaluation.
2. In the case of public institutions whose management evaluation scores are poor in this period, more efforts (for example, an improved debt ratio or welfare benefits expenditure) should be made to address moral hazard in the next term.
3. Management evaluation outcomes should not be affected by factors that are irrelevant to efforts (e.g. the age or size of an institution) to promote the publicness or efficiency of a given public institution.
4. Management evaluation results should not be greatly affected by items that do not considered to contain accurate information about the institution's efforts to improve its efficiency or publicness (e.g. the evaluation category of leadership and responsible management.)

To verify the first proposition, we have operationalized the level of moral hazard as debt ratio and spending on welfare benefits. If the debt ratio of a public institution is high or the expenditure on welfare benefits is large, it is expected that there is a problem of moral hazard within the institution. Our analysis indicates that institutions with higher debt ratios do not receive poor scores on management evaluation. It also suggests that a high level of spending on welfare benefits does not lead to a disadvantage on evaluation. Neither is the second proposition supported by our analysis; poor evaluation outcomes in the present term do not lead to a decrease in debt ratio in the next term.

To verify the third proposition, the size of an institution has been operationalized as the size of its assets. The results of our analysis reveal that public institutions with larger assets tend to receive better evaluation outcomes compared to smaller institutions. As for the age factor, we have found that older institutions tend to receive better evaluation results. In other words, management evaluation outcomes are influenced by factors that are irrelevant to an institution's efforts made to fulfill the purpose of its establishment. With respect to the variability of management evaluation rankings, the probability of staying in the same ranking group—i.e. either high, middle or low—turns out to be only 45%; and the probability of deviating from where a given institution used to belong was amounts to 55%. In particular, the probability that an institution previously categorized as the high-ranking group fall down to the low-ranking group in the following year—or vice versa— was 13%. To put it differently, 13 out of 100 institutions will experience a significant change in their evaluation rankings in just one year.

The analysis of the fourth proposition indicates that the results of management evaluation are significantly influenced by leadership and responsibility management indicators. This means that management evaluation rankings are affected the most by indicators that do not provide accurate information about the efforts of public institutions. All in all, the results of our analysis show that the current PIMPES satisfies none of the four propositions concerning the optimal design of a management evaluation system.

2 Policy Implications and Directions for Institutional Improvement

In this section, we propose several policy directions for improving the existing PIMPES based on the results of our analysis above.

A. Suggestion on the Purpose of Operating the PIMPES

First, we need to keep the main objectives of the PIMPES as concise as possible. As mentioned earlier, it has been designed to be inefficient in the sense that it is difficult for public institutions to respond swiftly to changing business environments. This inefficiency inevitably impedes the role of the PIMPES as an incentive contract. Under the current system, a performance-based pay scheme cannot play the same role as stock option contracts signed with the top management in the labor market. To address this problem, we would need a comprehensive overhaul of the PIMPES. The PIMPES has been around for more than 30 years, which means reorganizing the system would take a lot of time and effort.

If the government wants to operate the PIMPES as it is now, then the purpose of operating the system should be clearly re-defined as a control mechanism concerning moral hazard among public institutions. For example, we would need to cull evaluation indicators down to several objective financial indicators—e.g. debt ratio or spending on welfare benefits—and to impose penalty on institutions according to the level of severity. The necessity of such regulation is fairly recognized given that the government would have to bear a considerable financial burden if it continues to neglect a rise in debt ratio or in operating costs like welfare benefits expenditure. In this case, an improvement in management efficiency can be achieved by addressing inefficiency arising from moral hazard.

The major advantage of taking this direction is that we can keep the existing framework of the current system and still reduce social costs by simplifying the evaluation indicators. It can also help public institutions better adapt to to the system by clarifying the goals they should achieve. In addition, what we propose here is in line with the government's current policy stance: i.e. strengthening evaluation criteria concerning debt management and reckless management.

B. Design Problem concerning PIMPES Indicators

Our second suggestion is related to the design of the evaluation indicators. If policymakers want to raise the level of efforts made by public institution to improve their performance through management evaluation, it is inevitable to overhaul the current system. According to contract theory, one of the most important prerequisites for an incentive contract to function effectively is that performance indicators must reflect changes in the agent's level of effort. If management evaluation results are affected by factors that are not related to efforts put in by public institutions, the role of the PIMPES as an incentive contract would be weakened.

In order for the PIMPES to serve as a mechanism to induce public institutions to put in more efforts to a socially desirable level, it is necessary to select proper performance indicators. This means that factors related to short-term responsiveness to management evaluation—e.g. human resources, the size of asset and the age of an institution—should not affect management evaluation results. To prevent this from happening, it is necessary to alleviate, among others, the burden of evaluation borne by institutions.

In this regard, the Public Institution Management Research Corporation (2015) sets out several ways to reduce such burden: to scale down management performance reports and provide a simple standard form; to conduct intensive investigations and hold joint meetings; and to simplify the structure of evaluation indicators and cut down on the number of indicators. Meanwhile, Young-jae La and Tae-beom Yoon (2013) suggest that the current evaluation cycle be readjusted to two to three years.

This study also proposes to decisively cut down on the number of evaluation indicators and simplify sub-items of evaluation indicators as much as possible. Let us take an example of the evaluation item of government-recommended policy. In practice, it functions not so much as an indicator to measure the degree of publicness achieved by institution concerned. Rather, it is more like a management control measure used to ensure that public institutions fulfill tasks delegated by certain ministries (Public Institution Management Research Corporation 2015). Indicators to measure the level of publicness should be limited in scope to focus on public institutions' performance in key management

activities that are directly related to individual institutions' unique and disparate purposes. Also, general policy issues concerning the management of public institutions should be excluded from PIMPES indicators.

In addition, it should be noted that evaluation indicators must provide undistorted information about an institution's level of effort. For example, Ji-young Kim (2011) points out that evaluation grades assigned to public corporations do not accurately reflect the level of actual productivity and changes thereof. Rather, it is observed that the direction of change reverses in many cases. The reason for this, argues Kim, consists in the fact that management evaluation indicators are based on target values that are pre-determined for individual corporations instead of the actual growth rates of productivity. The credibility of the PIMPES will fall apart when there are many more cases in which evaluation rankings rise despite decreasing productivity. If performance-based pay becomes more dependent on the results of management performance evaluation than on what financial indicators reveal, the general public would lean toward a perception that performance-based pay is portioned out by public institutions.

Some argue that the performance of public institutions should be adjusted by using "adjusted financial information" or performance figures because it is not fair that an institution's performance gets determined by risks that go beyond the institution's control. As a matter of fact, debts incurred in the course of implementing key government policies—e.g. the Four-River project and overseas resource development projects—were excluded from performance evaluation for the same reason.

From a perspective of optimal contract theory, however, such argument is problematic. Suppose that an institution receives performance-based pay *(s)* and that the performance pay is basically dependent on the institution's financial indicators *(s (q_2))*. Let us also add that the financial indicators observed *(q_2)* contain information about both the institution's efforts to improve efficiency *(a_2)* and the risk *(ε_2)* that the institution cannot control.

$$s = s(q_2)$$
$$q_2 = a_2 + \varepsilon_2,$$

Based on the formulae above, the use of "adjusted financial information" can be explained as follows. The level of uncontrollable risks being high means that q_2 should not be used as it is to determine the amount of performance-based pay due to the large value of ε_2. This is because pegging *(s)* at q_2 can cause a situation in which the level of effort (a_2^*) to be achieved gets lower compared to what is possible when the level of uncontrollable risks is not so high. The current PIMPES is practically using "adjusted financial information" $\tilde{q}_2$ instead of actual financial indicators (q_2). Yet, $\tilde{q}_2$ itself is created on an arbitrary basis; so $\tilde{q}_2$ cannot provide information necessary to deduce a_2, either. Then, what would be the solution to this problem? According to contract theory, performance -based pay $(s(\tilde{q}_2))$ should not be based on adjusted financial information. Rather, we need to reduce the proportion of performance-based pay in compensating public institutions.

C. Designing a Performance-Based Pay Scheme under the PIMPES

As previously noted, performance-based pay is not directly related to a given institution's efforts to improve its publicness or efficiency unlike such indicators as leadership and responsible management. Also, it is influenced the most by indicators that cannot guarantee the accuracy of measurement results. According to the conclusion of the Holmström and Milgrom model (1991), lowering the proportion of performance-based pay may be socially optimal in this case.

Let us review a few more issues in this regard. The first issue concerns the proper ratio of performance-based pay to the total wage and differentials in the amount of such pay. Under the current system, performance-based pay ranges from 100% to 300% of basic monthly wage depending on the type of institution. With regard to the current state, many have argued that the amount of performance-based pay is large and that differentials among institutions are excessive. According to a survey of the employees of quasi-governmental institutions conducted by Young-Jae La and Tae-Beom Yoon (2013), nearly 50% of the respondents said that there should be an improvement in both the amount of performance-based pay and differentials among institutions. In this respect, we need to readjust the level of performance-based pay and the extent of differentials thereof among institutions.

As suggested by Holmström and Milgrom (1991), we need to classify evaluation indicators into easily measurable ones and immeasurable ones and to accordingly delegate tasks to public institutions. In other words, if the nature of a given institution's activities is difficult to evaluate, a performance-based pay scheme should be abolished. Instead, only basic wage should be provided. On the other hand, if there exists a better mechanism for certain types of institutions (e.g., stock prices in the case of listed public corporations) to assess their activities, these institutions should be exempted from the PIMPES. This will naturally lead to a readjustment in the scope of institutions subject to management performance evaluation under the PIMPES. As for the latter, it is reasonable to convert the function of the PIMPES into a control device to address moral hazard and to have it evaluate the latter the institutions accordingly.

References

THE Board of Audit and Inspection of KOREA, “Audit Result Report: State of Public Finance Financial and Business Structure Management,” 2013.

Public Administration Research Institute, “Study on Improvement of Management Evaluation System of Public Entities,” 2015.

Kim Bitmaro, “Contribution to the Theory of Contract in 2016: Nobel Prize in Economics,” Financial Forum, No.245, Korea Institute of Public Finance, 2016.

Kim Wanhee, “A Study on the Improvement of the Measuring Indicator System for Evaluating the Core Performance of Public Institutions,” Korea Institute of Public Finance, 2010.

Kim Ju Chan and Lee Min-Chang, “A survey of the duties of the person in charge of the effectiveness of the public institution management evaluation results,” 2014.

Kim Ji-young, “A Study on the Productivity of Public Enterprise,” Korea Institute of Public Finance, 2011.

Ministry of Strategy and Finance, “Reorganization of public institution management evaluation system,” December 29, 2010

_____, “Public Administration Evaluation Manual for 2012,” 2011.

_____, “Decision on the evaluation result of public institution's management performance in 2015,” 2016A.

_____, “Guidelines for budget execution by public corporations and quasi-government agencies in 2016,” 2016B.

La Gyujae and Yoon Tae-Beom, “Analysis of Management Evaluation System of Public Institutions and Development of New Model,” Korea Institute of Public Finance, 2013.

Park, Mi-jung, “Present Status and Tasks of Public Institution Management Performance Evaluation System,” National Assembly Legislative Investigation Department, 2010.

Lee, Suk Won, “Performance Measurement of Public Institutions Using Coordination Performance Indicators: Application to Government Investment Management Evaluation,” *Korean Journal of Public Administration*, Vol. 39, No. 4, 2005.

Lee O and Seung-hyun Yoo, “A Study on the Problems and Improvements of the Public Management Evaluation System - Focusing on the 2008 Performance Evaluation System,” Audit and Inspection Research Institute, 2010.

Lee, Won-hee and La Gifted, “30 Years of Public Management Evaluation, Retrospective and Prospect,” Korea Institute of Public Finance, 2015.

Lee, Chang-gil and Choi, Sukwon, “Analysis of Possibility of Error in Relative Evaluation in Public Management Evaluation,” Journal of Public Administration, Vol. 48, No. 1, 2009.

Hae SJeong and Oh Young-min and La Gifford, “Policy Measures for Debt Management in Sustainable Public Institutions,” Korea Institute of Public Finance, 2014.

Korea Institute of Public Finance, “A Study on the Transition Process of the Public Management Evaluation System,” 2010.

Choi Han-min, “Public-sector debt and moral hazard: Estimation of the amount of subsidy incurred by the government's tacit guarantee,” Korea Institute of Public Finance, 2015.

Bolton, Patrick, and Mathias Dewatripont. Contract theory. MIT press, 2005.

Hölmstrom, Bengt. and Paul Milgrom. “Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design.” Journal of Law, Economics, & Organization 7 (1991): 24-52.

Wilson, James Q. Bureaucracy: What government agencies do and why they do it. Basic Books, 1989.