

The impact of quality of performance measures on the performance of public sector organisations: Quantitative analysis based on questionnaire survey

Takehiro Metoki

Faculty of Economics, Musashi University

Japan Cost Accounting Association

E-mail: t-metoki@cc.musashi.ac.jp

Abstract

Although the introduction of PMS has progressed in many organisations, it has not necessarily been effectively utilised. Since the introduction and use of PMS has become a global trend, it has been reported that perfunctory introduced PMS has a negative impact on organisational performance. The issue of the quality of performance measures, which is shown to have a significant impact on the effectiveness of PMS in public sector organisations, has not been fully discussed. Low-quality performance measures can lead to incorrect decisions and actions and can undermine the credibility of PMS. Furthermore, if PMS constructed with a large amount of tax is not used properly, then it will lead to tax waste. Therefore, it is necessary to maintain and improve the quality of performance measures when designing and using PMS in public sector organisations. Previous studies pointed out that the quality of performance measures have a significant impact on organizational performance, the relationship between the two has not been clarified quantitatively. This study presents empirical evidence that the quality of performance measures (validity, legitimacy, and functionality of performance measures) increase positive outcomes of performance management systems use in public sector organisation and suppress negative outcomes of performance management systems such as loss of usefulness and side effect. Data for analysis was collected by mailing questionnaire survey; it was conducted in FY2016 in 791 Japanese cities (designated cities, core cities, special cities, and other cities) Japan. The number of valid responses was 327 (valid response rate 41.3%).

Keywords

Performance management systems, Quality of performance measures, Validity of performance measures, Legitimacy of performance measures, Functionality of performance measures

(1) Introduction

The use of performance management systems (PMS) has become a global trend in public sector organisations (Pollitt 2006). The use of PMS not only creates incentives to align organisational objectives with individual goals but also provides valuable feedback on the progress of listed goals (Cavalluzzo and Ittner, 2004; Heinrich, 1999; Kravchuk and Schack, 1996). Owing to these benefits, PMS, also called public sector evaluation "gyousei-hyouka", has been introduced by many public sector organisations in Japan since the late 1990s (Matsuo, 2009; Metoki, 2012).

Although the introduction of PMS has progressed in many organisations, it has not necessarily been effectively utilised. Since the introduction and use of PMS has become a global trend, it has been reported that perfunctory introduced PMS has a negative impact on organisational performance (Bevan and Hood, 2006; de Bruijn, 2002; Newberry and Pallot, 2004). Furthermore, PMS in public sector organisations not only impacts positively, but also has the risk of negative consequences; if it is not designed and used appropriately, then it may lead to erroneous decision-making and actions. (Cuganesan et al., 2014; van Dooren et al., 2015). This leads to the question of what considerations should be taken when designing and using PMS.

In this regard, research has shown that the effect of PMS on the relationship between PMS design and use and results in public sector organisations differs depending on how to use of performance management information (Metoki and Senoo, 2012) and various contingency factors inside and outside

the organisation (van Helden and Reichard, 2013). These studies are increasingly clarifying the relationship between PMS design and use and organisational performance in public sector organisations. However, the issue of the quality of performance measures, which is shown to have a significant impact on the effectiveness of PMS in public sector organisations, has not been fully discussed (van Dooren et al., 2015). Low-quality performance measures can lead to incorrect decisions and actions and can undermine the credibility of PMS. Furthermore, if PMS constructed with a large amount of tax is not used properly, then it will lead to tax waste. Therefore, it is necessary to maintain and improve the quality of performance measures when designing and using PMS in public sector organisations.

In this regard, Bouckaert (1993) recognises the three aspects of validity, legitimacy, and functionality as the quality of performance measures important for PMS in public sector organisations. In the case of validity, performance measures are measured as objective measures without being distorted by external influences, and legitimacy means that all organizational members support the design and use of performance measures. Finally, functionality means that the measured performance measure is realistic and suitable for managerial use. In line with Bouckaert (1993), this research considers the quality of performance measures to be composed of 1) validity, 2) legitimacy, and 3) functionality. According to Bouckaert (1993), the quality of performance measures affects the performance of public sector

organisations. Despite their impacts, research only presents the concepts that make up the quality of performance measures, and there is no quantitative analysis of the impact on the performance of public sector organisations. Therefore, in this study, the relationship between the quality of performance measures and the performance of public sector organisations is analysed quantitatively based on the results of a questionnaire survey conducted in 791 cities in Japan.

The rest of this paper is structured as follows. Section 2 summarises related studies and presents the analysis framework of this study. Section 3 employs variables, and Section 4 discusses the results of the analysis. Finally, Section 5 presents the contributions and limitations of this study and draws conclusions.

(2) Previous research and analysis framework

1. Positive and negative consequences of using PMS

Metoki and Senoo (2012), using 264 observations based on a questionnaire survey on Japanese government organisations, revealed that the results achieved by using PMS differ depending on the use of performance management information (financial outcomes/improved responses to residents' needs/psychological empowerment of those in charge of public sector reforms). Additionally, Speklé and Verbeeten (2014) studied the relationship between PMSs usage and organisational performance using 101 observations based on questionnaires targeting Dutch public sector organisations.

It revealed that contractibility which measured by the three concepts of clarity of goals, measurability, and the degree to which managers know and control the transformation process moderates the relationship between the incentive-oriented use of the performance measurement system and performance. However, it has been pointed out that PMS in public sector organisations does not always produce positive results; as stated earlier, it can produce negative results if not properly designed and used (Cuganesan et al. 2014; van Helden and Reichard 2013). In this regard, Bouckaert (1993) pointed out that the use of PMS will facilitate false decision-making and behaviour if the quality of performance measures is not maintained at an appropriate level. Similarly, Melnyk et al. (2014) found that performance measures and target are not modified according to the situation in which the organisation is placed; in other words, in certain cases, organisations continue to use these measures without proper modification. (Melnyk et al. 2014). Therefore, it is necessary to clarify the relationship between the quality of performance measures in the PMS of public sector organisations and organisational performance.

2. Building an analysis framework

In this study, we will focus on the concepts of validity, legitimacy, and functionality, and conduct an exploratory analysis of the relationship between the quality of performance measures and the performance of public sector organisations. However, neither Bouckaert (1993) nor van

Dooren et al. (2015), who indicate the importance of the quality of performance measures in public sector organisations, provides a measure of the quality of performance measures. Therefore, we will address this gap in this study. First, validity is the first concept that constitutes the quality of performance measures. Bouckaert (1993) considers validity as a performance measure that is measured objectively, without being distorted by external influences. Additionally, validity is the consistency of measurement, and hence its design must ensure that the measurement's results do not vary due to the influence of external factors, other than the quality of staff behaviour. Therefore, it is important that performance measures are designed and measured appropriately to ensure that measurement results are not distorted by external influences. The second concept that constitutes the quality of performance measures is legitimacy. As per Bouckaert (1993), legitimacy supports the design and use of performance measures by all members of the organisation. Additionally, it is reported that legitimacy is ensured by the participation of the members of the organisation in the design and use of performance measures. This prevents unauthorized manipulation and gaming of performance measures. Therefore, legitimacy is expected to increase by the participation of organisational members in designing, using, and modifying performance measures. The third concept that constitutes the quality of performance measures is functionality. Bouckaert (1993) considers functionality to be a realistic measurement so

that measured performance measures are suitable for use. Therefore, if the performance measures becomes unrealistic because of changing situation, it will be necessary to stop using or redesigning the performance measures.

Conversely, recent studies have pointed out that the effect of PMS in public sector organisations may not be necessarily positive, but also negative (Cuganesan et al., 2014; van Helden and Reichard, 2013). Therefore, when analysing the relationship between the quality of performance measures and the performance of public sector organisations, we consider both positive and negative outcomes.

Based on the aforementioned studies, a multiple regression analysis is conducted with validity, legitimacy, and functionality as independent variables and the performance of the public sector organisation as dependent variables. In the analysis, environmental uncertainty, operational complexity, and the number of employees are taken as control variables in consideration of the varying environments surrounding public sector organisations and the diverse business contents of the responding organisations. The analysis framework of this study is shown in Figure 1.

(3) Employment of data and variables

1. Data

Data for analysis was collected by mailing questionnaire survey; it was conducted in FY2016 in 791 Japanese cities (designated cities, core cities, special cities, and other cities). Prior to sending the questionnaire, we asked two practitioners,

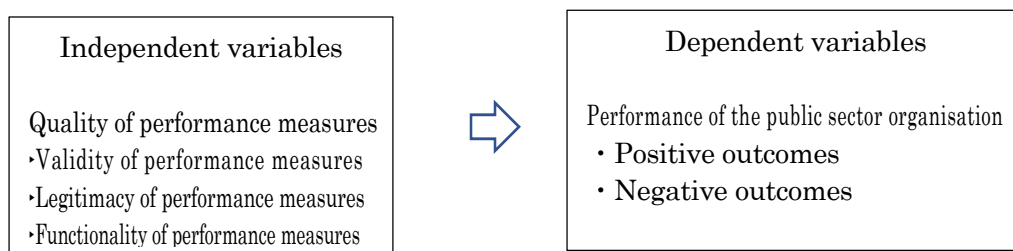


Figure1 The analysis framework

who supervised the related work in the public sector organisation, and two management accounting researchers with specialized knowledge in performance management to check the validity of the wording and question items. The questionnaire was received by the supervisor of the management department of the public sector organisation, and the questionnaire was sent with a request letter and reply envelope. In consideration of incentives for answering the questionnaire, it was clarified that a report on the analysis results will be sent to the applicant. In order to improve the collection rate, a reminder was sent before the collection deadline. The final number of responding organisations was 339 (recovery rate 42.9 %). In the analysis, we used data, excluding the responses of organisations that gave inadequate response and were judged inappropriate for analysis. The number of valid responses was 327 (valid response rate 41.3%). Details of the responding

organisations are shown in Table 1. Additionally, respondents comprised staff at a section manager level or higher in the responding department. A goodness-of-fit test was performed on the data collected from 327 public sector organisations used in the analysis. The results confirmed that the data of the public sector organisation was generally compatible with the group distribution ($\chi^2 = 7.770, df = 3, p = .051$). Additionally, we conducted a test of the difference between the size of the target organisation and the non-target organisation (public sector organization: number of employees), but no significant difference was found between the two. The results did not show any significant non-response bias in the data.

2. Variable manipulation

2.1 Validity of performance measures

Bouckaert (1993) regards the validity of performance measures as being measured

Table 1 Response distribution

Group classification	Number of sent	Number of valid response (rate)
designated cities	20	14 70.0%
core cities	48	20 41.7%
special cities	36	8 22.2%
other cities	687	285 41.5%
total	791	327 41.3%

as an objective measure without being distorted by external influences. In other words, there is a need to use objective performance measures to produce verifiable results and ensure that they are not changed by external influences. Therefore, we referred to Moers (2006), which analysed the relationship between performance measures' characteristics and performance (Table 2). Specifically, we determined whether the results of the performance measures are set so that they cannot be affected externally and whether the performance measures

accurately reflect the results of the staff's actions, or that the staff's work is good. The questionnaire asked whether it is directly reflected in the performance measures or whether these measures are objective; the responses were sought from 5 points, from '1 (not at all) to 5 (exactly)'. In the variabilisation, these average values are scored ($\alpha = .695$). Although Cronbach's α is slightly below the general level of .70, it is close to .70. Hence, we believe that there is no major problem in validity.

Table 2 Validity of performance measures

		Range	Mean	S. D.
VALID_1	The results of performance measures are appropriately designed to ensure that they are not affected by external influences, other than those emerging from staff's actions.	1 – 5	3.03	.816
VALID_2	Performance measures accurately represent whether the staff exhibited superior performance.	1 – 5	2.86	.784
VALID_3	If staff exhibited superior performance, it was directly reflected in performance measures	1 – 5	2.98	.746
VALID_4	Performance measures are objective and verifiable	1 – 5	3.50	.743

2.2 Legitimacy of performance measures

Bouckaert (1993) considers the legitimacy to be that the design and use of performance measures are supported by all members of the organization so that the performance measures can be designed and used to prevent unauthorized manipulation and gaming. Legitimacy is ensured by the participation of organizational members in the design and use of the PMS. Therefore, Abernethy and Bouwens (2005) also empirically clarified that the degree of employee participation in the management information system design has a positive impact on the satisfaction levels and

performance of the management information system usage. Based on Bourne et al. (2002), this study discusses the participation of organizational members in the design and use of management systems; moreover, the study sets and amends performance measures, collects data, and conducts analysis to achieve numerical targets for performance measures. Five questions are set regarding the degree of participation of non-managerial staff in processes such as the evaluation of the achievement status of the performance measures (Table 3). The above questions were responded on the basis of 5 points, from '1 (not participate at all) to 5 (Actively participating)'.

Table 3 Legitimacy of performance measures

	Range	Mean	S. D.
LEGIT_1 Setting of performance measures	1 – 5	3.72	.801
LEGIT_2 Selecting the data for setting of performance measures	1 – 5	3.63	.829
LEGIT_3 Amending performance measures	1 – 5	3.60	.822
LEGIT_4 Actions to achieve performance targets	1 – 5	3.88	.713
LEGIT_5 Analysis to achieve numerical targets for performance measures	1 – 5	3.65	.821

While constructing the variables, these average values were scored ($\alpha = .905$).

2.3 Functionality of performance measures

Bouckaert (1993) sees functionality as a realistic measurement so that measured performance measures are suitable for use. Therefore, when there is a difference between the initial performance measure setting and the actual situation, it is important to adjust the settings in accordance with the actual situation; it is also important to determine whether the monitoring can contribute toward facilitating this adjustment. The measures and targets in public sector organisations are often inflexible, which may not be conducive to adjustments made according to the situation and environment. Smith (1995) refers to this type of risk as ossification of performance evaluation. de Bruijn (2002) also argues that once performance measures are set, the continuous reproduction of existing services is encouraged and the demand for change can be ignored. If there is a change in the external or internal environments of organisations, then the measures must be modified appropriately,

failing which they can have a negative impact on organisational performance (Melnik et al., 2014). Therefore, we set four questions to determine whether the target can be changed flexibly in response to environmental changes (Table 4). These responses were given on the basis of 5 points, from '1 (not at all) to 5 (exactly)'. When constructing the variables, these average values were scored ($\alpha = .713$)¹.

2.4 Performance of public sector organisations

The use of PMS brings various benefits to an organisation. Non-financial as well as financial results have been used for performance assessment, as shown in many quantitative studies (e.g. Davis and Albright, 2004; Ittner and Larcker, 2001). However, public sector organisations cannot measure performance using financial results such as sales and profit measures. Therefore, many public sector studies have relied on the items of van de Ven and Ferry (1980), which are considered suitable for measuring public sector performance, and determined the accuracy of productivity and operations. Measurements were made using the following

¹ In order to change the functionality of performance measures, we used four questions. However, when the question item 'Change of target during the period is not

allowed (reversal scale)' was included, Cronbach's α was .596, which made the scale less reliable. Therefore, variablisation was performed, which excluded this item.

Table 4 Functionality of performance measures

	Range	Mean	S. D.
FNCT_1 Targets are updated regularly according to certain criteria	1 – 5	2.88	1.175
FNCT_2 When setting targets, adjustments are made throughout the year for situations that could not have been foreseen when setting targets unforeseen events.	1 – 5	3.15	.982
FNCT_3 Ongoing monitoring was conducted within the organization to assess whether targets are realistic or whether they should be changed.	1 – 5	3.14	1.010

seven items: productivity, accuracy of work, an increase in the number of innovative new projects, satisfaction from and reputation of the work, the achievement of targets, work efficiency, and improvement in staff motivation towards work. In this study, we added six items, such as budget adjustment and cost awareness, which were measured by major PMS studies such as Hall (2008). Additionally, it has been pointed out that PMS yields both positive and negative outcomes. Therefore, we also measured the seven negative outcomes emerging from use of PMS, as described by Cuganesan et al. (2014). We conducted an exploratory factor analysis (using the main factor method and promax rotation) on the performance of public sector organisations, based on these 20 questions. We identified the elements that contribute to the performance of public sector organisations. As a result, as shown in Table 5, it became clear that the performance of public sector organisations is composed of four factors. The first factor was named 'Task performance' because it comprised the result of work activities, such as work quality,

efficiency, and improvement of residents' satisfaction ($\alpha = .826$). The second factor was named 'Behavioural performance'² because items related to the impact on staff behaviour, such as cost awareness and motivation, showed a high factor loading ($\alpha = .806$). The third factor was termed 'Loss of usefulness' because items related to the loss of usefulness of PMS—such as the fact that performance measures are underutilised and poorly indicate organisational performance to residents and parliament—indicated a high factor loads ($\alpha = .707$). The fourth factor was named 'Side effects' because items restricting communication within the organisation, such as the inhibition of organisational flexibility and wrinkles within the organisation, showed a high factor loading ($\alpha = .704$). Additionally, in the variabilisation of each item, the average value of the question items that showed a factor load of .35 or more was scored.

2.5 Control variable

In order to control the analysis results, the number of employees was set as a control

² Items that make up the second factor include items that may not be directly linked to behavioural performance, such as an 'increase in the number of innovative new

project. However, these items are also considered to be indirectly generated by behavioural changes, and hence the second factor is named 'behavioural performance'.

Table 5 Factor analysis on the performance of public sector organisations

		Range	Mean	S.D.	Factor1 Task performance ($\alpha = .826$)	Factor2 Behavioural performance ($\alpha = .806$)	Factor3 Loss of usefulness ($\alpha = .707$)	Factor4 Side effects ($\alpha = .704$)
EFT_2	Quality or accuracy of work Improvement in the achievement	1 – 5	3.41	0.683	0.861	-0.149	0.036	-0.008
EFT_5	of business and departmental targets	1 – 5	3.65	0.694	0.755	-0.005	-0.008	0.008
EFT_4	Improvement in operational efficiency	1 – 5	3.47	0.749	0.611	0.203	0.105	0.017
EFT_1	Increase in business volume (production volume) and service provision	1 – 5	3.18	0.685	0.537	0.063	0.063	-0.034
EFT_12	Improvement in information collection and recording methods for measuring output and outcomes	1 – 5	3.23	0.732	0.421	0.222	-0.039	-0.015
EFT_6	Increase in the satisfaction and reputation of residents and users	1 – 5	3.21	0.751	0.398	0.174	-0.116	0.065
EFT_8	Coordination of Budgets and plan with other departments	1 – 5	3.25	0.831	-0.116	0.805	0.028	-0.067
EFT_13	Information exchange with staff in other departments	1 – 5	3.17	0.778	-0.033	0.679	-0.021	0.068
EFT_11	Reduction in the time incurred for making decision associated with targets, policies, and work schedule	1 – 5	3.01	0.750	0.004	0.574	-0.053	-0.040
EFT_9	Improvement in cost reduction awareness	1 – 5	3.38	0.764	0.160	0.570	0.014	0.015
EFT_7	Improvement staff motivation towards work	1 – 5	3.18	0.713	0.213	0.432	-0.147	0.011
EFT_3	Increase in the number of innovative new business projects	1 – 5	2.88	0.727	0.228	0.428	0.085	-0.011
EFT_15	Performance measures and related information are underutilised	1 – 5	2.91	0.891	0.082	-0.048	0.816	-0.070
EFT_16	Performance measures are not tied to organisational targets	1 – 5	2.69	0.889	-0.027	-0.016	0.753	-0.025
EFT_14	Performance measures tend to be poor measures of performance to residents and parliament.	1 – 5	2.45	0.887	-0.275	0.165	0.485	0.117
EFT_18	Performance measures are not designed to reflect business results or workload	1 – 5	2.80	0.862	0.154	-0.050	0.400	0.127
EFT_17	Measurement and evaluation of performance measures has led to the creation of traps and excessive competition within the organisation	1 – 5	1.99	0.698	-0.004	0.039	-0.021	0.789
EFT_19	Measurement and evaluation of performance measures has led to reduced organisational flexibility and disrupted change	1 – 5	2.29	0.752	-0.010	-0.082	0.089	0.662
Eigenvalue					6.068	1.858	1.260	1.008
Factor correlation					Factor1 —	—	—	—
					Factor2 0.670	—	—	—
					Factor3 -0.562	-0.545	—	—
					Factor4 -0.240	-0.111	0.427	—

variable; this is because items such as environmental uncertainty, business complexity, and organisation scale are expected to affect the performance management of public sector organisations. Regarding environmental uncertainties and operational complexity, we referred to Speklé and Verbeeten (2014) and Oura and Matsuo (2017) to set questions and conduct exploratory factor analysis (main factor method and promax rotation)³. Additionally, when conducting factor analysis, items for which the ceiling effect was confirmed and items that showed only factor loadings below .35 for all factors were deleted. As a result, the first factor denoting the uncertainty of the work environment had a high factor loading in the item indicating uncertainties related to the future predictability of work, and therefore this factor was termed 'Task predictability' ($\alpha = .666$). The second factor was named 'Business design difficulty' ($\alpha = .776$) because it showed a high factor loading in the items indicating difficulty in predicting needs and effects related to business design (upper part of Table 6). Although Cronbach's α of 'business predictability', which is the first factor, is below the general standard of .70, there is no major problem in validity because it is close to .70. It was judged. The first factor of business complexity is 'Business process clarity' ($\alpha = .790$) because it had a high factor

loading in the item indicating whether business procedures and processes were clarified by manuals and regulations. The second factor was named 'Measurable outcome' ($\alpha = .714$) because it showed a high factor loading on the items related to the measurable outcome of the department (lower part of Table 6). Additionally, in the variabilisation of each item, the average value of the question items that showed a factor load of .35 or more was scored. Finally, in order to control the scale of the organization, the value obtained by logarithm conversion of the number of employees is used in the analysis.

(4) Analysis results

Based on the above variables, we conducted a multiple regression analysis; validity, legitimacy, and functionality of performance measures were taken as independent variables, and the performance of public sector organisations was taken as a dependent variables⁴. Environmental uncertainty, operational complexity, and scale (number of employees) were set as control variables. The correlation coefficient between variables (Table 6) and the results of multiple regression analysis (Table 7) are as follows. As a result of the analysis, it became clear that any variable that makes up the quality of performance measures affects the performance of public sector organisations.

³ Speklé and Verbeeten (2014) do not perform factor analysis for each question item on both uncertainty of work environment and work complexity, but they are variabilised by the simple average method. However, as a result of performing the same operation in this study, the value of Cronbach's α became low,

and hence factor analysis was performed for enhancing the accuracy of the analysis.

⁴ The variance of inflation of the input independent variable was confirmed, but the maximum value was 1.325, and hence it is concluded that the problem of multicollinearity does not exist.

First, it improves operational results such as the validity of performance measures, legitimacy of performance measures, functionality of performance measures, productivity, and work efficiency (validity of performance measures: $B = .217, \beta = .228, t = 4.504$; legitimacy of performance measures: $B = .203, \beta = .262, t = 5.405$; and performance measures functionality: $B = .150, \beta = .242; t = 4.857$).

Table 6 Factor analysis on the control variables

Uncertainty of the task environment		Range	Mean	S.D.	Factor1 Task predictability ($\alpha=.666$)	Factor2 Business design difficulty ($\alpha=.776$)
ENVVDYN_1	There is a high possibility that unpredictable events will occur in daily operations	1 – 5	3.11	0.888	0.767	-0.130
ENVVDYN_2	It is difficult to predict when a new project will start	1 – 5	3.11	0.965	0.702	-0.001
ENVVDYN_4	It is difficult to estimate the required budget for the following year accurately	1 – 5	2.76	0.945	0.410	0.238
ENVVDYN_3	Achievement of targets depends heavily on the external environment	1 – 5	3.28	0.779	0.382	0.108
ENVVDYN_6	When designing a business, it is difficult to predict needs	1 – 5	2.72	0.739	-0.058	0.797
ENVVDYN_7	When designing a business, it is difficult to predict the effect	1 – 5	2.84	0.804	0.069	0.794
Eigenvalue					2.376	1.343
Factor correlation					Factor1 Factor2	— —

Business complexity		Range	Mean	S.D.	Factor1 Business process clarity ($\alpha=.790$)	Factor2 Measurable outcome ($\alpha=.714$)
ENVHOST_2	There is an appropriate and efficient process for conducting business	1 – 5	3.02	.930	.794	.010
ENVHOST_1	There is a standard procedure that should be referenced for addressing differences in business execution	1 – 5	2.93	1.067	.757	-0.019
ENVHOST_4	Procedures for conducting business are clearly shared among various departments	1 – 5	3.26	.902	.673	.071
ENVHOST_3	Procedures for conducting business are stipulated through laws, regulations, and rules, among others	1 – 5	3.21	.982	.534	-0.001
ENVHOST_6	Quantitatively understand what the department should achieve	1 – 5	2.88	.937	-0.060	.977
ENVHOST_5	A single quantitative measure can grasp the results of each business	1 – 5	2.53	.956	.027	.656
ENVHOST_7	Current performance measures accurately represent what the department should achieve	1 – 5	3.23	.781	.189	.362
Eigenvalue					3.104	1.281
Factor correlation					Factor1 Factor2	— —

ENVVDYN_1 to ENVVDYN_7 asked about uncertainties in the work environment and measured them on a five-point scale (“1 not at all” – “5 exactly”). ENVVDYN_5 was removed for factor analysis because both factors showed factor loadings less than .35. ENVHOST_1 to ENVHOST_10 asked about the complexity of the business and measured it on a 5-point scale (“1 not at all” – “5 exactly”). Since ENVHOST_9 confirmed the ceiling effect, ENVHOST_8 and ENVHOST_10 showed factor loadings less than .35 for both factors, and hence were excluded from the factor analysis.

Table 7 Correlation coefficient between variables

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1)TSK_PERF	1										
(2)BEH_PERF	.661**	1									
(3)RER_NPERF	-.446**	-.446**	1								
(4)SIDE_NPERF	-.218*	-.157**	.393**	1							
(5)VALID	.419**	.408**	-.342**	-.125*	1						
(6)LEGIT	.419***	.342**	-.289**	-.282**	.263**	1					
(7)FNCT	.415*	.377**	-.397**	-.075	.304**	.284*	1				
(8)PREDICT	-.084	-.029	.112*	.023	-.165**	-.081	-.112*	1			
(9)DIFFICULT	-.126*	-.144**	.118*	.127*	-.166**	-.118*	-.120*	.292**	1		
(10)CLARITY	.179**	.257**	-.216**	-.035	.283**	.155**	.293**	-.282**	-.191**	1	
(11)MEASURABLE	.350**	.309**	-.250**	-.069	.333*	.199**	.244**	-.237**	-.177**	.400**	1

1) Pearson's correlation coefficient

2) TSK_PERF: Task Performance, BEH_PERF: Behavioural performance, RER_NPERF: Loss of usefulness, SIDE_NPERF: Side effects, VALID: Validity of performance measures, LEGIT: Legitimacy of performance measures, FNCT: Functionality of performance measures, PREDICT: Task predictability, DIFFICULT: Business design difficulty, CLARITY: Business process clarity, MEASURABLE: Measurable outcome

3) **p<0.01, *p<0.05 (two-sided)

Table 8 The results of multiple regression analysis

	Task Performance			Behavioural performance			Loss of usefulness			Side effects		
	B	β	t-value	B	β	t-value	B	β	t-value	B	β	t-value
Constant term	.960		3.007***	.729		2.110**	5.299		12.476***	3.328		7.253***
VALID	.217	.228	4.504***	.243	.246	4.617***	-.209	-.178	-3.235***	-.045	-.040	-.658
LEGIT	.203	.262	5.405***	.159	.193	3.792***	-.123	-.130	-2.464**	-.239	-.257	-4.444***
FNCT	.150	.242	4.857***	.127	.201	3.822***	-.197	-.260	-4.809***	.015	.020	.340
PREDICT	.032	.038	.775	.088	.103	2.009**	.016	.015	.289	-.036	-.037	-.626
DIFFICULT	-.016	-.021	-.432	-.041	-.053	-1.055	.017	.018	.347	.095	.105	1.835*
CLARITY	-.045	-.065	-1.253	.049	.069	1.259	-.019	-.022	-.393	.031	.038	.605
MEASURABLE	.143	.195	3.749***	.096	.129	2.343**	-.088	-.098	-1.738*	-.013	-.015	-.238
Number of staff	.030	.047	1.010	.016	.024	.494	-.103	-.130	-2.607*	-.063	-.081	-1.476
R ²	.368			.306			.261			.097		
Adj. R ²	.352			.288			.241			.074		
F-value	22.500***			16.881***			13.531***			4.154***		

1) Estimation based on ordinary least squares (OLS) method

2) B: Non-standard regression coefficient, β : Standard regression coefficient, R²: Coefficient of determination, Adj. R²: Modified coefficient of determination

3) *** p < 0.01, ** p < 0.05, * p < 0.1 (two-sided)

Additionally, an improvement in positive outcomes was observed, such as improvement of staff motivation and the active exchange of information with other legitimacy of performance measures: B = .159,

$\beta = .193$, t = 3.792; and performance measures of functionality: B = .127, $\beta = .201$; t = 3.822).

As pointed out by Bouckaert (1993) and van Dooren et al. (2015), it is important to improve the quality of performance measures when designing and using PMS in public

sector organisations. This suggests that improving the validity, legitimacy, and functionality of measures improves the performance of public sector organisations. In other words, when designing and using PMS for public sector organisations, the performance measures are designed in a manner that they are not distorted by external influences. It is essential to build a mechanism that will allow managers and non-managerial staff to participate in the design and use of performance measures; this mechanism will contribute towards improving task and behavioural performance. Groen et al. (2017) clarified that the involvement of operational employees in the design and implementation of performance measures improves the job performance of employees. It can be said that this is consistent with previous studies. Additionally, the measurability of outcomes affects the relationship between each variable related to the quality of performance measures and task and behavioural performance. In other words, it is suggested that it is necessary to give sufficient consideration to the quality of performance measures when the business or project allows the quantitative measurement of the results of each department or project. These variables related to the quality of performance measures not only increase positive outcomes but also show the potential to suppress negative outcomes. In other words, the validity, legitimacy, functionality of performance measures suppresses negative outcomes of performance management systems such as loss of usefulness (validity of performance measures: $B = -.209, \beta =$

$-.178, t = -3.235$; legitimacy of performance measures: $B = -.123, \beta = -.130, t = -2.464$; and functionality of performance measures: $B = -.197, \beta = -.260, t = -4.809$). Additionally, legitimacy of performance measures, that is, staff participation in setting and using performance measures is also suppressed side effects of PMS use, such as excessive competition in the body and reduced flexibility of organisation ($B = -.239, \beta = -.257, t = -4.444$). Although Cuganesan et al. (2014) mentions the possibility that the use of PMS in public sector organisations not only produces positive outcomes but also negative outcomes, it has not been revealed that how to suppress these negative outcomes of PMS. In this respect, an improvement in the quality of performance measures, such as validity, legitimacy, and functionality of performance measures may suppress the negative outcomes.

In public sector organisations, once performance measures are set, they are often not reviewed for reasons such as year-on-year comparisons and comparisons with other organisations. Ensuring comparability is, of course, important in fulfilling accountability of the government. But leaving improvements of quality of performance measures can also lead to poor performance of public sector organisations. Therefore, it is necessary to verify and review the performance measures periodically to sustain the quality of the performance measures.

(5) Contribution of this research and future issues

Although previous studies pointed out that the quality of performance measures has

a significant impact on organizational performance (van Dooren et al., 2015), the relationship between the two has not been clarified quantitatively. In this regard, this study presents empirical evidence that the validity, legitimacy, and functionality of performance measures increase positive outcomes. This is the first contribution of this study. Additionally, although negative outcomes from the use of PMS have been pointed out (Cuganesan et al., 2014), the means to suppress them have not been presented. This study addresses this gap by demonstrating that negative outcomes can be suppressed by improving the quality of performance measures, that is, the validity, legitimacy, and functionality of performance measures. However, while making the above-mentioned contributions, it must be pointed out that this research has its limitations. This study analyses the quality of performance measures by focusing on the validity, legitimacy, and functionality of performance measures presented by Bouckaert (1993). Additionally, it attempts to grasp the quality of the performance measures by considering the validity of the construct, as much as possible, relying on previous research. However, since the concept of quality of performance measures is composed of various elements, this study does not capture all aspects. For example, the measures' characteristics such as measurement, importance, measurement cost, sensitivity, and noise discussed in Kawai (2016) are considered to be factors related to the quality of performance measures in public sector organisations. Additionally, when measuring negative outcomes, the study formulated

question items by relying on the literature on negative outcomes of PMS use, such as the study by Cuganesan et al. (2014). There is a possibility that the negative outcomes may not be fully understood. The points comprise the limitations of this research, which must be evaluated in future research.

Acknowledgements

This paper's original version (Metoki, 2018) won the Best Paper Award at the Japan Cost Accounting Association (JCAA) and was published in this journal with the recommendation of the president. The author would like to thank the president of JCAA, and two anonymous reviewers of Journal of Japanese Management.

References

- Abernethy, Margaret A. and Jan Bouwens. (2005). Determinants of Accounting Innovation Implementation. *Abacus*, 41 (3), pp.217-240.
- Bevan, Gwyn and Christopher Hood. (2006). What's Measured is What Matters: Targets and Gaming in the English Public Health Care System. *Public Administration*, 84 (3), pp.517-538.
- Bouckaert, Geert. (1993). Measurement and Meaningful Management. *Public Productivity and Management Review*, 17, pp.31-43.
- Bourne, Mike. John J. Mills, Mark Wilcox, Andy Neely, and Ken Platts. (2002). Designing, Implementing and Updating Performance Measurement Systems. *International Journal of Operations and Production Management*, 20 (7), pp.754-771.

- de Bruijn, Hans. (2002). Performance Measurement in the Public Sector: Strategies to Cope with the Risks of Performance Measurement. *International Journal of Public Sector Management*, 15 (7), pp.578-594.
- Cavalluzzo, Ken S. and Christopher D. Ittner. (2004). Implementing Performance Measurement Innovations: Evidence from Government. *Accounting, Organizations and Society*, 29 (2), pp.243-267.
- Cuganesan, Suresh. James Guthrie and Vedran Vranic. (2014). The Riskiness of Public Sector Performance Measurement: A Review and Research Agenda. *Financial Accountability and Management*, 30 (3), pp.279-302.
- Davis, Stan and Tom Albright. (2004). An Investigation of the Effect of Balanced Scorecard Implementation on Financial Performance. *Management Accounting Research*, 15 (2), pp.135-153.
- van Dooren, Wouter. Geert Bouckaert and John Halligan. (2015). *Performance Management in the Public Sector*. London and New York: Routledge and Taylor and Francis Group.
- Grafton, Jennifer. Anne M. Lillis and Sally K. Widener. (2010). The Role of Performance Measurement and Evaluation in Building Organizational Capabilities and Performance. *Accounting, Organizations and Society*, 35(7), pp.689-706.
- Groen, Bianca. A. C. Marc. J. F. Wouters and Celeste P. M. Wilderom. (2017). Employee Participation, Performance Metrics and Job Performance: A Survey Study Based on Self-determination Theory. *Management Accounting Research*, 36, pp.51-66.
- Hall, Matthew. (2008). The Effect of Comprehensive Performance Measurement Systems on Role Clarity, Psychological Empowerment and Managerial Performance. *Accounting, Organizations and Society*, 33 (2/3), pp.141-163.
- Heinrich, Carolyn J. (1999). Do Government Bureaucrats Make Effective Use of Performance Management Information? *Journal of Public Administration Research and Theory*, 9 (3), pp.363-393.
- van Helden, G. Jan and Christoph Reichard. (2013). A Metareview of Public Sector Performance Management Research. *Tékhné-Review of Applied Management Studies*, 11, pp.10-20.
- Ittner, Christopher. D. and David F. Larcker. (2001). Assessing Empirical Research in Management Accounting: A Value-based Perspective. *Journal of Accounting and Economics*, 32, pp.349-410.
- Kawai, Takaharu. (2016). Gyouseki Sokutei no Sokutei Taishou to Shihyou Tokusei (Performance Measuring Objects and Its Properties). *Journal of Cost Accounting in Japan*, 40(2), pp.87-99.
- Kravchuk, Robert and Ronald Schack. (1996). Designing Effective Performance Measurement Systems under the Government Performance and Results Act 1993. *Public Administration Review*, 56 (4), pp.348-358.
- Matsuo, Takami. (2009). Jichitai no Gyouseki Kanri Shisutemu (*Performance management systems in local government*). Chuokeizai-sha, Inc. Japan.

- Melnyk, Steven. A., Umit Bititci, Ken Platts, Jutta Tobias, and Bjørn Andersen. (2014). Is Performance Measurement and Management Fit for the Future? *Management Accounting Research*, 25 (2), pp.173-186.
- Metoki, Takehiro. (2018). Gyouseki Shihyou no Shitsu ga Gyousei Soshiki no Pafoumansu ni ataeru Eikyou: Sitsumonhyou Cyousa ni Motozuku Teiryuu Bunseki (The effects of performance measures quality on performance management systems in local government). *The Journal of Cost Accounting Research*, 42(9), pp.97-109.
- Metoki, Takehiro. (2012). Gyouseki Kanri Shisutemu to shite no Gyousei Hyouka (A Survey of Performance Management in Japanese Local Government). *The Journal of Musashi University*, 60 (2), pp.45-72.
- Metoki, Takehiro and Takeyoshi Senoo. (2012). Examining Managerial Uses of Accounting and Performance Measurement Information in Local Government. *Journal of Cost Accounting in Japan*, 36 (2), pp.115-129.
- Moers, Frank. (2006). Performance Measure Properties and Delegation. *The Accounting Review*, 81 (4), pp.897-924.
- Newberry, Susan and June Pallot. (2004). Freedom or Coercion? NPM Incentives in New Zealand Central Government Departments. *Management Accounting Research*, 15, pp.247-266.
- Oura, Keisuke and Takami Matsuo. (2017). Chihou Koukyou Danntai niokeru Gyouseki Kanrij Youhou no Riyuu to Yuukousei (The Use of Performance Management Information and Its Effectiveness in Japanese Local Government). *Journal of Cost Accounting in Japan*, 41(1), pp.103-115.
- Pollitt, Christopher. (2006). Performance Management in Practice: A Comparative Study of Executive Agencies. *Journal of Public Administration Research and Theory*, 16 (1), pp.25-44.
- Smith, Peter. (1995). On the Unintended Consequences of Publishing Performance Data in the Public Sector. *International Journal of Public Administration*, 18 (2/3), pp.277-310.
- Speklé, Roland F. and Frank H. M. Verbeeten. (2014). The Use of Performance Measurement Systems in the Public Sector: Effects on Performance. *Management Accounting Research*, 25 (2), pp.131-146.
- Van de Ven, Andrew. H. and Diane L. Ferry. (1980). *Measuring and Assessing Organizations*. Wiley, New York.

(Received: July 31, 2019)

(Accepted: November 1, 2019)