TOKYO METROPOLITAN UNIVERSITY

GRADUATE SCHOOL OF MANAGEMENT

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

**AN EMPIRICAL INVESTIGATION INTO
PUBLIC EFFICIENCY AND BUDGETING: THE CASE OF
TOKYO LOCAL GOVERNMENTS**

Tran Thien Vu

Supervisor: **Masayoshi Noguchi**

2019

ACKNOWLEDGEMENT

Summary of the Thesis

The New Public Management (NPM) doctrine has become the prevailing principle in public sector management reforms across the world since the 1980s. On the one hand, conventional public administration (PA) in the progressive era displayed bureaucratic and inefficient governance; on the other, a transformation from a conventional PA to NPM could bring about considerable merits based on the premise that public governance could be more efficient, effective, and accountable. The salient features of the NPM paradigm can be mainly analogously characterized in four folds along with four mainstreams in this thesis. First, the decentralization of fiscal responsibility from the higher tiers of government into local governments can result in optimum efficiency and accountability at the local level. Second, streamlining local expenditures in the shortage of resource can achieve greater public efficiency in the production process. Third, public sector organizations (PSOs) adopt private managerial practices, particularly private sector accounting norms, hence improving public efficiency and accountability. Fourth, stressing on output controls (NPM model) rather than input controls (conventional PA model) can enhance the performance (outputs and outcomes), which can link to budgeting procedures in planning resource. In this study, we greatly emphasize four main research topics, in accordance with four NPM principles, in the case of Tokyo local governments as follows:

In chapter 3, the first research topic investigates the public finance mechanism associated with the decentralization of revenue and expenditure assignments in Tokyo jurisdictions. We discern that local expenditure volatility can have significant implications for the health of local economies. Therefore, it is essential to understand how fluctuations in the various components of municipal revenue translate into expenditure volatility. We examine the association between the

revenue and expenditure volatility of Tokyo local governments, which are comprised of both special wards and Tama cities, through recourse to a six-year panel of fiscal data (2010–2015). We find evidence of statistically significant positive associations between the volatility of most local taxes and expenditure volatility, but negative associations between the volatility of grants and expenditure volatility. This suggests that grants play an important role in smoothing out local expenditures in Tokyo and that the prescription for greater reliance on local taxation, found in much of the literature, may not be appropriate for Tokyo local governments. The adjustments of intergovernmental grants are necessary to beat expenditure volatility in special wards.

Since public efficiency has become the focal point of interest in public sector management, the second research topic in chapter 4 examines the global efficiency of public service delivery in Tokyo metropolitan municipalities by using the nonparametric method of Data Envelopment Analysis (DEA) over 2001–2015. Furthermore, we compare efficiency scores before and after the introduction of the accrual accounting system in 2008. The findings reveal that public efficiency seems to have declined in special wards but increased in Tama cities since the adoption of the new accounting system.

In chapter 5, the third research topic explores what determinants or external factors are associated with the trend of productive efficiency scores. We regress estimated bias-corrected efficiencies on discretionary and non-discretionary variables at the second-stage analysis by using the double-bootstrapping truncated regression method outlined by Simar-Wilson (2007). The results indicate that public efficiency has a significantly positive association with asset utilization but no significant association with budget accuracy, suggesting two important public policy prescriptions. First, there is a need for increased focus on asset utilization to combat the efficiency decline. Second, owing to a mismatch between accrual accounting and cash

budgeting, we prescribe the introduction of accrual budgeting for policymakers and local authorities alike to more strategically manage public assets.

We continuously expand our study in two ways. On the one hand, the research on stratification of the geographical area in terms of public asset utilization between special wards and Tama cities presents evidence that, to maintain higher efficiency, special wards should cut down on public assets but Tama cities should consider more asset investment. On the other hand, a study on the decomposition of various assets indicates that while more investment in educational assets is required, other assets such living, hygiene, and industry assets should be given less investment in order to gain efficiency.

The salient policy implied in chapter 5 is to introduce the accrual budgeting system in Tokyo local governments. In principles, accrual budgeting is less likely to be effective without the support of performance management. Thus, in chapter 6, we aim to study a performance measurement system (PMS) *vis-à-vis* accrual budgeting – particularly the use of the PMS – by embracing the mixed-method sequential explanatory study (using a Structural Equation Model (SEM) at the first stage and semi-structured interviews at the second stage). We found that the PMS is virtually used for incentive-oriented use rather than exploratory use. We suggest that a shift toward strategic exploratory use of PMS is necessary to effectively implement accrual budgeting.

To sum up, this thesis delves into four main research topics along with four essential NPM concepts – fiscal decentralization, public efficiency, private sector accounting norms, and performance management for results – subject to 49 Tokyo local governments over 2001–2015.

Table of Contents

Figures

Tables

Abbreviations

BSC	Balance Scorecard
CB-SEM	Covariance-based Structural Equation Modeling
CLAIR	Council of Local Authorities for International Relations
CRS	Constant Return to Scale
DEA	Data Envelopment Analysis
DPJ	Democratic Party of Japan
FE	Fixed Effect
GFC	Global Financial Crisis
GPEA	Government Policy Evaluation Act
GPRA	Government Performance and Results Act
KPI	Key Performance Indicator
LAT	Local Allocation Tax
MLE	Maximum Likelihood Estimation
MIC	Ministry of Internal Affair and Communication
NPM	New Public Management
OECD	Organization for Economic Co-operation and Development
OLS	Ordinary Least Square
PA	Public Administration
PM	Performance Measurement
PMS	Performance Measurement System
PLS-SEM	Partial Least Square Structural Equation Modeling
PSO	Public Sector Organization
RE	Random Effect

SBM   Slack-Based Model

SFA   Stochastic Frontier Analysis

TMG   Tokyo Metropolitan Government

VfM   Value for Money

VRS   Variable Return to Scale

CHAPTER 1: INTRODUCTION

1.1 Introduction

In light of public management reforms since the 1980s, the notion of 'New Public Management' (NPM) has increasingly garnered special attention among scholarly communities and practitioners across the world. As noted, NPM paradigm is generally a manifestation of a necessity to reshape public sector organizations (PSOs) following private managerial practices (Dunleavy & Hood, 1994). Accordingly, two common features associated with the NPM paradigm are competitive markets and adoption of business-like management techniques (Hood, 1995; Speklé & Verbeeten, 2014). While the mechanism of competitive markets can stimulate more rational and efficient decision-making in individual agents, the adoption of business-like management techniques such as the Balance Scorecard[1] (BSC) can make PSOs more professional in operation and management (Hood, 1995; Speklé & Verbeeten, 2014; Ter Bogt, Budding, Groot, & Van Helden, 2010). Therefore, the NPM paradigm can influence PSOs to be more rational, efficient, and professional.

A transformation from the traditional public administration (PA) to NPM potentially instigates public sector performance and authority enhancement (Dunleavy & Hood, 1994; Hood, 1995; Moynihan, 2008). First, regarding public sector performance, Moynihan (2008, p. 27) asserts that PSOs have prolonged exposure to poor performance and low efficiency driven by the traditional PA system. This is because public officials are likely to be relatively disinterested in pursuing improvements in organizational performance unless they are motivated by political gains or benefits (Boyne, 2002; Bandy, 2015). Furthermore, in budgeting procedures, traditional

---

[1] The Balance Scorecard was proposed by Kaplan and Norton (1992) as an essential instrument of performance measurement, which was widely adopted in the private sector.

PA practices an input-oriented approach to spending rather than one based on outputs and outcomes; therefore, budget holders (public managers) barely eliminate budget slacks to enhance outputs and outcomes and might not strategically prioritize expenditure, resulting in less efficiency in public services and goods provisions. Second, regarding authority enhancement, public officials hold less decision-making authority in financial affairs. For instance, they are constrained by the budgeting appropriation, meaning that spending on public goods and services is confined to the budgeted amount approved at the beginning of the fiscal year. Hence, PSOs typically display relatively poor performance and have relatively less authority over financial affairs.

Despite poor public performance and less authority in the traditional PA system, NPM facilitated a radical change to improve public performance and authority (Hood, 1995). Indeed, PSOs have become more rational and efficient as a corollary of employing innovative public sector accounting systems, such as accrual-based accounting, and designing and implementing a management-for-results instrument such as the performance measurement system (PMS) (Moynihan, 2006, 2008, p. 27). The superior adoption of these novel instrumental practices can provide decision-makers and public officials with greater quantities of performance information and releases constraints on authority to some extent, encouraging them to be more efficient, effective, and accountable for dealing with an array of demands from stakeholders. Briefly, the NPM reform has motivated a change from the traditional PA system (low public performance and managerial authority) into performance management practices (higher public performance and managerial authority).

Therefore, the NPM reform can urge PSOs to improve public efficiency and accountability (Moynihan, 2008). Public efficiency is composed of technical and allocative

efficiency, while public accountability involves internal and external accountability. First, regarding efficiency, technical efficiency is simply the maximization of outputs with given inputs or minimization of inputs with given outputs, while allocative efficiency refers to the matching of public service supply and demand (Andrew & Entwistle, 2014, p. 5). In the face of economic recession in the last decade, because of limited income sources, PSOs have come under pressure to *do more with less* for resident demands. Second, regarding public accountability, external accountability describes the extent to which the public and citizen request the transparency of performance information from public governments while internal accountability holds bureaucrats or public servants accountable for performance information in response to elected officials (Moynihan, 2008, p. 36). Consequently, the notion of efficiency and accountability improvement becomes salient in NPM reforms.

In doing so, the public sector accounting innovation is the central of public sector reforms (Guthrie, 1998; Hood, 1995). This innovation entails the availability of financial and non-financial information in terms of quantity and quality among PSOs, people, public servants, and elected officials. In other words, this innovation demonstrates the need for 'accountingization' to distribute performance information amongst these stakeholders. Hence, a shift from traditional financial reports toward a more comprehensive picture of the financial position is essential in public management. Since the 1990s, some advanced economies such as New Zealand, Australia, and the UK have adopted accrual accounting regimes in financial statements supplemented by the accrual budgeting mode in strategic planning and controlling of local incomes and expenditures (Blöndal, 2004; Martí, 2013; Robinson, 2009; Warren, 2015). Other following adopters such as Canada, the USA, and Japan embraced this accounting regime later. The most considerable benefits of accrual-based accounting are that public authorities and

policymakers are capable of grasping more comprehensive information of the financial situation in regards to the medium and long term (for example, public assets, pensions), as opposed to cash-based accounting which only provides financial data for the specific fiscal year (Warren, 2015). Therefore, accrual basis adoption and implementation might be beneficial for users interested in increasing public efficiency and accountability − particularly improving more rational public asset management, accordingly leading to greater performance (Carlin, 2005; Robinson, 2009).

PMS focusing on measuring and monitoring performance is an essential instrument to underpin the accrual basis in financial reporting and budgeting (Martí, 2013; Schick, 2007). Performance budgeting is considered a perfect tool in the NPM era (Bandy, 2015, p. 68) because, compared to incremental and zero-based budgeting (frequently used in traditional PA) that aim at controlling inputs, performance budgeting encourages public servants to focus on outputs and outcomes instead. The performance information generated by the PMS could be integrated into the strategic planning and budgeting for the fiscal year ahead. Additionally, performance information could be used to evaluate and monitor local spending under the constraints of local incomes and helps users link performance targets and budgeting (Robinson, 2013). Thus, performance information generated by the PMS can facilitate not only *ex-ante* management activities (e.g. planning and budgeting) but also *ex-post* actions (e.g. evaluation, controlling, monitoring) in PSOs.

To sum up, this study highlights some prominent aspects of the NPM paradigm and delves into four main studies with regards to a Japanese prefectural government. The first study examines the delegation and devolution of revenue and expenditure powers from the higher tiers of government to local governments and the role of intergovernmental grants as an instrumental

policy. The second study examines public technical efficiency in the face of the financial crisis and public sector accounting change. The third study improves allocative efficiency and accountability through the prescription of accrual-based accounting in public asset management at the local level. The final study is the design and implementation of PMS in accordance with accrual-based accounting and budgeting system.

1.2 Research motivations

This section outlines the underlying motivations for this research. First, there is a dearth of academic works on public efficiency and performance associated with public finance management in Japanese municipalities as opposed to other advanced economies in Organization for Economic Co-operation and Development (OECD) countries (e.g., United Kingdom, Australia, New Zealand, Italy, Portugal, Germany, Belgium). Second, the Japanese government has embarked on the engagement of NPM reforms since the 1990s. Particularly, PA changes in the political lieu at the beginning of the 2000s under the Koizumi administration influenced the local administration. Moreover, the 2008 Global Financial Crisis (GFC) also adversely impacted the socio-economic condition of local governments. However, research on these effects on Japanese local governments seems to be limited and where research has been undertaken, it has been restricted to a regional or prefectural rather than local level. Finally, finding determinants of public efficiency and performance can be a benefit for local authorities and policymakers to have good policy prescriptions. These arguments are discussed in the following two subsections.

1.2.1 Research Gaps.

There are two research gaps in the corpus of literature that explain why Japanese local governments are primarily the focus of this study. First, in the wake of NPM reforms, Anglo-sphere countries (e.g., United Kingdom, Australia, New Zealand) – forerunners of this evolution

in the late 1980s – and followers in OECD countries (e.g., Netherland, Germany, Italy, Portugal, United States) have conducted administrative shifts to be more efficient and accountable. Myriad research on public sector finance and accounting, particularly fiscal decentralization, has been considerably proliferated in recent decades in these countries. For example, existing studies on public spending volatility are presented by Sachi and Salotti (2017) in Italy, Furceri (2007) for OECD countries, Denison and Guo (2015) in American states, and Albuquerque (2011) in Europe. Those on public revenue volatility have been illustrated by Afonso (2017) in North Carolina counties and Staley (2015) in American states. With respect to public performance research, a range of findings have been identified in OECD countries. For instance, Storto (2013) evaluated the technical efficiency of major municipalities in Italy. Afonso and Fernandes (2006) studied the efficiency of 51 Portuguese municipalities. Doumpos and Cohen (2014) analyzed the efficiency of Greek local government based on accrual accounting data. Cuadrodo-Ballesteros et al. (2014) showed the relationship between public services delivery and efficiency in Spain. Conversely, research on the same subject has been limited in Japan, although Nijkamp and Suzuki (2009) delineated the efficiency score for cities in Hokkaido prefecture and Fukuyama et al. (2017) raise the issue of public efficiency for nationwide estimation at the prefectural level. However, studies on Japanese local governments, specifically Tokyo local governments, are scarce.

Second, in the contextual situation, Japan municipalities have ever experienced dozens of far-reaching changes impacted by national policies and environmental shifts. From a national perspective, Japanese scholars and practitioners paid great attention to applications of NPM paradigms at national and local government levels in the mid-1990s under the Hashimoto administration (1996–1998), notably focusing on policy evaluation and 'agencification' (or fiscal

decentralization) (Yamamoto, 2003). The Law for the Promotion of Decentralization was enacted in 1995, followed by the Government Policy Evaluation Act (GPEA) in 2002 (Mochida, 2008). During the Koizumi administration in 2004–2006, the 'Trinity Reforms', relating to the innovation of a local allocation tax (LAT) grant, a national subsidy, and local tax, rendered local governments more flexible in gaining own revenues (Mochida, 2008). Conversely, Japan municipalities were adversely impacted by the economic downturn in 2008, causing the local financial situation to deteriorate. Despite these events, there is limited empirical evidence on the internal and external impacts on Japanese local governments. Subsequently, we aim to investigate these impacts on Tokyo local governments mainly since 2008.

1.2.2 Potential prescriptions for public policies.

Since Tokyo local governments have been confronted with challenges to improve public efficiency and accountability due to the increasing demands of the people, an investigation of public efficiency and performance in this study may provide four dimensions of critical considerations. First, local authorities should take into account restructuring the components of revenues, intergovernmental grants in particular, to resist budget volatility in local governments. Second, the reevaluation of public efficiency and asset utilization is required to adjust the provisions of public goods and services associated with public assets. Third, the current cash-based budgeting system needs to be upgraded in order to strategically support performance budgeting. Finally, in accordance with performance budgeting, a performance measurement system should be examined to identify what factors drive its adoption and implementation. Therefore, it is noteworthy that this empirical evidence can fill the research gap in Japanese literature and contribute to the formation of policy implications for both local authorities and policymakers alike.

7

1.3 Research objectives

This study aims at three basic research objectives. First, we investigate into four main principles of NPM paradigms into Japanese local context, where limited studies have been scholarly conducted. Four NPM principles − a shifting from unified management system toward decentralized units, a pursuit of 'do more with less', an adoption of private sector accounting norms, and explicit standard and measure of performance in public budgeting, go along with four research questions below. Second, we expect to contribute to literature in terms of NPM practice in an industrialized Asian OECD country. We further examine the fiscal difference between sub-regions in this context. Third, because of fiscal austerity since the late 1990s, both national and local governments had some policy changes (e.g., an adoption of accrual accounting mode in municipalities), but its effects on public organizations have been not analyzed. This study attempts to explore the effects and results of these changes, and propose some important public policy amendments for local governments.

Accordingly, we set out four principal objectives with regards to public finance management, public service provisions efficiency, public asset utilization, and public performance measurement in tandem with eight research questions, as follows:

Public finance management

1. Which components of the volatility of revenues and grants are associated with local spending volatility in Tokyo municipalities?

2. Are the fiscal systems (taxes, grants, etc.) different between urban (special wards) and suburban (Tama cities) areas in the Tokyo metropolis, due to the different fiscal arrangement system in the face of the financial crisis?

Public service provisions efficiency

3. What are the trends or patterns of public (technical) efficiency scores of Tokyo local governments over the 2001–2015 period, before and after the introduction of accrual accounting?

Public asset utilization

4. What is the impact of asset utilization on municipal efficiency in Tokyo local governments since the introduction of accrual accounting and compare the effect of accrual accounting mode as compare to cash accounting on public efficiency improvement?

5. What is the role of current budgetary control (cash budgeting) functions for financial stability or efficiency improvement?

6. What types of assets significantly impact municipal efficiency?

Public performance measurement

7. What are the driving factors to determine PMS use in case of Tokyo Metropolitan Government (TMG)?

8. Is the adoption and implementation of PMS for incentive or exploratory use?

We attempt to solve these research questions in chapter 3 (questions 1 and 2), 4 (question 3), 5 (questions 4, 5 and 6), and 6 (questions 7 and 8), respectively.

1.4 Research methods

This study seeks to explore and explain the contemporary condition under which public sector accounting and finance regimes in a specific jurisdiction have evolved over a certain

period. Thus, positivism befitting realist ontology is best suited for this aim (Easterby-Smith, Thorpe, & Jackson, 2012). The empirical-analytical approach, 'gathering and interpreting knowledge about real life in an objective manner' (Van Thiel, 2014, p. 33), was intentionally used for research design and procedures. Utilizing a solely quantitative method, however, can be a shortcoming because a deterministic model of existing theories testified over a certain period hardly seems to express realistic situations over a long historic period and a limitation of data relevance is problematic. Therefore, we mixed the deductive and inductive approaches but emphasized deductive ways. The research methods conducted in this study involve econometric analysis, two-stage DEA, and mixed research methods.

1.4.1 Econometric analysis of panel data.

The panel data econometric method involves estimating the association between the dependent variable, namely expenditure volatility, with various independent variables in this study. Panel data refers to the pooling of observations by cross-sectional units over several periods, also called cross-sectional time-series data or longitudinal data (Baltagi, 2005). According to Baltagi (2005, p. 4), panel data techniques can 'control for individual heterogeneity' and 'give more informative data, more variability, less collinearity among the variables, more degrees of freedom and more efficiency'.

In panel data methods, the unobserved effect due to the heterogeneity of observations accounts for a discrepancy between the fixed effects (FE) model and random effects (RE) model. As defined by Wooldridge (2013), the FE model refers to a pooled Ordinary Least Squares (OLS) regression based on time-invariant variables, where the unobserved effects can be correlated with the explanatory variables. The FE model treats unobserved heterogeneity as a part of the intercept which varies over each period. Contrarily, the RE model is defined as a

panel data model in which unobserved effects are assumed to be uncorrelated with the explanatory variables (Wooldridge, 2013). The RE model considers unobserved effects as an error component of the composite error term. While the RE model can estimate the effect of time-invariant variables, the RE model faces a potential problem related to the inconsistency of estimated parameters in the presence of correlation and cannot reliably estimate unobserved heterogeneity. Though the FE model is less efficient than the RE model, we can eliminate the possible variance by stratifying observations into multiple groups (Drew & Dollery, 2016). However, Kennedy (2008, p. 286) argues that RE estimators are biased because individual effects are possibly correlated with independent variables whereas FE estimators are not biased. Hence, the FE approach is a better choice for estimation.

1.4.2 Data Envelopment Analysis (DEA)

There are two approaches to measure efficiency in the economic theory of production: parametric and non-parametric methods. Parametric techniques include Stochastic Frontier Analysis (SFA), Ordinary Least Square (OLS), and Fixed effects (FE) regressions, whereas the non-parametric techniques are Data Envelopment Analysis (DEA) and m-order. In our research, we adopt the DEA because of its popularity in measuring municipal efficiency (Worthington, 2000; Storto, 2016)

DEA measures typically rely on mathematical linear programming techniques. Farrell's seminal work (1957) provided a definition of technical and allocative efficiency in production. On the one hand, the technical component refers to the ratio of outputs and inputs in the production. The tactic to improve technical efficiency is either to eliminate resources (inputs) translated into goods or services (outputs) with the given outputs and technology or to maximize productive outputs as much as possible with the given inputs and technology. Thus, the analysis

of technical efficiency is relative to the input-oriented (waste elimination) or output-oriented (result maximization) approach. On the other hand, estimating technical efficiency under the condition of prevailing prices is referred to as the allocative component (Fried, Lovell, & Schmidt, 2008).

DEA computes relative efficiency scores under the assumption of convexity. It means that observed efficiency scores are convex or bounded by a technology frontier or boundary (data are enveloped by the estimated boundary, so-called Data Envelopment Analysis) (Charnes, Cooper, & Rhodes, 1978). The Constant Returns to Scale (CRS) DEA model with the input orientation was initially proposed by Charnes et al. (1978) (hereafter referred to as the CCR model). The CRS assumes that an increasing (or decreasing) input can proportionally produce an increasing (or decreasing) output. However, as this phenomenon scarcely occurs in the real world, Banker, Charnes and Cooper (1984) (hereafter referred as the BCC model) have extended the CCR model into the Variable Returns to Scale (VRS) model where the weights between outputs and inputs are constrained. Since the introduction of the CCR and BCC model, there has been an exponential growth in the corpus of literature focusing on theoretical and empirical studies in various fields (e.g., education, health) in both private and public sectors. In particular, the employment of the DEA method to evaluate public efficiency for benchmarking public entities has garnered considerable critical attention (Worthington, 2000; De Borger & Kerstens, 1996).

DEA measures the relative economic performance among entities but cannot explain the determinants affecting their efficiency. The second stage of analysis was introduced to regress potential environmental factors on the efficiency estimated in the first stage (also called two-stage DEA). In the second stage, an array of such methods has been feasibly adopted, such as

OLS, COLS (Corrected OLS), MLE (Maximum Likelihood Estimation), Tobit regression, and truncated regression with bootstrapping procedures. Banker and Natarajan (2008) advocate the parametric regression estimation (OLS and MLE, or Tobit regression) at the second stage since these procedures can yield consistent estimators of the impact of contextual factors. In contrast, Simar and Wilson (2007) are in favor of the truncated regression with a bootstrap algorithm by using a coherent data-generating process (DGP), stating that this approach can avoid the serious problems in the second stage analysis that involves serial correlation among estimated efficiency and the creation of boundary issue (whereby efficiency scores cannot exceed 1 and are hence constrained when regressing against contextual factors). The bootstrap method can generate statistical confidence intervals by resampling data, resulting in more feasibly consistent inference (Simar & Wilson, 2007). In this study, we employ the DEA to estimate efficiency scores and truncated regression with double bootstraps to regress various environmental variables on estimated efficiency.

1.4.3 Mixed research methods

In the study of performance measurement and management, we prefer to adopt mixed research methods, particularly explanatory sequential mixed methods. As defined, mixed research methods combine both qualitative and quantitative data collection and analysis (Creswell, 2014). Both are mutually complementary in research interpretation since separate use of either quantitative or qualitative methods cannot provide comprehensive research results. We therefore select an explanatory sequential mixed method where we employ a quantitative method, particularly Partial Least Squares Structural Equation Model (PLS-SEM), in the first procedure, followed by a qualitative method with interviews.

Structural Equations Modeling (SEM) is a family of multivariate techniques that seeks to simultaneously investigate a set of dependence relationships (Hair et al., 2009). Different from traditional multivariate methods (e.g., multiple regression), SEM can appropriately explain the separate multiple regression equations among multiple independent and dependent variables, even incorporation of mediating and/or moderating effects. Two types of models are estimated: (1) the structural (inner) model and (2) the measurement model (outer) model. The structural model illuminates the path diagram among variables, whereas the measurement model is exposed to the indicators or items reflecting the latent variables or constructs. It is noteworthy to outline the differences in the reflective and formative measurement model. Regarding the reflective measurement model, indicators represent the effects of a latent variable or construct. All the indicative measures reflect the character of the underlying construct. In contrast, indicators form the underlying construct by the casual effects from indicative measures to the construct in the formative measurement model (Hair et al., 2017).

SEM analysis includes two types of multivariate techniques: (1) Covariance-based SEM (CB-SEM) and (2) Partial Least Squares SEM (PLS-SEM). While the CB-SEM is widely used to empirically test a set of relationships supported by theories (in other words, it is employed for confirmatory purposes), PLS-SEM is primarily used for exploring a set of relationships to develop theory (in other words, it is used for explanatory purposes) (Hair et al., 2017). Hence, PLS-SEM is widely taken into consideration in the case of ongoing theory development, such as that involved in management accounting research (Nitzl, 2016). According to Chin (2009), PLS-SEM is preferable over CB-SEM due to several advantages. First, PLS-SEM relieves the hard assumptions of distribution and independence of observations associated with CB-SEM and utilizes soft distributional assumptions. Second, while CB-SEM depends on full information

maximum likelihood techniques to generate correct parameters estimated, PLS-SEM can employ a limited-formation by using least-squares algorithms. Thirdly, as a result, PLS-SEM's sample size requirements could be smaller relative to those of CB-SEM. In fact, Chin and Newsted (1999) state that 'minimum recommended ranges from 30 to 100 cases' for PLS analysis. Subsequently, due to its superior features, we opt for the PLS-SEM technique in performance management research in this study.

In the second stage, in order to attain a deeper understanding of interrelated associations among constructs, we design and implement interviews by selecting some of the participants in the first stage to respond to further semi-structured questionnaires. We use content analysis and conceptual mapping to analyze verbal data. Taken together, we adopt mixed research methods, including SEM analysis and semi-structured questionnaires, in this study.

1.5 Structure of the thesis

In order to have a profound insight into the public finance, public sector accounting, and public performance management of Tokyo local governments, we deliberately separate this dissertation into seven sections as outlined below (see Figure 1).

Chapter 1 introduces the research background in public efficiency and performance-oriented management simulating the business management style in the wake of public management reforms, urging many countries across the world to carry out innovations and improvements in public finance, public sector accounting, and public performance management. Particularly, with a primary focus on Tokyo local governments in Japan, we raise research motivations, research gaps, and potential policy implications for local authorities. We outline eight contemporary research questions that we address in four main chapters (Chapters 3-6).

Chapter 2 presents a scholarly corpus of literature regarding the research topics. First, we introduce NPM principles from international perspectives and, particularly, NPM in Japan. Second, in the field of public finance, we synthesize the theory of fiscal federalism and decentralization, which support the research of budgetary volatility in Chapter 3. Third, the public efficiency and public sector accounting innovations are introduced. Finally, we depict PMS in the field of management accounting and focus on its implementation and use. In addition, underlying theories such as agency theory, institutional theory, and contingency theory are introduced to support the use of PMS.

Chapter 3 aims at solving the problem associated with the impact of revenue volatility on expenditure volatility in the case of 49 Tokyo local governments. The public incomes, including various taxes, intergovernmental grants, fees, charges, subsidiaries, and local bonds, could potentially affect public expenditure in the financial austerity aftermath of the GFC. The volatility of many types of incomes makes the spending decisions of public authority harder in terms of pursuing fiscal sustainability and stability. It is necessary to fine-tune some incomes to smooth out the spending volatility.

Chapter 4 merely introduces the empirical analysis of public efficiency scores in 49 Tokyo local governments. Based on the logic production model, in which inputs are local demanding expenditures (described in Chapter 3) translated into public goods and services provision, we estimate the public efficiency by the means of Data Envelopment Analysis (DEA) over the 2001–2015 period. We also benchmark the efficiency scores prior to and after the introduction of the accrual accounting system in Tokyo local governments in 2008.

Chapter 5 extends the research in Chapter 4 by implementing the second stage analysis after DEA estimation. We recognize the decline of public efficiency in 49 Tokyo local governments since the introduction of the accrual accounting system in 2008. We examine the regression of public efficiency on various environmental variables and identify which determinants affect the public efficiency scores by using the Simar-Wilson (2007) approach. The analysis results in terms of cash-based and accrual-based accounting are presented. We notice that the budgeting in accrual-based accounting is better than that in cash-based accounting. Therefore, we recommend the adoption of accrual budgeting in public management reform. We also deconstruct the study on public asset utilization into special wards and Tama cities and investigate the various assets affecting efficiency scores.

In the NPM, the key benefit of accrual budgeting might be vis-à-vis performance budgeting, provided a justification for PMS has been carried out in the public sector. Thus, Chapter 6 studies the use of PMS and its driving factors at TMG and local governments. We employ mixed research methods, particularly sequential exploratory research, in which we use the PLS-SEM at the quantitative stage and then semi-structured interviews to get insights into PMS use at TMG and its local level.

Finally, Chapter 7 provides concluding remarks associated with smoothing out expenditure volatility, the decline of public efficiency scores, the importance of asset utilization, and the incentive use rather than exploratory use of PMS. Following this, we prescribe some public policy implications for local governments as well as the prefectural local government (TMG). Some research limitations and future studies are additionally described in the final section.

| Chapter 1
Introduction | **Research backgrounds** | **Research motivations** | **Research objectives** | **Research methods** |

Chapter 1
Introduction

Research backgrounds　**Research motivations**　**Research objectives**　**Research methods**

Chapter 2
Literature Review

Public finance and localism　**Public sector accounting**　**Performance measurement**

Chapter 3
Budgetary Volatility

Revenue and expenditure volatility　　　　**Fixed-Effects panel data**

Chapter 4
Public Efficiency

Expenditure as inputs

Public efficiency estimation　　**Data Envelopment Analysis**

Chapter 5
Asset Utilization

Asset utilization　　**Truncated regression with bootstrap**

Chapter 6
Performance Measurement

Accrual Budgeting

Performance measurement system use　**PLS-SEM and interviews**

Performance Budgeting

Chapter 7
Conclusions & Implications

Conclusion remarks　▸　**Policy implications**

Note.

◂ – – – – –　: Indicating the research method used in each chapter

◂————　: Indicating the underlying theories and studies built for ahead subjects

Figure 1. Structure of the thesis

CHAPTER 2: LITERATURE REVIEW

This chapter features a discourse on literature relevant to this thesis. First, we analyze the doctrine of NPM in an international and local context. Second, we discuss the theory of fiscal federalism in public finance and localism, mainly concentrating on the principles of fiscal decentralization and fiscal volatility at the local level. Third, in relation to the notion of fiscal decentralization presumably causing the Pareto-efficient enhancement of public service provisions, we emphasize municipal efficiency, resource allocation, expenditure priority, and, particularly, private sector accounting norms. Fourth, the body of knowledge on performance measurement and management associated with performance budgeting is introduced. Finally, we encapsulate the theories that provide the guidelines for our research.

2.1 New Public Management (NPM)

2.1.1 NPM backgrounds.

The concept of NPM has become a predominant doctrine in public sector reforms across the globe, garnering international scholarly attention since the 1980s (Lapsley, 1999). NPM paradigms lead to greater public efficiency, effectiveness, and accountability (Dunleavy & Hood, 1994; Hood, 1995; Lapsley, 1999). Initially, UK Prime Minister Margaret Thatcher adopted the NPM model in her government in the late 1980s, from where it widely dispersed to advanced economies such as Australia, New Zealand, the USA, and Canada. A decade later, the NPM principles *per se* were found outside the border of the Anglosphere, in countries such as Japan and developing countries in Asia (Diefenbach, 2009; Hood & Dixon, 2013).

The principal objectives of public sector management NPM are similar to those used in private sector management, with two basic management doctrines (Hood, 1991, 1995). First, a shift from traditional public management towards private sector management will avoid the

bureaucratic procedures, wastes, and incompetence typically associated with public sector practices (Hood, 1995). For example, PSOs have embraced business management techniques such as Balance Scorecard (BSC) and Key Performance Indicators (KPIs) as essential management tools. Second, the NPM theory inspires public authorities to actively seek more autonomy in spending, staffing, and contracting out (Dunleavy & Hood, 1994; Hood, 1995). As a result, Hood (1995, p. 96) proposed seven critical principles of the NPM theory as follows: (1) 'unbundling the public sector into corporatized units organized by product', (2) 'more contract-based competitive provision, with internal markets and term contracts', (3) 'stress on private sector management styles', (4) 'more stress on discipline and frugality in resource use', (5) 'more emphasis on visible hands-on top management', (6) 'explicit formal measurable standards and measures of performance and success', and (7) 'greater emphasis on output controls' (Hood, 1991, 1995). Public sector reforms over the past decades have adopted these seven principles, bringing about positive consequences of efficiency, effectiveness, and accountability improvement.

The NPM theory has continuously evolved with the extensive body of literature. Indeed, Dunleavy et al. (2006) encapsulated the NPM as disaggregation, competition, and incentivization. The NPM possibly appears at two levels. At the doctrinal (or high) level, NPM undertakes business concepts, techniques, and values toward public sector management. At the mundane (middle) level, NPM emphasizes output-oriented measurements, disaggregated organizations, substitution of contracts for hierarchical relations, market-type mechanisms, and treatment of public service recipients as customers (Dunleavy et al., 2006). Diefenbach (2009) rendered the basic NPM principles regarding business environment and strategic objectives more explicit by categorizing outside- and inside-orientations. For external determinants, PSOs are

driven by the market (providing public services and goods with 'value for money' (VfM)), stakeholders (being influenced by external parties), and customers (supplying satisfactory products). For internal management, PSOs concentrate not only on boosting efficiency, effectiveness, and productivity but also increasing 'cost-effectiveness' (or running cost-cutting measures) (Diefenbach, 2009). Furthermore, Pollitt (2016) suggested mixing the NPM model and managerialism principles to unanimously support public management reforms.

Although the NPM doctrine has been continually developed, certain undisputable features such as fiscal and administrative decentralization, efficiency and effectiveness improvement, public sector accounting innovations, and performance management have remained largely unchanged. Being compatible with local context (history, culture, politics, and economics), NPM principles vary from country to country. In our research scope, we characterize these principles in terms of Japanese local governments.

2.1.2 NPM in Japan

In the wake of NPM, Japan's democracy has embarked on central and local administrative reforms since the late 1990s with some unique features among OECD countries (Kudo, 2003; Yamamoto, 2003). Apart from English-speaking nations who positively adopt the NPM, Japanese governance has put less emphasis on NPM in some ways (Eshima, Katayama, & Ohno, 2001; Hood, 1995; Yamamoto, 2003). Firstly, Hood (1995) insists that Japanese public management reforms accentuate privatization, deregulation, and tax reform, which are slightly different from the aforementioned NPM principles. Secondly, while public management reforms in New Zealand and the UK promote the NPM through a market-oriented mechanism, those in Japan highlight social equity and customer needs rather than market forces (Eshima et al., 2001). Thirdly, New Zealand and the UK exercise 'make manager manage' in contrast to Australia and

Sweden who 'let manager manage'. In this regard, Japanese public management has followed the

second line of practice ('let manager manage' but with less flexibility and responsibility)

(Eshima et al., 2001, p. 3). Fourth, Japanese public administration has high opportunity but low

motivation shifting toward NPM (Yamamoto, 1999). Lastly, Hori (2004) asserts that bureaucrats,

instead of politicians, have driven Japanese public management reforms. Subsequently, Japanese

public management is more prone to new public administration rather than NPM.

In Japanese administration at the central level, two salient actions, policy evaluation and

agentification, may influence local governments. First, regarding policy evaluation, the National

Diet enacted the Government Policy Evaluation Act (GPEA) in 2002. The GPEA is a guideline

for governmental agencies in evaluating policy and performance in the executive branch (Talbot,

2006; Yamamoto, 2003). Second, in agentification, autonomous or semi-autonomous agencies

are delegated to lower levels as representatives of ministerial organizations to implement several

specific functions. This system was applied in Japan following the UK practice (slightly altered

for the Japan context) (Yamamoto, 2003). In performance management, the UK model focuses

on the executive branch while the US model concentrates on the legislative branch. The Japanese

case sets performance evaluation in the middle of the UK and US extremes. Moreover, unlike the

UK and US, which concentrate on performance, Japan undertakes not only performance but also

evaluation (Yamamoto, 2003). Accordingly, Japanese NPM-oriented governance seems to be

unique across the world due to the differential environmental context.

Due to the particular local context of Japanese public administration, a shifting from

theoretical NPM doctrines to practical issues is subject to some necessary conditions. First,

bureaucrats or public servants have tendency to oppose to public administrative reforms

(Yamamoto, 2003). Even if readily accepted to these reforms, they seemed to have low incentive

to new adoption (Yamamoto, 1999). Hence, some governmental rules and regulations should be enacted to coercively urge bureaucrats to engage in NPM implementation. Second, NPM reforms entail public servants to have advanced knowledge such as accrual accounting or double-entry bookkeeping system, whereby elite groups of bureaucrats can grasp the understanding of these knowledge. Therefore, offering training courses for lower-level public servants is necessary to lessen the likelihood of rejection on NPM reforms adoption. Consequently, NPM becomes a global paradigm in public management, but not a panacea or "one-size-fit-alls". Some requisite conditions applied in a specific local governments are necessarily important.

In a nutshell, following the NPM wave, Japanese administration has eventually innovated in public management. NPM became an underlying theory for our four main research points, focusing on fiscal decentralization in public finance (chapter 3), public efficiency (chapter 4), public sector accounting reform (chapter 5), and performance management (chapter 6). The next three sections capture the background of our research.

2.2 Public Sector Finance

2.2.1 Fiscal federalism

Much attention has been paid to the theoretical framework of fiscal federalism, which has experienced two generations of theory, since the 1950s. The first generation lasted from the 1950s to the mid-1990s and the second dates back to the mid-1990s. Fiscal federalism refers to 'assignment of functions to different levels of government and the appropriate fiscal instruments for carrying out these functions' (Oates, 1999, p. 1121). Accordingly, the fiscal federalism theory seeks out an explanation for decentralization among countries. In general, four theoretical frameworks were developed with respect to fiscal federalism, including the sorting or Tiebout

model (Tiebout, 1956), Olson's fiscal equalization (Olson, 1969), Oates' decentralization theorem (1972), and Brennan and Buchanan's constraining Leviathan (1980).

2.2.1.1 First generation of theory.

Tiebout model (1956)

The Tiebout model derived from the book *A Pure Theory of Local Expenditure*, which did not receive much attention initially, but significantly contributed to the body of knowledge in fiscal federalism (Grant & Drew, 2017, p. 138). The Tiebout sorting model argues that consumers or voters can decide the place where they can satisfactorily reside in response to their demands. Regarding fiscal decentralization, the Tiebout model captures two salient features: community participation and inter-jurisdictional mobility. First, people are eligible to reject the power of local authorities via voting if they are unsatisfied. In other words, people have a 'voice' to reveal their preferences. Second, people can decide to move to a jurisdiction which they feel befits their needs and preferences, also known as 'voting with their feet'. Hence, people can 'exit' the place where tax is levied and public service provisions do not respond to their requirements (Litvack, Ahmad, & Bird, 1999). Accordingly, the Tiebout model promotes municipal separation, provision for mobility, and flexibility in the revenue-expenditure structure.

Olson's fiscal equivalence (1969)

Olson (1969) presumably observed what generates a discrepancy between the boundaries of collective goods and government boundaries, defined as 'fiscal equivalence'. Pareto efficiency would be not optimized if the collective good provision either exceeded or was under government boundaries. To achieve this Pareto optimal supply of public goods, the boundaries of collective goods should match the government boundaries. Therefore, to mitigate the mismatch

24

between two entities, it is important to delegate responsibilities from the central government to lower governments who can provide public services corresponding to their boundaries (Olson, 1969). Olson's seminal work (1969) addressed two prominent points. First, the government should provide a bundle of collective goods corresponding to various cultures and tastes towards different racial and ethnic groups. On doing so, fiscal equivalence and economic efficiency could be improved (Olson, 1969). Second, Olson (1969) proposed that the central government can offer grants and subsidies to lower governments, which can alleviate losses due to the external benefits of adjacent local governments. Overall, Olson (1969) advocated the need for assignment of responsibilities among different governmental levels (decentralization) to reach the Pareto-efficient optimal in public service provision but did not advocate for too extreme levels of centralization or decentralization.

Oates' Decentralization Theorem (1972)

In the same vein, Oates (1972) emphasized economy of scales and local preferences to establish the Decentralization Theorem. The theory, from the normative economic view, states that 'the level of welfare will always be at least as high if Pareto-efficient levels of consumption are provided in each jurisdiction than if any single, uniform level of consumption is maintained across all jurisdictions' (Oates, 1972, p. 54). Essentially, decentralized governments may obtain economic welfare gains in providing people with public services fitting for the locality rather than national uniform public services. This occurs, firstly, because local governments close to the people can understand and better comprehend local preferences. Secondly, local tastes and costs vary among jurisdictions, so setting the same services at the same prices for all local governments can have a negative effect. In short, fiscal decentralization is necessary for differential local preferences and cost in order to improve overall social welfare.

Leviathan constraints of Brennan and Buchanan (1980)

Brennan and Buchanan's work (1980), *The Power to Tax*, further clarified the realistic concept of the Tiebout model's inter-jurisdictional mobility across jurisdictions in the presence of a competitive market. Brennan and Buchanan's model (1980) proposed the federation as a way to avoid the Leviathan government (the government intends to maximize revenue by coercively taxing the people). A feasible approach to prevent the revenue-maximizing Leviathan is to introduce decentralization, whereby the lower tiers of governments attempt to induce more taxpayers to reside in their local area with lower taxation. Thus, there is a competition among local governments with respect to tax power. Brennan and Buchanan (1980) posit that smaller decentralized governments are associated with economic efficiency.

Briefly, the underlying theoretical frameworks – Tiebout model, fiscal equivalence, decentralization theorem, and Leviathan hypothesis – which were the foundation of public finance theory, particularly fiscal federalism at the first generation, are in favor of decentralization. In vogue, the second-generation theory of fiscal federalism has attracted much scholarly attention.

2.2.1.2 Second generation of theory

The second generation of fiscal federation literature has remained somewhat controversial. Indeed, Qian and Weingast (1997) initially ignited the second-generation economic theory of federalism, suggesting that federalism with an appropriate level of decentralization from the central to local government could provide efficient public goods and also preserve the market. Qian and Weingast (1997) proposed a public finance mechanism of rewards or punishment in case of economic success or failures in an attempt to preserve market

incentives. If failing, the government sought to bail its lower tier of governments out of fiscal difficulties. This reaction is described as a 'soft budget constraint' problem.

Oates (2008) extended the practical issue of soft budget constraint as the first approach of second-generation theory. This line of thought explores the structure of incentives at decentralized levels of government, in which local authorities have exhibited deliberately detrimental and perverse behavior in running up deficits and accumulating a stock of debts. The local market and its neighbors also possibly go bankrupt and have an adverse effect on the national economy. Hence, to avoid the collapse of fiscal outcries at lower levels of government, the central government must support or bail them out. This was a typical experience in Japan; for instance, Yubari city in Hokkaido prefecture, which suffered from the exaggeration of sub-national borrowing, called for a municipal bailout in 2007 (Aoki, 2008; Mochida, 2008). Local authorities deliberately manipulated the general account by hiding the truth of snowballing debts in 2006, causing the fiscal insolvency (Hattori & Miyake, 2015; Mochida, 2008). Briefly, the first strand of the new theory focuses on soft budget constraint (what is different from hard budget control), whereby destabilizing fiscal behavior at the decentralized government level might precipitate an array of fiscal crises.

From the viewpoints of political economy, the second strand of the second generation of theory is likely to be more traditional and theoretical as compared to the first strand, which tends to be practical (Oates, 2008). In this strand, utility-making behaviors of public agents are motivated by fiscal decentralization, causing welfare gains and loss. Hence, the ultimate purpose of this conventional theory indicates the extent to which local government is decentralized by the central government, particularly 'a basic tradeoff between the gains from improved coordination

under centralization and the greater sensitive of local outputs to local tastes (or costs) and perhaps increased accountability under decentralization' (Oates, 2008, p. 322).

Collectively, two strands raise similar incentives of fiscal decentralization, but have different concerns. The first strand advocates the positive effects of fiscal decentralization in the public sector, while the second posits that fiscal decentralization can induce the rent-seeking behavior of public agents, causing social welfare gains or losses (Oates, 2008).

2.2.2 Fiscal decentralization

Fiscal decentralization refers to the devolution of taxation assignments and expenditure responsibilities from the central government to lower levels of government (Bird & Vaillancourt, 2009; Tanzi, 1995). In recent decades, a considerable number of advanced economies and developing countries have adopted the principles of fiscal decentralization at varying degrees of autonomy, according to fiscal and political institutions and contextual circumstances. There are three distinguished types of autonomous decision-making: de-concentration, delegation, and devolution. First, de-concentration refers to the 'dispersion of responsibilities' from the central government to provincial and state branch offices or local administrative entities (Bird & Vaillancourt, 2009; Litvack et al., 1999). This policy shifts responsibilities to the lower levels, but does not assign autonomy to them (common in many unitary systems) (Litvack et al., 1999). Second, delegation refers to the transfer of responsibilities associated with several functions from the central government to local governments. Alternatively, local governments act as public agents in executing certain public services and goods deliveries on behalf of the central government (Bird & Vaillancourt, 2009; Litvack et al., 1999). For instance, the Japanese government previously implemented a system of delegated functions at the local government level, but this system has been abolished since the Local Autonomy Law enactment (Ohsugi,

2009). Third, devolution refers to the assignment of authority corresponding to the transfers of responsibilities to local governments (Bird & Vaillancourt, 2009; Litvack et al., 1999). In this case, local governments have legal discretions on their own fiscal policy implementation to perform public functions.

Intriguing effects of fiscal decentralization

International experience has proven the importance of fiscal decentralization in bringing about welfare gains but, in some countries, fiscal decentralization makes the pursuit harder. In the corpus of literature on public choice, fiscal decentralization has been exposed to promote economic efficiency, macro-economic stability, fiscal equity (Litvack et al., 1999; Mclure & Martinez-Vazquez, 2000; Tanzi, 1995), and potentially foster economic growth (Martinez-Vazquez & McNab, 2003). First, as opposed to the central government (providing uniform levels of public service regardless of local demands), local governments (closer to the people and communities) can gain a better understanding of local needs and preferences. Better public provisions may be supplied, improving allocative efficiency (Oates, 2008). Furthermore, decentralization enables competition among jurisdictions in taxing and spending and tax-payers can decide the most suitable place for their preferences (Tiebout, 1956). Consequently, Pareto efficiency could be optimized. Second, though the macro-economic stabilization is attributable to national government duty, fiscal decentralization plays a role in mitigating fiscal volatility. Indeed, local officials – politically competitively voted electorates – have to design a balanced budget with respect to tax assignment and public spending, so they can be democratically re-elected by electorates (who are satisfied with this fiscal policy in taxing and spending). Thus, the local budget can be balanced. Typically, the less volatile the local budget, the more stable the macro-economic status. Third, fiscal decentralization can commit to interregional and

interpersonal equity by redistributing grants toward poor regions. When the poor suffer from the financial burden (higher tax but lower spending), the central government can lessen this fiscal gap by means of intergovernmental transfers to preserve fiscal equity for the poor (Litvack et al., 1999). Fourth, even though the theory of decentralization influencing economic growth is under-developed, Martinez-Vazquez and McNab (2003) propose that there is a direct and indirect nexus between these two factors, in which macroeconomic stability, economic efficiency, and the distribution of resources might be intermediating factors in this relationship.

Despite the aforementioned merits of fiscal decentralization, decentralized governments can face a number of fiscal difficulties concerning political and constitutional arrangements and destabilizing fiscal behavior. First, decentralization can lead to jurisdictional disparities of fiscal policies, which might overlap or be contradictory among jurisdictions. As a result, the central government faces a harder task in monitoring the fiscal situation as a whole (Tanzi, 2003). Second, according to the second strand of fiscal federalism, the rent-seeking behavior of public officials (maximizing their utility for their own purposes) might inherently impede fiscal policy prescriptions (jeopardizing stability and efficiency) for their constituency (Oates, 2008). Therefore, Tanzi (2003) concludes that fiscal decentralization becomes a good choice when the institutions associated with tax policy, expenditure responsibility, and budget are well-established and/or local governments become less reliant on the possibility of bailouts from the central government as a result of excessive spending.

Taken together, fiscal decentralization can be beneficial to welfare gains (mostly in developed countries), but also susceptible to pitfalls because of the weak institutional framework (mostly in emerging economies). Therefore, it is argued that '…decentralization is neither good

nor bad for efficiency, equity, or macroeconomic stability; but rather that its effects depend on institution-specific design.' (Litvack et al., 1999).

2.2.3 Fiscal reform in Japan

Since postwar 1945, public governance in Japan has experienced a series of reforms. From the postwar period to the 1990s, decentralization and centralization did not coexist, but neither were they profoundly antagonistic (Mochida, 2006). The interrelation between the central and local governments was implemented by public agents delegated by the central government, namely 'agency-delegated functions', regulated by the Local Government Law (article 150) (Mochida, 2006). Akizuki (2001) named this approach 'controlled decentralization'. However, the enactment of the Law for the Promotion of Decentralization in 1995 (the first stage of decentralization) abolished the system of delegated functions (Mochida, 2012, 2008, p. 1). Additionally, the deregulation and decentralization of administrative control and intergovernmental relations were innovated to relieve the burdens of local revenue on central government funding (Mochida, 2008).

In the second stage of decentralization, the Koizumi administration revitalized the remainder of the first stage fiscal reform and embarked on the 'Trinity Reform', aiming at the reduction of earmarked grants, local taxes, and local allocation tax within the 2004–2006 period (Mochida, 2008). Consequently, municipalities became reliant on central government funds and raised the capacity of revenue powers. Since 2009, the Democratic Party of Japan (DPJ) has set 'regional sovereignty' in motion by replacing the earmarked grants with lump-sum block grants in response to regional demands. The DPJ's policy was aligned with fiscal decentralization, possibly affecting the central government's support toward local government (Mochida, 2012). Thus, the degree of decentralization in Japan has been exposed to radical changes since the post-

31

war period, from 'controlled decentralization' to more decentralized functions through deregulation and de-concentration.

As a result, Japanese democracy, although regarded as a unitary system, attained a high degree of fiscal decentralization among OECD countries (higher than the US and Sweden and equal to Canada and Germany) (Mochida, 2008, p. 13). The ratio of central government to local government expenditure is 40 to 60; in contrast, the ratio of central government to local government revenue is 60 to 40. Hence, intergovernmental transfers necessarily reduce the fiscal gap in local governments, consequently resulting in fewer fiscal imbalances among prefectures (Mochida, 2008). Mochida (2008) concludes that Japan central-local relations play an important role in tax and expenditure harmonization, which is controlled by a higher level of government.

2.3.4. Fiscal decentralization in Tokyo local governments

Tokyo local governments have properly implemented fiscal decentralization reforms in central-local relations since the late 1990s. The national policy of 'Trinity Reforms' in 2004 navigated fiscal reforms in Tokyo, which focused on fiscal adjustment of earmarked grants, local taxes, and national subsidies (Mochida, 2008). However, local governments are still somewhat reliant on intergovernmental transfers, restricted in terms of tax administration, and subject to borrowing controls (Mochida, 2008).

Regarding the fiscal decentralization study, we have selected the Tokyo local governments, governed by the Tokyo Metropolitan Government (TMG) for several reasons (Tran et al., 2018). First, Tokyo is the locomotive of Japan's economy, contributing to 19.5% (JPY 92 trillion) of the country's Gross Domestic Product in 2014 – the highest portion of prefectural contribution to the national economy (Bureau of Finance, 2014). Second, Tokyo's

economy suffered due to the 2008 GFC. Local taxes plummeted by as much as 20% of the total revenue (JPY 1 trillion) in 2009. In response to the GFC, the TMG executed staff cuts and implemented controls to public spending in the pursuit of financial soundness (Bureau of Finance, 2014). Third, there are two disparate public administration systems within Tokyo local governments: 23 special wards (urban area) and 26 cities (suburban area, namely Tama). Under the Local Autonomy Law, while the former are categorized as special public entities – enjoying privileged financial coordinating grants from TMG to compensate for the fiscal gap – the latter are ordinary public entities, similar to other Japanese municipalities (Ohsugi, 2011). Additionally, the special wards are virtually limited in some public service responsibilities (e.g., fire prevention, water supply, and sewerage system), which are governed by the TMG on their behalf (CLAIR, 2013; Ohsugi, 2011). This unique fiscal system in Tokyo is the focus of our study.

At the local level, there has been growing interest in the volatility of local expenditure. For instance, Furceri (2007) maintained that expenditure volatility has a negative effect on long-run growth. In the OECD and EU countries, the volatility of indirect taxes, government spending, subsidies, and investment negatively impact economic growth (Afonso & Furceri, 2010). As the bigger countries or governments are capable of resilience, they tend to smooth out their public spending volatility (Albuquerque, 2011). In American states, revenue volatility and debt outstanding volatility act to lessen budget expenditures (Dension & Gou, 2015). Similarly, Sacchi and Salotti (2017) study how various revenue (e.g. income tax, sale tax, property tax, grant) volatilities affect local spending volatility in OECD countries. However, empirical studies on income and expenditure volatility are still scant in a Japanese local government context, particularly Tokyo municipalities. Hence, in the first main section of research, we aim to address

two research questions: (1) what are the components of revenues and grants associated with local spending volatility in Tokyo in the aftermath of the 2008 GFC? and (2) does the existence of the fiscal system (taxes, grants, etc.) differentiate between the urban (special wards) and suburban (Tama cities) areas in the Tokyo metropolis due to the different fiscal arrangement systems in place? To answer these questions, we employ fixed effects regression analysis over a six-year panel of data (2010–2015).

The importance of research

Besides addressing the empirical research gap, this study contributes to the literature regarding taxes and transfers. Regarding tax, some empirical studies show that local governments tends to have fewer local expenditures if they are financed by their own tax revenue (Cassette & Paty, 2010), congruent with the Leviathan hypothesis in the theory of public finance. Likewise, higher tax autonomy would restrain local spending and could improve the budgets of local governments in the face of an economic downturn (Leberati & Sacchi, 2013; Bartolini et al., 2017). In terms of ramifications of various taxes, local sales taxes are subject to fluctuations in line with the current financial situation, increasing the volatility of own source revenue, while property taxes seem to be less volatile (Afonso, 2017). In OECD countries, the volatility of income and sales taxes has a positive impact on the volatility of local expenditures, whereas a greater reliance on property taxes could lessen the spending volatility (Sacchi & Salloti, 2017). In Japan, studies by Bessho and Ogawa (2015) and Martin-Rodriguez and Ogawa (2017) indicated that own resource revenues (e.g., local taxes) serve a limited role in lessening budget volatility. Accordingly, prior literature shows the role of various taxes (incomes tax, sales tax, property tax) in OECD countries, but similar empirical evidence seems scant in the case of Tokyo local governments, where their local economy – particularly own source revenues – was

adversely influenced by the GFC. Therefore, we expect to contribute to the literature through an examination of the effect of various revenue categories on the volatility of local spending.

Existing studies indicate a strong relationship between grants delivered from the central to local governments and local expenditure, named 'flypaper effects' (money sticks to where it hits) (Bailey & Connolly, 1998). Worthington and Dollery (1999) concluded that, in Australian local governments, no evidence of flypaper effect was found in the case of New South Wales local governments, though Australia had attempted to gain local autonomy. Similarly, evidence was found in EU countries that intergovernmental transfers can make local government larger (the common pool theory) (Cassette & Paty, 2010). Furthermore, grant volatility might result in more local spending volatility in many OECD countries (Sacchi & Salloti, 2017). In Japan, the central government tried to fill the fiscal gap between local governments' own revenues and expenditures through various intergovernmental grants, specifically towards low-income municipalities (Shirai, 2006). The results show evidence of the flypaper effect in the Japanese local government fiscal system (Shirai, 2006). However, the same phenomenon was not present for municipalities in the Tokyo metropolis (Doi, 1996; Miyana & Fukushige, 2001; cited by Shirai (2006)). Accordingly, although empirical evidence on intergovernmental transfers affecting local spending is mixed, we expect to shed light on the impact of grants on spending in the case of Tokyo local governments, particularly special wards.

In conclusion, fiscal federalism, particularly fiscal decentralization, on the grounds of the public finance theory has been widely studied in localism. Empirical studies on decentralized fiscal arrangements associated with tax assignments, intergovernmental transfers, and expenditure responsibilities provide policymakers and practitioners alike with valid public

policies. Nevertheless, a few studies have been undertaken on the same topic in the case of the Tokyo metropolis. We delve into this topic in Chapter 3 of this thesis.

2.3 Public sector financial management

The ongoing economic and financial crisis over the last two decades has urged many governors to innovate public sector financial management. The principal innovations and reforms are relevant to fiscal rules, medium-term budget frameworks, fiscal risk management, accrual-based accounting, and performance budgeting (Schick, 2013). Essentially, these reforms aim to achieve three basic objectives: fiscal sustainability, effective allocation of resources, and efficiency of public service provision (Schick, 2013; Bandy, 2015). In our research project, we mainly focus on public financial management innovations on an accrual basis (public sector accounting reforms) as well as performance budgeting (performance measurement and management) with respect to public effectiveness and efficiency.

2.3.1 Public efficiency

Efficiency is one of the most important notions in the '3E' model (economy, efficiency, and effectiveness), which is commonly characterized as 'Value for Money' in the public sector (Figure 2). Economy means minimizing the costs of inputs. Efficiency relates to maximization of outputs transformed from inputs, while effectiveness refers to intended achievements derived from the outputs of service delivery (Bandy, 2015, p. 274). In the public sector, constituencies frequently have growing demands on public service provisions such as healthcare services, education, infrastructure, welfare, administrative affairs (outputs) corresponding to local resources such as human resources, and capital (inputs). The goal of public managers is to improve the capability to supply more needs within budget constraints (do more with less),

36

correspondingly enhancing their public (technical) efficiency in public service provisions (Da

Cruz & Marques, 2014).
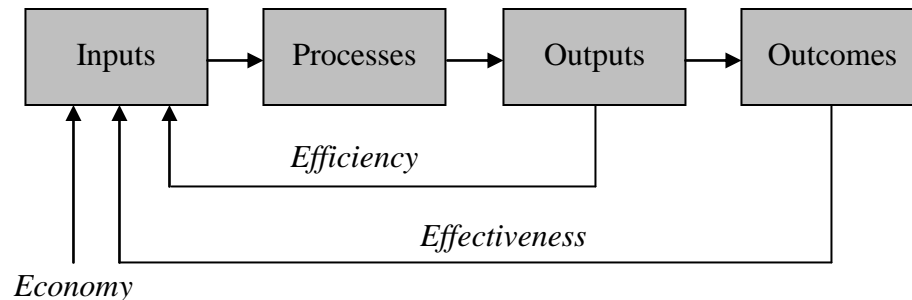
Figure 2. The production performance model

Over past decades, there has been a plethora of empirical studies on local government

performance, varying in different local contexts. Basically, there are two strands of empirical

research at the local level: (1) efficiency of specific (single) local service, and (2) efficiency of a

bundle of (multiple) local services. In the first strand, some studies focus on local road

maintenance (Kalb, 2014), waste management, and recycling services (Worthington & Dollery,

2001). In the second strand, studies evaluate the overall local government efficiency in Australia

(Worthington, 2000; Fogarty & Mugera, 2013), Belgium (De Borger & Kerstens, 1996),

Germany (Kalb, 2010; Kalb et al., 2012; Lampe, Hilgers, & Ihl, 2015), Greece (Dumpos &

Cohen, 2014), Italy (Storto, 2013, 2016), Portugal (da Cruz & Marques, 2014; Afonso &

Fernandes, 2006; Cordero et al., 2017), and Spain (Balaguer-Coll et al., 2013; Balaguer-Coll et

al., 2007; Balaguer-Coll & Prior, 2009). In Japan, there are a number of studies on local

efficiency such as Nijkamp and Suzuki (2009), Fukuyama et al. (2017), Nakazawa (2013, 2014),

Haneda et al. (2012), and Otsuka et al. (2014). It is apparent that these studies on public

performance in Japan are at prefectural rather than municipal level. To address this gap, we aim

to investigate the public efficiency in local governments in the Tokyo metropolis from 2001 to

2015. Specifically, we attempt to observe the efficiency trends of these municipalities before and after the introduction of the accrual accounting system in 2008. This is the focus of Chapter 4.

2.3.2 Public sector accounting reforms

The advent of NPM was primarily motivated by changes in public sector accounting (Carlin, 2005; Guthrie, 1998; Hyndman & Lapsley, 2016; Lapsley, 1999), also known as the 'accountingization' phenomenon (Hood, 1995). The central change seeks to improve public accountability and best practices of PSOs, encouraging the shift from traditional cash-based accounting (adopted in the conventional public sector) to accrual-based accounting (widely used in the business environment) (Diefenbach, 2009; Hood & Dixon, 2013; Hyndman & Lapsley, 2016). Some exemplary pioneers in this reform are New Zealand (Carlin, 2005; Newberry & Pallot, 2004), Australia (Guthrie, 1998), the Netherlands (van der Hoek, 2005), and Japan (Kobayashi, Yamamoto, & Ishikawa, 2016). The extension of these reforms was central to transform cash-based budgeting to accrual-based budgeting in order to match financial reports and fiscal planning in the public sector.

Accruals and cash stand on two extremes in a range of possible accounting and budgeting modes (Blöndal, 2004). As defined, cash-based accounting and budgeting recognizes transactions when cash is received or paid, while accrual-based accounting and budgeting treats transactions when revenues are generated or resources consumed without regard to actual cash inflow or outflow (Blöndal, 2004). Essentially, the difference between the two modes is the time of recording the transactions. Additionally, it is important to distinguish accounting and budgeting. While accounting is considered retrospective financial reporting of transactions and events or *ex-post* portraying of the financial situation and performance achieved, budgeting means planning for future financial scenarios, also known as *ex-ante* financial arrangements. In

the public sector, budgeting and accounting should work in tandem to report, monitor, and control public money. In some advanced economies such as New Zealand, Australia, and Canada, budgeting and financial reports are prepared using the same mode (accrual-based regime) (Warren, 2015), but some countries such as France and Japan still adopt the combined mode, mixing accruals for accounting and cash for budgeting (Kobayashi et al., 2016).

Accrual-based accounting (double-entry bookkeeping mode) prevails over cash-based accounting (single-entry bookkeeping mode) for several reasons. First, the introduction of an accrual accounting regime in the public sector can improve internal and external transparency and accountability (Blöndal, 2004; Carlin, 2005; Guthrie, 1998). This is because accrual information can provide a more accurate view of the cost of public service provisions. Second, the accrual system can facilitate an improved organizational performance, particularly leveraging the capability to allocate resources (Carlin, 2005). Because accrual accounting information is recognized in the medium and long-term rather than the short-term, it can provide better decisions on priority setting of resources used. Resource allocation can become more appropriate and rational under the accrual system. Third, the accrual system provides a comprehensive picture of cost, leading to greater efficiency (Carlin, 2005; Guthrie, 1998). As a result, the adoption of accrual accounting could improve transparency and accountability, enhancing efficiency and effectiveness and, ultimately, raising organizational performance.

Budgeting is a future-oriented activity and process of allocating resources. As defined by Warren (2015), accrual budgeting 'explicitly forecasts and makes decisions about how rights to resources are established and received, and how obligations are incurred and settled', while cash budgeting 'focuses on the forecasting and allocation of one economic resource, that is, cash' (p.

115). In practice, some countries still maintain cash-based budgeting in the public sector, while others have evolved to use accrual-based budgeting in tandem with accrual financial reporting.

Accrual budgeting is highly advocated due to its merits. First, it can synchronize with accrual reporting (Blöndal, 2004; Reichard & Helden, 2017). Some countries still using cash budgeting necessarily articulate year-end accrual financial reports with cash-based budgeting. Although the articulation is seamless, there is still a mismatch. Second, accrual budgeting provides comprehensive information of full costs, resulting in better decision-making (Blöndal, 2004; Reichard & Helden, 2017; Warren, 2015). According to Blöndal (2004), accrual budgeting calculates not only cash-in-hand items but also non-cash consumption (e.g., depreciation of assets, pensions payments). Therefore, Schick (2007, p. 118) regards 'accrual budgeting requires that spending be measured on a cost basis rather than on a cash basis'. Third, accrual budgeting facilitates long-term fiscal sustainability of public finances (Blöndal, 2004). Fourth, accrual budgeting is considered 'a catalyst for other management reforms' in the public sector (Blöndal, 2004). For instance, accrual budgeting adds value to performance budgeting (Martí, 2013). Most importantly, Robinson (2009) argues that accrual budgeting not only improves the effectiveness and efficiency of expenditure but also improves discretions in public assets. Therefore, the nature of accrual budgeting is far 'superior' to cash budgeting in public management reforms.

Public asset management

One of the main putative advantages of the accrual system in efficiency and effectiveness enhancement is better decision-making with respect to public assets (Robinson, 2009; 2016). Decisions relating to the acquisition, disposal, maintenance, and management of assets particularly entail full information, such as depreciation and opportunity cost, to value assets at

40

market price (Robinson, 2009, 2016; Van de Hoek, 2005). However, most local governments not only focus less on the role of balance sheets in decision-making for assets but also remain vague about their valuation of assets (Kaganova & Nayyar-Stone, 2000). Therefore, adoption of private sector accounting practices such as an accrual system can improve public asset management (Kaganova & Nayyar-Stone, 2000; Van der Hoek, 2005).

It is a common belief that PSOs manage and operate their public assets inefficiently and a large amount of assets are underutilized (Kaganova & Nayyar-Stone, 2000; Phelp, 2011). This is because the misuse of public assets is attributed to not imputing a cost to asset utilization. Hence, many advanced countries such as Australia, New Zealand, the UK, and the US have introduced cost efficiency in public asset management since the late 1980s (Kaganova & Telgarsky, 2018). Accordingly, there are some theoretical and empirical studies on public asset management by Bond and Dent (1998), Kaganova and Nayyar-Stone (2000), Tanzi and Prakash (2000), Grubišić, Nušinović, and Roje (2009), Phelp (2010; 2011), and Kaganova and Telgarsky (2018).

However, surprisingly, little is known about asset utilization in the public sector. Moreover, there are few empirical studies on the association of asset utilization and public efficiency in particular. Therefore, we focused the second main research section, in Chapter 5, on asset utilization associated with public efficiency in Japanese local governments, particularly in the Tokyo metropolis. In this study, we expect to contribute to the importance of asset utilization in improving efficiency scores and the role of an accrual-based system in valuing municipal assets for decision-making.

2.4 Performance measurement and management

In relation to accrual budgeting, performance measurement and management also play an important role in supporting budgetary activities (Martí, 2013; Speklé & Verbeeten, 2014; Verbeeten, 2008; Moynihan, 2008; Robinson, 2013). Indeed, goal achievement, resource allocation, and performance measurement are the main functions in managerial accounting relevant to budgeting. In this section, we discuss performance measurement and its use within performance management in the public sector.

2.4.1 Background on performance measurement and management

2.4.1.1 Performance

The definition of performance is derived from multiple perspectives. Simply, by adopting the metaphor of a production model, performance is referred to as the outputs and outcomes of various activities in a production process (Van Dooren, Bourkaert, & Halligan, 2015). Other definitions of performance are reflected through multiple dimensions: productivity, accuracy of work produced, number of innovations, process improvements, reputation for work excellence, attainment of production or service, efficiency of operations, and morale of unit personnel (Speklé & Verbeeten, 2014; Van de Ven & Ferry, 1980; Verbeeten, 2008). Lebas (1995) broadly defined performance as 'the potential for future successful implementation of actions in order to reach the objectives and targets' (p. 23). Following this definition, performance is about the future represented by diversified criteria, such as employment creation, society good, innovativeness in processes and products, customer satisfaction, and growth of market share (Lebas, 1995). In the public sector, De Bruijn (2007) posits that performance can be reflected through effectiveness, efficiency, and legitimacy.

2.4.1.2 Performance measurement

According to management guru Peter Drucker, 'if you cannot measure it, you cannot manage it'. Hence, performance necessarily needs to be measured. The early definition of performance measurement as the process of quantifying efficiency and effectiveness of action was delineated by Neely et al. (1995). From this point of view (in the context of marketing), efficiency is defined as how economically resources are spent to meet the given level of customer need, while effectiveness is the degree to which customer demands are met (Neely et al., 1995). Performance measurement is a metric used to quantify the efficiency and/or effectiveness of an action (Neely et al., 1995). Likewise, Radnor and Barnes (2007) define performance measurement from an operational perspective as 'quantifying, either quantitatively or qualitatively, input, output, or level of activity of an even or process' (p. 393). Hence, performance measurement includes quantitative (financial) and qualitative (non-financial) indicators aligning with the organizational strategy (Franco-Santos, Lucianetti, & Bourne, 2012).

2.4.1.3 Performance measurement system (PMS)

Neely et al. (1995) define a PMS as 'the set of metrics used to quantify both the efficiency and effectiveness of actions' (p. 81). A PMS is composed of an array of individual performance measures and can interact with the external environment in which a PMS operates (Neely et al., 1995). A PMS focuses on developing performance metrics, setting goals, and collecting, analyzing, and interpreting performance information. The objective of this process is to draw performance information on evaluating efficiency and effectiveness of action (Melnyk, Bititci, Platts, Tobias, & Andersen, 2014).

While performance measurement deals with four aspects – what to measure, how to measure, how to interpret data collection, and how to communicate results (Fryer, Antony, &

Ogden, 2009) – a PMS is an instrument that works out the functions of performance. Two exemplar PMSs are BSC and KPI (Franco-Santos et al., 2012; Fryer et al., 2009). A traditional PMS, such as a budgeting system or Activity-Based Cost (ABC) system, cannot be called a PMS because it only copes with financial information without involving non-financial information. Hence, a measurement system is viewed as a PMS when it incorporates the co-existence of financial and non-financial data measured (Franco-Santos et al., 2012). Finally, many researchers summarize the functions of a PMS: linking with the organizational business strategy through financial and non-financial performance indicators and supporting the rationale of decision-making and performance evaluation (Franco-Santos et al., 2012).

2.4.1.4 Performance management

Performance management is a broad concept because it evolved from a performance measurement recommendation, framework, and system (Folan & Browne, 2005). From the view of operations management, Radnor and Barnes (2007) regard performance management as 'action, based on performance measures and reporting, which results in improvements in behavior, motivation and processes and promotes innovation' (p. 393). Similarly, Van Dooren et al. (2015) define performance management as a kind of management using performance information extracted from a PMS for decision-making. This performance information deliberately serves as policy-making, budgeting, and contract management which has three functions: learning, steering and controlling, and account-giving (Van Dooren et al., 2015). Moynihan (2008) defines performance management as 'a system that generates performance information through strategic planning and performance measurement routines and that connects this information to decision venues.' (p. 5). Adoption of doctrinal performance management through the design and use of a PMS in the public sector has created managerial benefits,

improving technical and allocative efficiency and improving bureaucrats' responsiveness to elected officials and the government's accountability to the public (Moynihan, 2008).

It is worth noting a distinction between performance measurement and management. According to Lebas (1995), performance measurement and management are mutually supported and cannot be separated. Performance management, considered a philosophy of management, precedes and follows performance measurement and offers an external management context to the design and implementation of a PMS (Lebas, 1995). Conversely, performance measurement supports performance management (Lebas, 1995). Broadbent and Laughlin (2009) notice that performance measurement consists of ex-post activities while performance management consists of ex-ante actions. In other words, performance measurement concerns the past but performance management is relevant to future affairs (Lebas, 1995).

2.4.1.5 Performance management system

Broadbent and Laughlin (2009) built a conceptual model of a performance management system with regards to the management of results (ends) and their determinants (means) and asserted that 'performance management system is concerned with defining, controlling and managing both the achievement of outcomes or ends as well as the means used to achieve these results at a societal and organizational, rather than individual, level' (p. 283). Hence, the performance management system plays a dual role as ends (ultimate results) and means (instruments to support management reaching to the ends). Ferreira and Otley (2009) extended the performance management system literature by proposing a framework to holistically describe the structure and operations of performance management systems, including vision and mission, key success factors, organization structure, strategies and plans, key performance measures,

target setting, performance evaluation, and reward systems. These factors are externally influenced by the organizational culture and contextual factors.

While a PMS directs development of a metrics set, implementation of the measurement, and translation of performance data, a performance management system gauges the gap between actual and desired outcomes and identifies the rationale for the gap (Melnyk et al., 2014). Accordingly, a PMS is an important but conditional requirement for performance management. A performance management system is a sufficient condition and becomes a complementary tool for a PMS (Melnyk et al., 2014).

All in all, the literature on performance management is still emerging, so the notion of performance, performance measurement, and performance management still calls for additional holistic research. In our study, we focus on PMS use in local governments.

2.4.2 Performance measurement system

2.4.2.1 Benefits and risks of PMS

Implementation and use of PMS aligning with strategies and management control can have strengths and weakness for organizations. De Bruijn (2007) listed a numbers of strengths: (1) A PMS can create transparency and an incentive for innovation, (2) a PMS can reward performance and block bureaucracy, (3) a PMS can promote learning, and (4) a PMS can augment intelligence. Johnsen (2005) presents some benefits in the implementation, use, and outcomes of performance measurement. Regarding implementation, a PMS can monitor efficiency, effectiveness, and equity as well as enhance control. Regarding use, a PMS can reduce asymmetry of performance information, provide an incentive for information users, and enhance learning. For the outcomes step, a PMS can increase accountability, improve resource

allocation, and increase transparency, credibility, and legitimacy (Johnsen, 2005). Essentially, OECD countries draw upon PMSs in public management reform as 'vital catalysts for performance' (Verbeeten & Speklé, 2015). Accordingly, a PMS has the potential to promote performance.

However, when overusing or misusing a PMS, PSOs could face difficulties. According to de Bruijn (2007), a PMS can be an incentive for strategic behavior, hamper innovation, block ambition, deceive actual performance, leave out professional management principles regarding quality, responsibility, and bureaucracy, facilitate copying without learning, and punish performance. Additionally, Cuganesan, Guthrie, and Vranic (2014) maintain that a PMS can be risky when its adverse effects outweigh the benefits. The first risk could be a mismatch between a PMS and its strategies and goals. The second could be the incentive for 'gaming' behavior such as data manipulation and ratcheting. The third could be an obstruction of flexibility and restriction on innovation. The final one could be the potential for distortion or inaccurate reporting of performance (Cuganesan et al., 2014). Van Dooren et al. (2015) describes some dysfunctional behavioral impacts of a PMS – one can distort performance information such as over-representation, under-representation, 'mushrooming' indicators, cream skimming (deliberately select inputs), tunnel vision (synecdoche measurement), polluted indicators, misrepresentation, and misinterpretation. To sum up, the literature on performance management conveys wariness of implementing and using PMSs that could have not only positive impacts but also unintended consequences for PSOs. Hence, public authorities must weigh the beneficial effects and negative impacts of a PMS before its adoption.

2.4.2.2 The use of PMS

PSOs, in general, utilize PMSs for multiple operational and strategic purposes. A number of conceptual frameworks have been raised in scholarly literature relating to the role of PMSs in management control systems (MCS), such as Levers of Control (Simons, 1995), and models of rationality (Broadbent & Laughlin, 2009). Moreover, empirical studies also suggest some findings of PMS used in practice, including works by Hansen and Van der Stede (2004), Henri (2006), Franco-Santos et al. (2007), and Speklé and Verbeeten (2014).

The Levers of Control proposed by Simons (1995) conceptually describes four systems as four levers of control linked to an organization's business strategy. The first lever is the diagnostic control system, which efficiently and effectively pursues organizational goals and objectives. By means of diagnostic functions, managers can monitor the progress of employees and departments' achievements. The second lever is the belief system, in which top managers direct the organizational values that individuals must follow; less strictly, top managers can encourage individuals to create organizational values. The third lever is the boundary system, acting as the 'organization's brakes' to adjust organizational performance. The final lever is an interactive control system, supporting subordinates' decisions and learning in important strategic matters. While the diagnostic control system strictly monitors organizational goals and objectives, the interactive control system relieves constraints and promote strategic development.

The model of rationality presented by Broadbent and Laughlin (2009) indicates the means and ends of a PMS by using the 'middle range' theory. The nature of PMS stands on two extreme continua, from 'transactional' to 'relational' use, responsive to the characteristics of instrumental and communicative rationalities. The transactional uses of PMS 'have a high level of specificity about the ends to be achieved and often a clear specification of the means needed

to achieve these defined ends'. The relational use of PMS suggests that 'the ends and means are deliberately subject to a discourse between the stakeholders and chosen by them' (Broadbent & Laughlin, 2009, p. 289). Accordingly, the use of PMS ranges from less discretionary (transactional use or instrumental rationality) to more discretionary (relational use or communicative rationality).

Empirical studies additionally address the use of PMS in practice. For instance, Hansen and Van der Stede (2004) explore the four reasons for budgeting associated with organizational performance, including (1) operational planning, (2) performance evaluation at the operational level (short-term), (3) communication of goals, and (4) strategy formation at the strategic level (long-term), which are positively associated with the reason-to-budget performance. Apart from prior research, Hansen and Van der Stede (2004) provide a holistic view on the multiple uses of budgeting with an emphasis on performance evaluation. The use of budgeting-based performance evaluation relates to target-setting and incentives, because enabling stakeholders to participate in setting targets motivates budgeting information users to achieve their targets for the rewards (Hansen & Van der Stede, 2004). Moreover, formulating a clear strategy and communicating goals to employees also contributes to budgeting performance (Hansen & Van der Stede, 2004). Indeed, the study by Hansen & Van der Stede (2004) makes a distinction between operational and strategic use in budgeting.

Henri's study (2006) investigates the relationship between organizational culture and two attributes of PMS (the nature of use and diversity of measurement). There are four classifications of the use of performance measures: First, monitoring function as the feedback of performance information to top managers, enabling fine-tuning of performance. Second, strategic decision-making as a facilitator for the justification of decisions and actions to ensure legitimacy of the

PMS. Third, performance information that attracts attention focusing (equivalent to interactive control to urge communication insight of the organization in the Levers of Control theory by Simon (1995)). Therefore, monitoring use is classified as operational and incentive, while attention focusing and strategic decision-making are considered exploratory.

The findings of contemporary PMS by Franco-Santos et al. (2012) additionally exhibit the incentive and exploratory use of PMS. Speklé and Verbeeten (2014) summarized five groups of PMS use based on Franco-Santos et al.'s research: measuring performance (monitoring, measuring, and evaluating performance); strategic management (strategic formation, implementation, execution, and alignment); communication and compliance with regulations; influencing behavior (rewarding); learning and improvement (double-loop learning for performance improvement).

Finally, Spekle and Verbeteen (2014) categorize PMSs by incentive-oriented, operational, and exploratory use. The incentive use motivates individuals to attain the targets outlined in the rewarding mechanism as in the function of Hansen and Van der Stede's performance evaluation (2004). In contrast, exploratory use concerns priority setting and policy development, along with double-loop learning and improvement (Speklé & Verbeeten, 2014). The operational use (such as operational planning, process monitoring) is given less attention in this research because it is widespread, commonplace and always present in any organization (Speklé & Verbeeten, 2014; Hansen & Van der Stede, 2004). In line with this research, we focus on incentive-oriented and exploratory use of PMSs in PSOs.

All in all, while incentive-oriented use of PMS can be implemented for operational functions, exploratory use can be strategically deployed in the medium/long-term. These two

conceptual variables are the foci of the empirical study in Chapter 6 for the TMG and Tokyo

local governments, which have been given less attention so far. This study expects to explain that

the incentive-oriented use rather than exploratory use appears in the context of local

governments still adopting the traditional cash-based budgeting regime.

2.4.3 Theory supporting performance measurement in the public sector

In principle, aligning the PMS with the business strategy can improve organizational

performance. It is necessary to discuss underlying theories supporting this argument; because the

discourses on performance management literature are emerging and ongoing, viewing PMS from

the perspective of a specific theory seems biased and incomprehensive. Hence, a combination of

various theories is necessary in relation to PMS. In Chapter 6, we rely on three theories in

particular: principal-agency theory, institutional theory, and contingency theory.

2.4.3.1 Principal-agency theory

Principal-agency theory indicates the relationship relying on the agreement between

individuals (principals) and their representatives (agents), in which principals delegate rights and

responsibilities to agents (Baiman, 1990). It is assumed that individuals are fully rational and

have well-defined preferences, conforming to the axioms of expected utility theory (Bonner &

Sprinkle, 2002). Furthermore, Baiman (1990) posits that individuals tend to be motivated by

their self-interest to gain their utility. For instance, individuals can contribute to performance

because they expect to receive payments, bonuses, or career promotions. Otherwise, they ignore

work that does not improve their economic well-being.

There is an asymmetry of information between principals and agents (Stede, Wim, Chow,

& Lin, 2006). While agents such as PSOs are frequently close to people and communities and,

thus, have a full understanding of their needs, principals have ambiguous information. Therefore, using a PMS to elicit information helps reduce the information asymmetry among principals and agents (Franco-Santos et al., 2012; Stede et al., 2006). Furthermore, aligning a PMS with strategic planning and target-setting can motivate individuals to achieve their goals and objectives (Franco-Santos et al., 2012; Verbeeten, 2008). Therefore, agency theory demonstrates the underlying foundation to explain the extrinsic motivation of individuals who attempt to pursue their individual and organizational performance.

2.4.3.2 Institutional theory

Institutional theory suggests that the design and implementation of performance measurement and management might be influenced by contextual institutions (Brignall & Modell, 2000; Cavalluzzo & Ittner, 2004; Modell, 2009). The basic premise of institutional theory maintains that PSOs influenced by external and internal constituencies should comply with external rules or laws to gain legitimacy, subsequently attaining beneficial resources and long-term success (Brignall & Modell, 2000). Hence, maintaining legitimacy-seeking behavior along with some rationalized institutional practices can lead to survival and success.

Legitimacy is the central point of organizational institutionalism. A PMS is designed and used for seeking legitimacy such that PSOs can fit with the social and cultural environment. According to Modell (2009), from the institutional perspective, a PMS deals with three substantive aspects. First, goal-directed performance measurement and management, using management techniques, is embedded in performance measurement research. Second, reliance on the one-sided dimension of financial performance measures is insufficient; rather, incorporation of multiple performance measures, including financial and non-financial performance measures such as BSC, is necessary. Finally, performance evaluation producing

outcomes such as rankings, league tables, and benchmarking is the third focus of performance measures associated with the institutional theory. Briefly, three main points of goal-directed, multi-dimensional performance measurement and evaluation are our foci in this study.

2.4.3.3 Contingency theory

Contingency theory proposes that the nexus between a PMS and organizational performance is subject to environmental changes (Ferreira & Otley, 2009; Franco-Santos et al., 2012). The fundamental premise of contingency is that the design and implementation of a PMS is not universal for any PSOs around the world, but should be adjusted in accordance with the local context of institutions and administration. In fact, Chenhall (2007) summarized the contextual variables related to MCS: the external environment, technology, organizational structure, size, strategy, and culture.

Ferreira and Otley (2009) proposed the PMS framework and argued that 'external environment, strategy, culture, organizational structure, size, technology, and ownership structure have an impact on control systems design and use' (p. 267). Indeed, in the inter-relationship among environmental factors, business strategy, and PMS, any change of environment leads to the need for a revised PMS (Melnyk et al., 2014). However, in reality, business environment changes merely cause the modification of a strategy rather than performance measures. Hence, PMS implementation could lag behind the strategy; therefore, innovation of PMS through learning and improvement is necessarily emphasized (Melnyk et al., 2014). Some scholars suggest that contingency-based research would be more useful if it engaged with economic theories (e.g., agency theory and psychological theory) (Chenhall, 2007).

Overall, research on PMS is ongoing, most of it in empirical studies reliant on various theories such as agency theory, contingency theory, and institutional theory. The agency theory explains the incentive and motivation of individuals in using PMS from an economic perspective. The institutional theory describes the external influence on the use of PMS from a legacy perspective. The contingency theory is exposed to learning and improvement under the flicking changed contextual factors and business strategy needs and requires adaption to the varying environmental context.

2.5 Summary of underlying theories

This study emphasizes four inter-disciplinary research topics: (1) public finance, (2) public efficiency, (3) public sector accounting, particularly public budgeting, and (4) PMS use.

In the first research topic (Chapter 3), prevailing characteristics of NPM are described as an assignment of financial responsibility to lower tiers of governments. Following this NPM principle, the public finance corpus of literature presents fiscal federalism in the first and second generation. The underlying theories are the Tiebout model (1956), Olson's fiscal equalization (1969), Oates' decentralization theorem (1972), and Brennan and Buchanan's Leviathan constraints (1980). Accordingly, our study on public finance takes on fiscal decentralization, which draws on theories of fiscal federalism. To explain further, our study focuses on the devolution of fiscal power associated with four inter-jurisdictional fiscal issues: taxation assignment, spending responsibilities, intergovernmental grants, and sub-national borrowing (Vo, 2010) in the Tokyo local governments and separately between urban and suburban areas.

In the second research topic (Chapter 4), we examine the municipal efficiency for entire local governments in Tokyo so as to detect the efficiency status of each local unit and compare

municipal efficiency among urban and suburban areas before and after the introduction of an accrual accounting system in 2008.

In the third research topic (Chapter 5), we extend our study on various determinants of efficiency scores relevant to public sector accounting and budgeting. As discussed in prior literature, financial reporting and budgeting in the public sector seem to lag behind that in the private sector. In nature, the private sector's management styles are likely to be more efficient and effective than that of the public sector. Hence, our study raises some prominent advantages of an accrual-based system, which is widely used in the private sector, that PSOs could follow. Consequently, accrual accounting and budgeting are a foundation of the public sector accounting and budgeting study in Chapter 5.

In the fourth research topic (Chapter 6), associated with performance budgeting, performance measurement is necessary to enhance the budgeting procedure. The emerging theory regarding the use of PMS (incentive and exploratory use) is introduced. Moreover, theories regarding performance measurement that are still under-researched, such as agency theory, institutional theory, and contingency theory, form a basic framework in Chapter 6.

Overall, the encapsulated body of literature in relation to fiscal decentralization, public efficiency and public sector accounting, and performance measurement becomes the underlying theory for our research in following chapters.

CHAPTER 3. THE ROLE OF REVENUE VOLATILITY ON LOCAL EXPENDITURE VOLATILITY

Since the 2008 GFC, most municipalities have been confronted with financial challenges and budget constraints; Tokyo's local governments are no exception. This study examines budgetary volatility in terms of local revenues and expenditures across the jurisdictional region in Tokyo and emphasize the public finance difference between the urban and suburban areas. Consequently, some prescribed public policies are to smooth out the volatility of local spending.

3.1 Introduction

There has been much recent attention on decentralization of government in most advanced economies in response to a widespread belief that fiscal decentralization reforms – assigning autonomy from the central government to local governments in terms of raising revenues and controlling expenditures – can offer potential gains in economic growth (Martinez-Vazquez & McNab, 2003), social welfare improvement (McLure & Martinez-Vazquez, 2000), and efficient provision of public services (Oates, 2008; Martinez-Vazquez & McNab, 2003). Fiscal decentralization principally relates to matters of revenue raising and taxation powers, decisions on spending, and intergovernmental relations (Vo, 2010). The merit of fiscal decentralization has been shown in recent empirical studies – In particular, it has been asserted that "more tax autonomy would improve the budget of all tiers of government" (p. 1) and expenditure decentralization could improve responsiveness in the face of economic shocks (Bartolini, Sacchi, Salotti, & Santolini, 2017). Moreover, tax decentralization can help sub-national governments avoid deficits (as shown for several European countries) (Foremy, 2014), while expenditure decentralization can help reduce regional income disparities (Sacchi & Salotti,

2016). However, because devolution of autonomy falls most heavily upon local governments, there may be unanticipated consequences for this tier of government (Tanzi, 2003) and it is, therefore, important to empirically investigate expenditure volatility for a municipal system.

In some countries, the devolution of spending responsibilities is not matched by the corresponding devolution of tax revenues, which can exacerbate vertical fiscal imbalance (Eyraud & Lusinyan, 2011). According to McLure and Martinez-Vazquez (2000), it is important to clearly determine expenditure portfolios in advance, before designing revenue assignments and transfers, to ensure that municipalities have adequate revenues to match expenditure responsibilities. There is evidence to suggest that residents are increasingly demanding more public services despite low willingness to pay, thus expanding local expenditures (Eyraud & Lusinyan, 2011; Grant & Drew, 2017, p. 274). If this is indeed the case, the other sources of revenue (and other cash flows) may hence be required to match expenditures. Thus, there is a good reason for supposing that the variance of taxes and transfers might be related to local spending, but few scholarly studies so far have explored the potential associations.

In general, measures of volatility can provide a picture of an entity's economic progress, since this approach allows for an understanding of the fluctuations from equilibrium (Staley, 2017). Expenditure volatility is defined as the standard deviation of the annual growth rate of local spending for a given fiscal year (Staley, 2015; Sacchi & Salotti, 2017). Research on expenditure volatility at the local level is particularly salient, given the need for local authorities and policymakers to arrive at efficacious public policy prescriptions. For instance, Sacchi and Salotti (2017) argue that local spending volatility probably impedes the health of local economies. If expenditure tends to be volatile and unpredictable, it is hard for local businesses to plan their spending (with respect to staffing needs, inventory, and the like). This is especially

true if municipal staff numbers are volatile – which casualized staff probably feel most keenly –

because less money is being injected into the local economy. Moreover, expenditure volatility,

even revenue volatility, might give rise to uncertainty concerning future fiscal periods, which

will hamper the "selection of efficient production processes" (Crain, 2003, p. 96). Thus, local

expenditure volatility is worthy of study given its implications for local economic performance.

The nascent scholarly literature on expenditure volatility has identified some of its

determinants. Potential determinants of expenditure volatility can be attributed to variation in

taxes and intergovernmental grants (Sacchi & Salotti, 2017), degree of fiscal decentralization in

most developed and developing countries (Furceri et al., 2016), quality of fiscal institutions

(Albuquerque, 2011), and revenues and debt outstanding (Dension & Guo, 2015). However,

these academic works have been largely focused on Europe and America and there is a gap in the

scholarly literature relating to the local expenditure volatility in Japan, particularly local

governments. Our paper is motivated by a desire to address this gap, especially concerning

business cycle fluctuations in the Tokyo municipalities sequent to the 2008 GFC.

This study examines: (i) the association between the volatility of various revenues (e.g.,

local taxes and grants) and local spending volatility, and (ii) how the fiscal differences between

special wards and cities affect the spending volatility in the context following the 2008 GFC. To

do so, we employ fixed effects regression analysis over a six-year panel of data (2010–2015).

The remainder of the chapter is organized as follows: the next section introduces the

related literature on Tokyo public finance. Thereafter, we outline our empirical strategy.

Following this, we discuss the statistical results and findings. We conclude with a discussion on

the public policy implications arising from this study.

3.2 Tokyo Local Government Remit, Revenue, and Expenditure

Local government in Japan is organized in two tiers: prefectures and municipalities. At the prefectural level, the TMG administratively governs Tokyo metropolis. At the municipal level, there are 23 special wards, 26 cities, five towns, and eight villages[2]. The Tokyo metropolis covers an area of 2,191 square km and has an estimated population of 13.491 million as of October 2015 (TMG, 2018). While 9.241 million reside in the 23 wards with a population density of 14,746 persons per square km, 4.233 million live in 26 cities with 3,640 persons per square km (TMG, 2018).

In Japan, the national and local government's assignment of responsibilities was clarified by the Omnibus Local Autonomy Law in 1999. The abolition of the system of delegated functions — a representative agency of municipalities appointed by the national government — led to the elimination of intervention by the national government. Hence, local governments now have broad responsibilities for their administrative functions autonomously and comprehensively (CLAIR, 2013; Mochida, 2008). Prefectures are responsible for prefectural roads, high schools, public health centers, and police, whereas municipalities are responsible for urban planning, municipal streets, schooling for children under 15, health care services, social welfare, garbage disposal, and fire services. In contrast to the 26 cities' responsibilities for public services, several public provisions such as water supply, sewerage, and fire protection in the 23 wards are co-shared and undertaken by TMG (TMG, 2018). Thus, an incorporated council between TMG and special wards has been established to facilitate continuous negotiations and discussions regarding

---

[2] The five towns and eight villages — rural and island areas in Tokyo — are designated in the geographical area, but their scope of public services and finances are comparatively small as compared with those provided by the special wards and cities. Therefore, we eliminated these units from our analyses.

co-sharing activities and other fiscal interrelation (TMG, 2018). Therefore, the relationship between TMG and special wards can be considered unique and mutually interdependent.

The primary local revenue sources for Japanese local governments are local taxes, intergovernmental grants, and local bonds. Local taxes are imposed by both prefectural and municipal governments with various types of taxes: inhabitant, business, consumption, property tax, etc. In general, tax bases and rates are regulated consistently across local governments. Intergovernmental grants are composed of Central Government Subsidies (CGS) and Local Allocation Tax (LAT). The CGS is an obligatory share from the national government for specific purposes (e.g., educational assistance, post-natural disaster alleviation) (MIC, 2017). The LAT, allocated as a fixed portion of national taxes such as income, corporate, alcohol, and consumption, plays an essential role in horizontal fiscal equivalence to narrow the gap between poor and affluent regions. It is estimated by the gap between basic fiscal needs and revenues. Tokyo is such a wealthy region (the income per capita exceeds the national fiscal standard) that LAT grants are basically not allocated to municipals in Tokyo from the national government (MIC, 2017); however, they are still recipients of LAT grants from TMG. Finally, the Japanese principle of local autonomy allows municipalities the authority to issue bonds (Tanaka, 2011). Local bonds perform a deficit adjusting function, thus shifting the debt burden to the next generation and augmenting general revenue sources (Tanaka, 2011).

Importantly, this study presents two main differences in the revenue-raising powers of the 23 wards and 26 cities. Regarding local taxes, there are 16 types of charges for the 23 wards, but 22 for the cities (TMG, 2018). Moreover, the rate of corporation tax in the special wards is around four times higher than the rate imposed in the cities (Bureau of Taxation, 2017). The TMG collects three kinds of charges (property tax, corporation tax, and special land acquisition

and holding tax) on behalf of the special wards, while each city collects these taxes by themselves (Ohsugi, 2011). The TMG divides the pool of these collected taxes for two functions: 55% for grant allocations towards special wards and 45% for co-sharing administrative works. The grant funds are further divided into special grants (5%) and ordinary grants (95%). Special wards receive financial coordinating grants or LAT grants from the TMG, while cities do not. Therefore, special wards differ from cities regarding revenue sources.

The volatility of revenue and expenditure has increasingly been attracting the attention of scholars. Regarding revenue, the constant fluctuation of income can make it hard for local authorities and policymakers alike to plan, budget, and provide efficient and sustainable goods and services (Staley, 2017). Thus, it is considered essential to identify factors affecting revenue volatility. For instance, tax limitations are positively linked with state revenue volatility in the US states (Staley, 2017; 2015). Moreover, Afonso's empirical study (2017) asserts that greater reliance on local sales tax can increase the volatility of own source revenue, whereas more significant reliance on property tax can decrease the volatility of individual source revenue, consistent with the underlying theory of fiscal federalism (Oates, 2011). Additionally, increasing property tax can reduce capital expenditures, which are not affected by the local sales tax (Afonso, 2017).

Regarding expenditure, some critical factors have been identified. For example, as pointed out by Furceri (2007) and Afonso and Furceri (2010), the effect of government expenditure volatility on economic growth is generally negative and statistically significant for European countries. However, some advanced countries are able to absorb expenditure volatility because of their better taxation system and powerful domestic stabilizers (Furceri, 2007). Similarly, Furceri and Poplawski-Ribeiro (2008) conclude that country size is negatively

associated with government spending volatility, which is also consistent with Albuquerque's findings (2011).

However, there is only nascent literature regarding factors contributing to expenditure volatility at the local level. Sacchi and Salotti (2017) have investigated the influence of local taxes (income tax, property tax, and sales tax) and grants on local spending volatility in 20 OECD countries. Their results suggest that while volatilities of grants and income tax are positively associated with local spending volatility, the volatility of property tax has a negative effect on spending. Another study, by Denison and Gou (2015), elaborated on the association between outstanding debt and expenditure volatility for local US jurisdictions and found that debt had a statistically significant impact on the expenditure volatility for 13 states.

The above literature suggests several factors that might be associated with revenue and expenditure volatility. However, it also suggests that the relationship between revenue and expenditure volatility is not uniform. Furthermore, only a few studies present a clear link between revenue and expenditure volatility, even though the importance of this issue has been stressed by Thompson and Gates (2007, p. 825), who state that "volatile, unpredictable revenue growth causes all sorts of unpleasant governmental responses, most commonly manic-depressive patterns of spending and taxing."

In addition to studies from abroad, some related empirical studies in the Japanese context, such as Bessho and Ogawa (2015) and Martin-Rodiguez and Ogawa (2017), investigated local fiscal adjustments and found that own-source revenue plays a limited role in balancing provincial budgets because it tends to be offset by the supply of grants. Subsequently, it has been suggested that municipalities can induce grants by expanding current expenditure. These scholarly works

provide some insight into the fiscal deficit associated with own-source revenues, grants, and debt; however, the specific association between revenue and expenditure still requires investigation. Moreover, this association has not yet been examined in detail at the specific local level of Tokyo where the significant amount of tax was decreased because of the GFC since 2008 (Figure 3). To address these gaps in the literature, we analyze the association between the volatilities of various revenue components (e.g., local tax, grants) and expenditure volatility of local governments in Tokyo. The next section will outline the empirical strategy we employ.
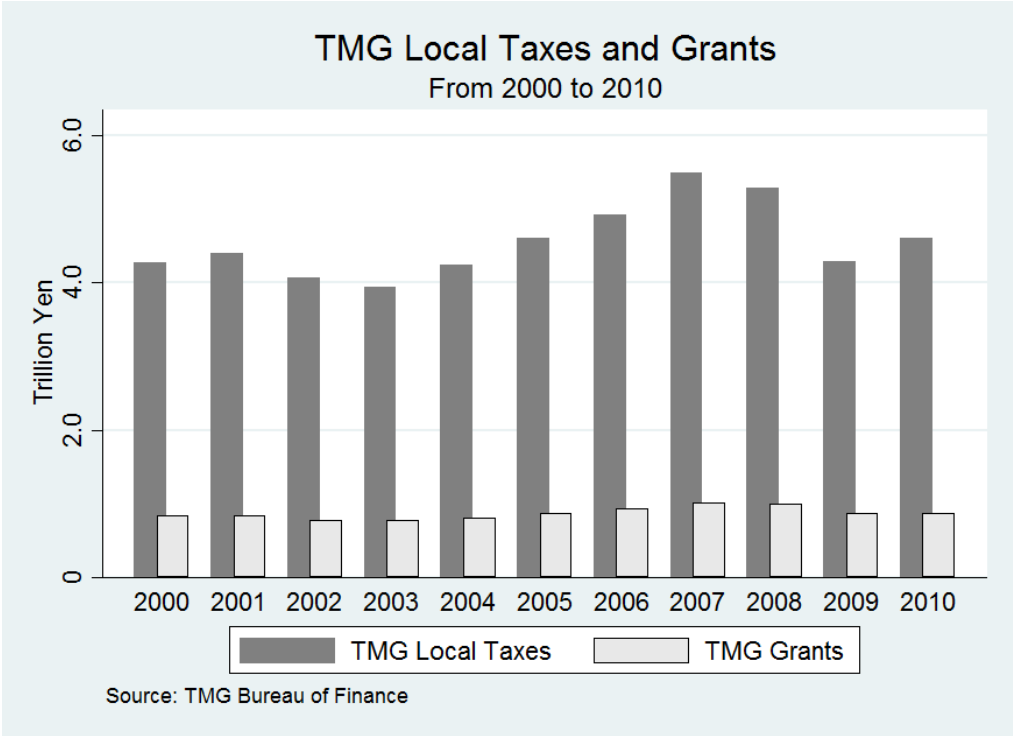
Figure 3. The TMG Local Taxes and Grants for Special Wards from 2000 to 2010

3.3 Empirical Strategy

Local expenditure volatility is defined as the standard deviation of the annual growth rate of local spending for a given fiscal year (Sacchi & Salotti, 2017). The measurement of the volatility of revenue components is estimated in a similar manner. In contrast to the Sacchi and

Salotti (2017) model, which examined the expenditure volatility of local governments in the context of OECD nation states, our specification instead concentrated on local governments in the context of the TMG, Japan. Hence, several revenue components from the Sacchi and Salotti (2017) model were modified to reflect the circumstances of the Tokyo local governments. Data were extracted from the cash-based accounting system of 49 administrative units in Tokyo from 2010 to 2015. Fixed Effects (FE) panel regression for data analysis was adopted in this empirical study because it controls for time-invariant latent variables that might influence the dependent variable. Although Random-Effect (RE) models might be more efficient, they needed to overcome the problem of a composite error possibly correlated with the explanatory variables (Drew & Dollery, 2016). According to Thompson and Gates (2007), there is a possibility that a feedback loop might operate between revenue and expenditure and readers should remain cognizant of this potential interaction when interpreting our results, which are primarily directed at investigating the association that runs from revenue to expenditure. The regression specification was estimated by the following equation:

$$Exp_{i,t} = \alpha_{i,0} + \beta_1 LT_{i,t} + \beta_2 CT_{i,t} + \beta_3 OT_{i,t} + \beta_4 SG_{i,t} + \beta_5 G_{i,t} + \beta_6 S_{i,t} + \beta_7 FC_{i,t} + \beta_8 OR_{i,t} + \beta_9 B_{i,t} + \mu_{i,t}$$

(1)

Where independent variable $Exp_{i,t}$ is the annual volatility of municipality i expenditure at the period of t fiscal year. In the parsimonious model (1), the dependent variables − $LT_{i,t}$, $CT_{i,t}$, $OT_{i,t}$, $SG_{i,t}$, $G_{i,t}$, $S_{i,t}$, $FC_{i,t}$, $OR_{i,t}$, and $B_{i,t}$ − denote the volatility of local tax, consumption tax[3] (ln), other taxes (ln), special grants, grants (ln), subsidies, fees and charges (ln), other revenues, and

---

[3] In Japan, 6.3% and 1.7% of the consumption tax pool go to national and local governments, respectively. These tax rates are preset before tax collection, which is different from some OECD countries, such as Australia, where the national government collects the consumption tax first, then transfers funds to state governments on the basis of horizontal fiscal equalization. In this study, we only accounted for the 1.7% consumption tax directed to local government.

local bonds (ln), respectively (transformations were applied to correct skewed distributions

where indicated). The set of β_k (k =1,...,9) represents the estimated parameters in the regression

model and $\mu_{i,t}$ is an independent identically distributed random error term. To ensure the

robustness of the results, we also controlled for municipal specific features under an alternative

specification. Therefore, specification (2) with additional control variables has been proposed:

$$Exp_{i,t} = \alpha_{i,0} + \beta_1 LT_{i,t} + \beta_2 CT_{i,t} + \beta_3 OT_{i,t} + \beta_4 SG_{i,t} + \beta_5 G_{i,t} + \beta_6 S_{i,t} + \beta_7 FC_{i,t} + \beta_8 OR_{i,t} + \beta_9 B_{i,t}$$
$$+ \sum_{i=1}^{s} Control_{i,t} + \mu_{i,t} \qquad (2)$$

In this study, we employ Full-Time Employees (FTE) as a proxy for municipal size

(Drew et al., 2016). This control variable is estimated by the natural logarithm of the number of

FTE staff at each municipality i at year t. Some relevant empirical studies use different sets of

control variables, for instance: population to measure country size, urbanization as a percentage

of each municipal population over the total population, population density (Sacchi & Salloti,

2017), population older than 65, population growth rate, or unemployment rate[4] (Staley, 2017).

The population-related control variables seem inappropriate for the Tokyo case, mainly because

Tokyo metropolis has a unique population distribution; it is located at the center of the Greater

Tokyo Area, surrounded by three neighbors − Saitama, Chiba, and Kanagawa, the most populous

areas in Japan. People in these neighborhoods commute daily to their offices in the center of

Tokyo. Accordingly, Tokyo's population gap between daytime and nighttime was around 2.89

million in 2010 (TMG, 2018). These commuters essentially export metropolitan corporation tax

forward to Tokyo, but rarely benefit from public services to the same degree as Tokyo's

residents. It is hence argued that control variables related to population parameters would fail to

---

[4] The data of unemployment rate for the municipal level are basically published in five-year intervals. Accordingly, the availability of data did not fit with our panel.

accurately reflect either the scale or substance of the expenditures made by the Tokyo local governments. Therefore, FTE staff is considered a more accurate proxy for municipal scale in the Tokyo metropolis.

With respect to specification (1) and (2), we stratify each of the specifications into two groups (23 wards and 26 cities) and observe how the different fiscal arrangements across the two groups affect the volatility of local spending (please see Table 1 for disaggregated summary statistics). The summary statistics suggest that local spending volatility in the special wards is relatively higher than for the cities. Regarding local revenue compositions, while local tax for the 23 wards is more volatile than for 26 cities, income from consumption tax and other taxes in both areas is relatively stable. This is because the core of the local tax is corporation tax, to which most of Japan's large corporations located in the special wards or Central Business Districts contribute. The corporation tax fluctuates according to corporation income, which is vulnerable to the business cycle. Hence, the local tax in the special wards seems less stable. For grants, the volatility in the special wards is lower than that of the cities. Moreover, the volatility of subsidies supplied to the special wards are somewhat higher than that of the cities, but the volatility of fees and charges in the special wards are quite stable relative to the cities. The volatilities of the two remaining revenues – other revenues and local bonds – are considerably higher in the special wards. In sum, the magnitude of volatility is quite disparate between the two groups when it comes to local tax, grants, other revenues, and local bonds.

Table 1. Summary statistics of 23 special wards and 26 cities.

Variables	23 special wards (n =92)				26 cities (n = 104)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Dependent variable								
Expenditure	3.973	5.098	0.0174	31.77	3.539	2.256	0.119	10.741
Independent variables								
Local tax	1.610	1.277	0.250	7.768	1.176	0.837	0.0008	3.888
Consumption tax	8.856	10.640	0.150	32.525	11.080	12.461	0.111	33.106
Other taxes	14.959	7.835	1.477	27.882	12.426	6.256	0.166	26.117
Special grants	40.672	8.934	24.055	68.044	28.95	9.899	8.949	55.659
Grants	6.562	13.012	0.254	81.4182	70.736	152.491	0.838	783.317
Subsidies	6.145	5.468	0.473	28.166	5.940	5.014	0.262	21.613
Fees and charges	2.389	1.573	0.158	9.057	3.679	4.231	0.144	25.524
Other revenues	25.36	27.082	2.595	150.615	8.269	4.410	0.079	19.927
Local bonds	330.31	1366.55	0.000	8237.91	35.655	28.624	0.158	139.014
Control variables								
FTE	2478.576	1011	901	4856	787.644	524.437	310	2656

3.4 Research Results and Findings

3.4.1 Research results

In Table 2 Model 1, which refers to specification (1) for the entire Tokyo municipal cohort, we can see that local expenditure volatility was statistically significant and positively associated with the volatilities of other taxes, other revenues, and local bonds, but negatively associated with that of grants. However, as we have outlined, the two systems of Tokyo local governments are very different in both their revenue streams and remits; thus, it is necessary to stratify the regression into special wards (Model 2) and cities (Model 3), respectively.

On doing so, we find that, in Model 2, the statistically significant and positive explanatory variables for the special wards are volatilities of local tax (β=0.193, p<0.10), subsidies (β=0.173, p<0.05), other revenues (β=0.022, p<0.001), and local bonds (β=0.082, p<0.05). Specifically, the volatility of public spending is associated with an increase of 19.3% for an additional one-standard deviation in volatility of local taxes, reflecting the critical role of special ward local tax, rather than consumption tax and other taxes, in financing expenditures. Then, a change of one-standard deviation in the volatility of subsidies is associated with an increase in expenditure volatility of around 17.3%, *ceteris paribus*. For the other revenues, the local expenditure volatility is expected to increase by just 2.2% in response to a one-standard deviation increase in other revenues.

A 1% increase in the volatility of local bonds is associated with a mere 0.08% increase in volatility of local expenditure and implies that local bonds have little effect on public spending. It is also important to note that the volatility of grants has a significantly

negative association with the volatility of local spending (β=-0.704, p<0.001). Thus, a 1% increase in the volatility of grants is associated with a relatively strong response of 0.7% (decrease) in the volatility of local spending, *ceteris paribus*. This result does not support the common pool theory, being inconsistent with Sacchi and Salotti's findings (2017). In general, it is expected that municipalities will increase their expenditures in response to offerings from higher-tier governments. However, this behavior is not recognized in the case of special wards. In this sense, they differ from other Japanese municipalities to induce grants by expanding expenditure (Bessho & Ogawa, 2015; Martin-Rodiguez & Ogawa, 2017). To provide further illumination, we present additional analysis in the subsection.

In Model 3, we found evidence of statistically significant positive associations for consumption tax (β=0.093, p<0.05), other taxes (β=0.267, p<0.001), and other revenues (β=0.062, p<0.01) concerning the local expenditure volatility in the cities. Specifically, our results suggest that increases in the consumption tax volatility of 1% are associated with local spending increases of just 0.09%. For other taxes, a 1% increase in volatility is expected to lead to a relatively strong response of a 0.26% increase in the volatility of local spending. We can see that, compared to the consumption tax, fluctuations to other taxes (e.g., golf course tax, vehicle tax) tend to elicit a stronger expenditure response by local governments. By virtue of cities being located in suburban areas, where land is available for development of recreational facilities (e.g., golf courses) and transportation infrastructure, expenditures for these public works are significantly higher and financed by other taxes. The local spending volatility is associated with an increase of 6.2% for every additional one-standard deviation in the volatility of other revenues.

Table 2. Results of the parsimonious model with FE

Total expenditure *(ln)*	*Model 1*	*Model 2* (Special wards)	*Model 3* (Tama cities)
Local tax	0.0990	0.1930$^+$	-0.0659
	(0.0698)	(0.1021)	(0.0892)
Consumption tax *(ln)*	0.0667$^+$	0.1512	0.0935*
	(0.0395)	(0.0949)	(0.0385)
Other taxes *(ln)*	0.2182**	0.2452	0.2670***
	(0.0766)	(0.1660)	(0.0697)
Special grants	-0.0216	-0.0123	0.0106
	(0.0167)	(0.0360)	(0.0186)
Grants *(ln)*	-0.4081**	-0.7039***	-0.0709
	(0.1301)	(0.1943)	(0.1607)
Subsidies	0.0206	0.1729*	-0.0004
	(0.0216)	(0.0791)	(0.0187)
Fees and charges *(ln)*	0.0140	0.0607	-0.1276
	(0.0899)	(0.1703)	(0.0935)
Other revenues	0.0216***	0.0221***	0.0620**
	(0.0033)	(0.0041)	(0.0202)
Local bonds *(ln)*	0.0729**	0.0821*	0.1401
	(0.0273)	(0.0335)	(0.0903)
Observations	196	92	104
Coefficient of Determination	0.6116	0.7285	0.5736
n	49	23	26

Note: Significant levels are *** p< 0.001, ** p<0.01, * p<0.05, + p<0.1. The denote (ln) represents the natural logarithm. Numbers in parentheses are standard errors.

Our results confirm that there are clear differences between Tokyo's two local systems. In terms of positive and statistically significant associations, special wards are associated with local taxes, subsidies, other revenues, and local bonds while cities are associated with a consumption tax, other taxes, and other revenues. The positive relationship between the volatility of various kinds of taxes and that of local spending seems to fit with Sacchi and Salotti's findings (2017). However, the negative association between the volatility of grants and local spending found in the special wards (but not cities) seems inconsistent with much of the literature. Moreover, the associations with the

volatility of subsidies and local bonds (with respect to public spending) in the special wards are significant at the conditional level, but this pattern does not appear in cities. Finally, in both areas, we found evidence of a positive association between the volatility of other revenues and that of public spending. It might be argued that both types of local government can utilize other revenues, whose principal part is the saving of money from the previous year, as a possible policy instrument to stabilize public spending volatility.

Table 3. Results of FE models with government size

Total expenditure *(ln)*	Model 4	Model 5 (Special wards)	Model 6 (Tama cities)
Local tax	0.1062	0.1932^{+}	-0.0654
	(0.0698)	(0.1029)	(0.0911)
Consumption tax *(ln)*	0.0675^{+}	0.1486	0.0934*
	(0.0394)	(0.0960)	(0.0388)
Other taxes *(ln)*	0.210**	0.2376	0.2667***
	(0.0766)	(0.169)	(0.0706)
Special grants	-0.0203	-0.0117	0.0106
	(0.0167)	(0.0363)	(0.0188)
Grants *(ln)*	-0.4127**	-0.7068***	-0.0708
	(0.1298)	(0.196)	(0.1619)
Subsidies	0.0202	0.1639^{*}	-0.0003
	(0.0215)	(0.0849)	(0.0189)
Fees and charges *(ln)*	0.0076	0.0699	-0.1279
	(0.0897)	(0.1742)	(0.0945)
Other revenues	0.0221***	0.0225***	0.0618**
	(0.0033)	(0.0042)	(0.0212)
Local bonds *(ln)*	0.0778**	0.0839*	0.1404
	(0.0275)	(0.0342)	(0.0918)
FTE *(ln)*	-4.5505	-2.5894	-0.1077
	(3.3702)	(8.3743)	(3.1352)
Observations	196	92	104
Coefficient of Determination	0.6167	0.7289	0.5736
n	49	23	26

Note: Significant levels are ***$p<0.001$, ** $p<0.01$, * $p<0.05$, + $p<0.1$. The denote (ln) represents the natural logarithm. Numbers in parentheses are standard errors.

In addition to the parsimonious specification (1), the literature presents a strong *prima facie* case to suggest that local size may affect the volatility of both revenues and expenditures (Sacchi & Salotti, 2017). Therefore, in Table 3, we repeat our estimations with the addition of our size proxy (FTE staff). In Table 3 Model 4, positive statistically significant associations persist between the volatility of local spending and other taxes, other revenues, and local bonds. Moreover, the volatility of expenditures is still negatively associated with the volatility of grants for the entire Tokyo municipality cohort. In general, the magnitude of the coefficients attenuates only slightly when the municipal size proxy is included.

We also stratified specification (2) and reported the same in Table 3 Model 5 for the special wards and Model 6 for the cities. In Model 5, statistically significant associations persist between the volatility of local spending and those of subsidies ($\beta=0.164$, $p<0.05$), other revenues ($\beta=0.023$, $p<0.001$), local bonds ($\beta=0.084$, $p<0.05$), and grants ($\beta=-0.707$, $p<0.001$). Once again, despite the inclusion of the local size variable, the magnitude of coefficients attenuates only slightly. Similarly, we found with Model 6 that statistically significant associations persisted for consumption tax ($\beta=0.093$, $p<0.05$), other taxes ($\beta=0.267$, $p<0.001$), and other revenues ($\beta=0.062$, $p<0.01$) and the size of the coefficients attenuated only slightly.

3.4.2 Further evidence on grants

To delve further into the negative association between the volatility of local spending and grants of the special wards, we conducted additional analysis as follows. In general, intergovernmental grants transferred from TMG to the special wards rely on the

fiscal gap between basic fiscal needs and revenues. Hence, the intergovernmental relationship is a crucial issue in grant allocation and there appears to be constant tension between competing objectives of the special wards and TMG. From the perspective of the special wards, the emphasis tends to be on expanding basic fiscal needs by incorporating expected expenditures to maximize the grants received, while basic fiscal revenues derived from fixed percentages of some principal taxes are relatively unchangeable. However, from the perspective of TMG, the focus is on constraining growth in grant distributions.

When excluding the grants, basic fiscal revenues reflect 39.5% of actual revenues and are significantly correlated with a coefficient of 0.898 (t-test=28.54), whereas basic fiscal needs account for 59.5% of actual expenditures, less than transfers to reserve funds, and their correlation coefficient is significant, with a coefficient of 0.987 (t-test=28.01). This suggests that the gap between the basic fiscal revenues and actual revenues is greater than that of the basic fiscal needs and actual expenditures. Hence, expected deficits to be covered by grants seem to be overestimated, as compared with the actual deficits. Therefore, it seems possible for special wards to save funds from grants, creating a significant amount of surplus.

The special wards do not seem to be disposed to expanding local spending despite strong surpluses. Perhaps this reticence arises because of previous experiences, such as when the local tax revenues of TMG plummeted by 20% in 2009 as a result of the 2008 GFC (Figure 1). Consequently, TMG lost around JPY 1 trillion, suffering fiscal difficulties, which also brought about a significantly reduced pool for grant distribution. Indeed, the special wards' grants decreased by around 15% in 2009 as a response to the reduction in

local tax revenues collected by TMG. It seems that special wards may have become fearful of similar sudden decreases associated with unpredictable future shocks. If this is the case, then what we observe may be a rational response by the authorities of special wards to accumulate surpluses by saving a certain part of the grants.
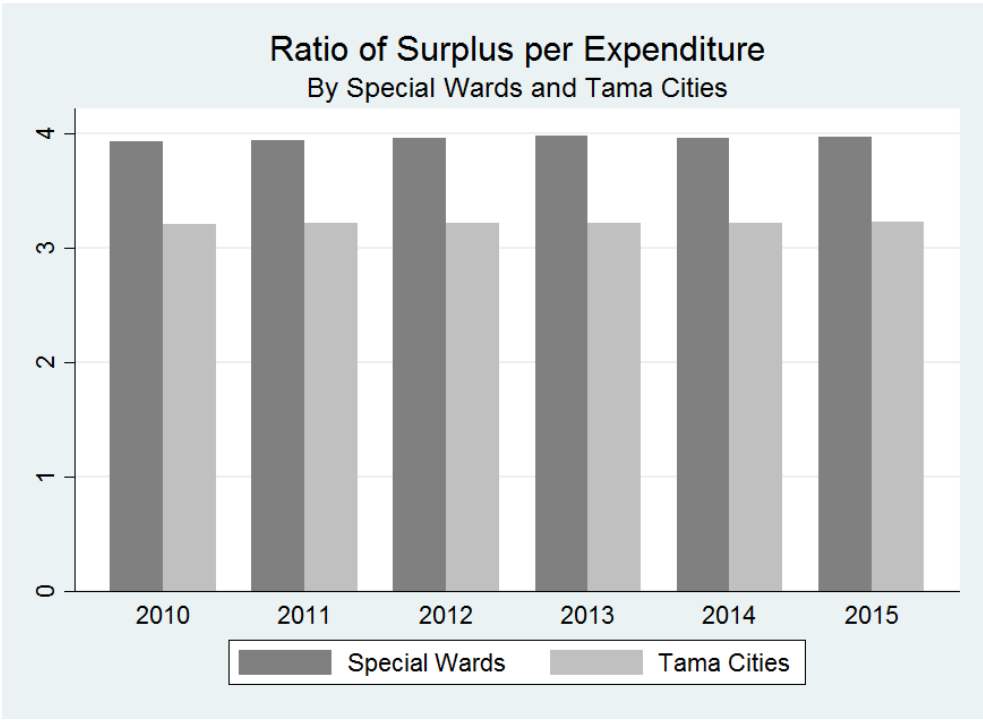
Figure 4. Ratio of surplus per expenditure between special wards and cities.

Our analysis suggests that, over 2010–2015, the surplus per expenditure by the special wards (3.92%) was higher than that of the cities (3.22%) (Figure 4). There is a significant difference between the two systems (t-test=3.77). Furthermore, the authorities of the special wards tend to prefer accumulating reserve funds for non-specific purposes (e.g., reserve funds can be withdrawn to compensate for future income shortages). The ratio of reserve funds per expenditure for the special wards is around four times higher than it is for the cities (Figure 5). Indeed, by 2015, preserved reserves for special wards (JPY 128

billion) were over six times greater than those of cities (JPY 19 billion). There is also a significant difference between the two systems (t-test=9.58). It can be inferred that special wards tend to finance their expenditure by using tax revenues and other revenues rather than grants, which are saved for reserve funds. In sum, this data seems to point to the special wards directing expanded grant revenues to reserves as a way of insulating against future financial shocks, rather than responding with higher levels of local expenditure.
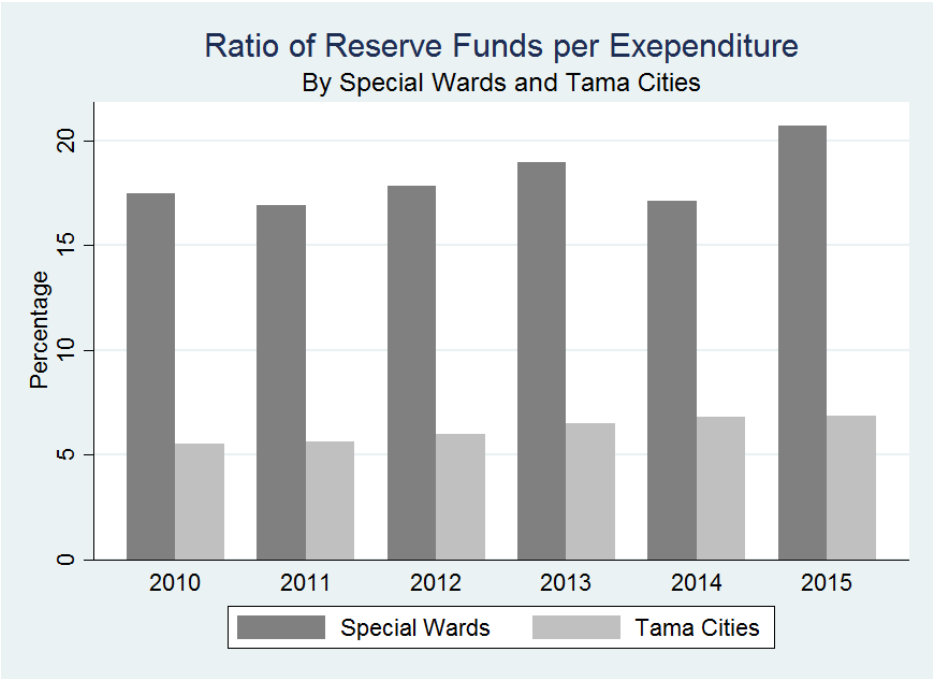
Figure 5. The ratio of reserve funds per expenditure between special wards and cities.

3.5 Policy Implication and Concluding Remarks

In this study, we examined the association between local expenditure volatility and the volatilities of various revenues for the Tokyo local governments through panel data covering 2010–2015. The evidence confirms that there are distinct differences in the determinants of volatility between the special wards and cities due to the different fiscal systems. A number of public policy implications arise from our major findings.

First, local tax volatility was found to be positively associated with local expenditure volatility in the special wards and a similar association was determined for the volatility of consumption tax and other taxes with respect to the volatility of local spending in the cities. It might be generalized that local expenditures become volatile when they are financed by tax revenues. This finding is somewhat congruous with Sacchi and Salotti's extant study (2017), arguing that there are positively significant linkages between the volatility of local expenditure and that of various own sources (local tax and consumption tax). To mitigate volatility in local expenditures in the special wards, it might be necessary for the local governments to control the volatilities of these tax revenues by focusing on tax bases that are less volatile to business cycle fluctuations, such as property tax rather than income taxes (Oates, 2011; Afonso, 2017). In doing so, local expenditures could be more stable.

Second, ordinary grants volatility had a statistically significant negative association with local expenditure volatility in the special wards, which implies that the response to greater volatility of ordinary grants is lower volatility in local spending. The evidence suggests that the special wards may be exercising anticipatory resilience concerning future economic shocks (Steccolini et al., 2017). Specifically, the special wards have attempted to accumulate reserves from the source of ordinary grants in the aftermath of the GFC, while attempting to maintain public expenditures at a certain level. This finding argues against much of the existing literature that tends to suggest that provision of intergovernmental grants exacerbates local spending (Sacchi & Salotti, 2017; Martin-Rodriguez & Ogawa, 2016; Bessho & Ogawa, 2015). This finding also strengthens the findings of Doi (1996)

and Miyana and Fukushige (2001) (cited by Shirai (2006)) that there was no "flypaper effect" in Tokyo metropolis.

Given this evidence of a saving tendency amongst special wards, TMG might respond by increasing the weight to the grant allocation (while maintaining subsidies) and, thus, mitigate spending volatility. Moreover, the practice of accumulating reserves could put the special wards in a dilemma: if ordinary grants (95%) for non-earmarked expenditures are overestimated with the intention of preserving part of the same for future contingencies, this may result in the special grant quantum (5%) being insufficient to cover specific purpose expenditures, such as disaster recovery. To cope with this possible antinomy, TMG needs to either transfer some funds from ordinary grants to special wards where there is a surplus of the former and an insufficiency of the latter (some local governments could cover the expenses for unforeseen events) or make flexible, where appropriate, the allocation rate between ordinary (becomes lower than 95%) and special grants (becomes higher than 5%) for special wards. The more the grant fluctuates, the more stable local spending becomes. Therefore, TMG could exert influence over special wards spending to be more stable and, importantly, achieve a more efficient allocation of the grants (to attain horizontal fiscal equalization) among the local governments. Moreover, because of the merits of fiscal stabilization relying on the intergovernmental grant policies in the wake of business cycle fluctuation, TMG might strengthen its justification with respect to the *raison d'etre* of tax collection on behalf of the special wards without being reconciled to tax reduction.

Third, local bonds of special wards have a statistically significant positive impact on expenditure volatility, although the magnitude of the association is relatively small. Several municipalities can utilize bonds as fiscal adjustments to redress short-run deficits. However, a side-effect of this practice might be to exacerbate local spending volatility. Therefore, trade-off issues between costs and benefits should be carefully considered before the bonds are issued.

Finally, the methodology outlined in this paper would be suitable for studies abroad and examining expenditure volatility in other contexts would allow scholars to identify the effect of different revenue structures. Comparative analyses, particularly with jurisdictions such as Australia that operate distinct intergovernmental grant transfer systems, would allow scholars to further explore the importance of grants as a determinant of expenditure volatility.

CHAPTER 4. PUBLIC EFFICIENCY IN TOKYO METROPOLITAN LOCAL GOVERNMENTS

The previous chapter examined the translation of various collected revenue volatilities into local expenditures volatility. However, further research on how well these permitted expenditures are used for provision of public goods and services is still required. Hence, this study aims to investigate the concept of municipal efficiency, in which local expenditures as inputs yield the public service delivery as outputs for 49 Tokyo local governments over the 2001–2015 period. In doing so, we use the non-parametric linear programming approach, Data Envelopment Analysis (DEA), to estimate the level of efficiency and, furthermore, make a comparison of efficiency scores before and after the introduction of accrual accounting in 2008 at the local level.

4.1 Introduction

Many advanced economies worldwide (e.g., United Kingdom, Australia, New Zealand, Japan) have properly implemented a policy of efficiency promotion at the local level, following the NPM tenets, since the late 1980s. The main principle of NPM theory is the adoption of private sector practices at the public sector level, with an emphasis on the measurement of performance or efficiency and greater emphasis on output controls for better resource allocation, thus optimizing the outputs and outcomes of a PSO's production outcomes (Hood, 1991, 1995). Hence, an evaluation of public efficiency and of the implications for efficiency enhancement become essential for both local authorities and public policymakers alike.

In practice, local governments have come under pressure to respond to ever-increasing citizen demands in the presence of mounting financial constraints. On one hand, the rising living standards of the people consequently lead to demands not only for higher quality but also larger quantities of public goods and services provisions (Zhu & Peyrache, 2017). On the other, local governments are confronted with budget austerity and resource constraints resulting from the economic and financial crisis (Cordero et al., 2017; Storto, 2016; Da Cruz & Marques, 2014). Thus, in the pursuit of greater operational public efficiency, local authorities in recent decades have fundamentally responded in two ways: (1) amalgamation or merger of local governments, and (2) decentralization of fiscal power (Storto, 2013; 2016). In the first strand, municipalities in Australian states such as New South Wales, Queensland, and Western Australia have undergone municipal amalgamation or consolidation of small councils into jurisdictions for substantial cost-saving (in turn increasing public efficiency) due to the economies of scale in service provisions (Drew, Dollery, & Kortt, 2016; Drew, Kortt, & Dollery, 2017; Fogarty & Mugera, 2013). In the second strand, some administrative governments such as Japan enacted policies targeted towards fiscal decentralization, including the "Promotion Law of Decentralization" in 1995 and the later "Trinity Reform" in 2004, that concentrated on eliminating the role of agency-delegated functions, reducing the reliance on central government, strengthening the fiscal power of local revenues, and allowing more discretion in local spending (Mochida, 2008).

The focus of this research is on the second strand — fiscal decentralization of Japanese local governments, particularly Tokyo municipalities — which is motivated by two reasons. First, in the corpus of empirical studies on efficiency in Japan, though there

are a number of academic works measuring public efficiency at the prefectural level, few studies exist measuring public efficiency for Tokyo municipalities. Second, despite the adoption of the new accrual accounting systems in 2008, there is little evidence for the implications for municipal efficiency. Therefore, our study attempts to address this gap in the literature and directly aims to answer the following research question: how have the trends or patterns of public (technical) efficiency scores of Tokyo local governments (23 special wards and 26 cities) over the 2001–2015 period changed before and after the introduction of accrual accounting.

The remainder of this chapter is structured as follows. Section 2 reviews previous empirical studies on local government efficiency, particularly focusing on the Tokyo public sector setting. Section 3 presents the DEA methodology. Section 4 provides empirical results of the productive efficiency analysis. The final section consists of remarks and our study's further research.

4.2 Public sector setting

Public efficiency is concerned with the functional relation between public sector inputs and outputs. Indeed, this involves the conversion of compositional resources as capital and human into public initiatives, programs, and services. The tactics associated with efficiency enhancement either diminish inputs while maintaining constant output levels or increase the quantity of outputs as much as possible with the given level of inputs. Accordingly, given the relative inability of municipalities to influence output levels (due to legal requirements, public mandate, etc.), the minimization of input resources to maximize

productive efficiency (input-oriented approach) is a common practice in the public sector's "do more with less" context (see Da Cruz & Marques, 2014).

Regarding public finance at the local level, each municipality is primarily responsible for revenue raising and public service spending decisions. Tokyo local governments earn their income through various kind of taxes (e.g., corporation tax, inhabitant tax, consumption tax) in addition to intergovernmental grants and subsidies from TMG and the national government as well as municipal bonds (Tran, Drew, & Noguchi, 2018). Conversely, Tokyo local governments are in charge of providing a bundle of public services, including mandatory schooling for children under 15, public social security, public health, local infrastructure, and urban planning (Mochida, 2008). By virtue of the decentralization reforms since 2000, local authorities in Japan have been given more discretion on expenditures for public service delivery (Mochida, 2008) with an assumption that higher fiscal power can facilitate positive consequences of public efficiency.

Given that local governments are responsible for public service provision, constituencies have prioritized creating "Value for Money" (VfM), maximizing the benefits residents receive from payments of taxes, fees, and charges. Hence, the topic of public efficiency and performance has attracted scholarly attention internationally to offer some policy implications of VfM for local authorities and practitioners. For example, Storto (2013) evaluated the technical efficiency of major municipalities in Italy. Afonso and Fernandes (2006) conducted a study on the efficiency of 51 Portuguese municipalities in the region of Lisbon and Vale do Tejo and concluded that Vale do Tejo municipalities could improve performance without necessarily increasing public expenditure. Doumpos

and Cohen (2014) analyzed the efficiency of local Greek governments based on accrual accounting data of local governments. The findings of Cuadrado-Ballesteros, García-Sánchez, and Prado-Lorenzo (2013) showed the relationship between public services delivery and efficiency of local governments in Spain, where the functional decentralization and externalization have a negative association with efficiency. In Japan, there has been limited research regarding municipal efficiency, although academic attention has been increasing in recent years. For instance, Nijkamp and Suzuki (2009) delineated the efficiency score for cities in Hokkaido prefecture; Fukuyama et al. (2017) investigated the 47 Japanese prefectures, observing average efficiency of 81.8% and determining that Okinawa prefecture was considered the most efficient in Japan. However, the investigation into public efficiency of Tokyo local governments has been scant so far. Therefore, we conducted an empirical analysis of municipal efficiency in Tokyo metropolis.

4.3 DEA method

The measurement of productive efficiency has an origin in the seminal works of Farrell (1957), introducing *technical efficiency, allocative efficiency,* and *overall efficiency* of firms. A non-parametric method to estimate the efficiency of production using the linear programming approach where data are enveloped within the estimated boundary or production frontier, known as Data Envelopment Analysis (DEA), was initially developed by Charnes, Cooper, and Rhodes (1978) (hereafter referred as CCR model). The model to evaluate *relative efficiency* with multiple inputs and outputs introduced by Charnes et al. (1978) is also known as the constant returns to scale (CRS) model. The model has been

extended with the addition of variable returns to scale (VRS) by Banker, Charnes, and

Cooper (1984) (hereafter abbreviated as BCC model) (Figure 6).

The DEA estimator is characterized by a set of technology assumptions (or axioms):

(A1) "no free lunch"; (A2) producing nothing is possible; (A3) boundedness of the output

sets; (A4) closedness of the technology set; (A5) free disposability of all inputs; (A6) free

disposability of all outputs. Axiom 1 means that there is no possibility of generating outputs

without using any inputs. However, it is possible to produce no outputs with given inputs in

axiom 2. Axiom 3, associated with the boundedness concept, means that the technology

frontier is made feasible and bounded. It is possible to seek the optimal position of the

technology compact set presented, given axiom 4. Axioms 5 and 6 mention the capability to

alter production levels without constraint (Sickles & Zelenyk, 2019, p. 15).
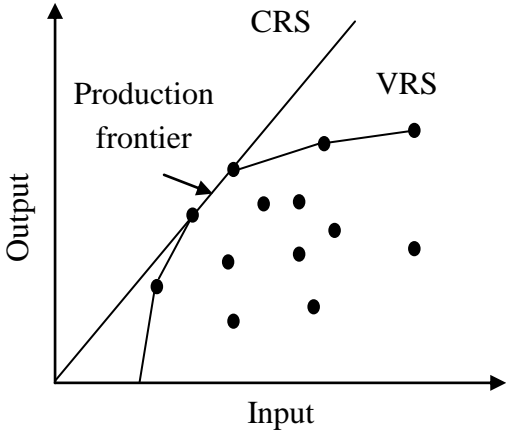
Figure 6. Production set for CCR and BCC model.

Alongside these assumptions, the DEA method benchmarks the relative efficiency

of a group of entities (e.g., firms, institutions, local governments), also called decision-

making units (DMUs), associated with numbers of inputs (resources) and outputs (products

or services). We denote the number of DMUs, inputs, and outputs as *n, m and s,*

respectively. For each DMU$_j$ (j =1,2,…,n), there exist *m* inputs x$_{ij}$ (i=1,2,…,m) and *s*

outputs y_{rj} (r=1,2,...,s). Let the weight of the non-negative value of inputs and outputs be

$v=(v_1,v_2,..., v_m)$ and $u=(u_1,u_2,..., u_s)$, respectively. The mathematical problem of the linear

programming problem of the CCR model to estimate efficiency θ is as follows:

$$\max_{v,u} \theta = \frac{\sum_{r=1}^{s} u_r y_{r0}}{\sum_{i=1}^{m} v_i x_{i0}} \qquad \text{Subject to} \quad \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \ \ (j=1,2,...,n)$$

$$v \geq 0, u \geq 0$$

In which θ is a scalar representing the technical efficiency score for each DMU,

ranging from 0 to 1. If a DMU attains a score of 1, it would be positioned on the production

frontier boundary, considered the best practice. Otherwise, it would be positioned inside the

boundary and considered inefficient. When we constrain the $\sum_{i=1}^{m} v_i y_{i0} = 1$ and λ intensity

vector, the mathematical problem of the input-oriented CCR model can be simplified as

follows:

CCR model - Multiplier form	CCR model - Envelopment form
$\max_{v,u} \theta = \sum_{r=1}^{s} u_r y_{r0}$ Subject to $\sum_{i=1}^{m} v_i y_{i0} = 1$ $\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0 (j=1,2,...,n)$ $v \geq 0, u \geq 0$	$\min_{\lambda} \theta$ Subject to $\sum_{j=1}^{n} x_{ij}\lambda_j - \theta x_{i0} \leq 0 (i=1,2,...,m)$ $\sum_{j=1}^{n} y_{rj}\lambda_j \geq y_{r0} (r=1,2,...,s)$

The BCC model is extended by constraining the sum of intensity value λ_j to be equal to 1; then, the envelopment form of the input-oriented BCC model is similar to the input-oriented CCR model with the additional condition $\sum_{j=1}^{n} \lambda_j = 1$.

BCC model - Multiplier form	BCC model - Envelopment form
$\max_{v,u} \theta = \sum_{r=1}^{s} u_r y_{r0}$ Subject to $\sum_{i=1}^{m} v_i y_{i0} = 1$ $\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0 (j = 1, 2, ..., n)$ $\sum_{j=1}^{n} \lambda_j = 1$ $v \geq 0, u \geq 0$	$\min_{\lambda} \theta$ Subject to $\sum_{j=1}^{n} x_{ij} \lambda_j - \theta x_{i0} \leq 0 (i = 1, 2, ..., m)$ $\sum_{j=1}^{n} y_{rj} \lambda_j \geq y_{r0} (r = 1, 2, ..., s)$ $\sum_{j=1}^{n} \lambda_j = 1$

In economic discourses on public efficiency, *technical efficiency* has been widely adopted in the public sector rather than *price or allocative efficiency*. This is because the nature of the market does not exist in the public sector as local governments are largely unable to determine or influence the market price for public goods and services like in the private sector (Fried, Lovell, & Schmidt, 2008). Furthermore, Farrell (1957, p. 261) argues that "price efficiency is […] both unstable and dubious of interpretation" in a public sector setting. Hence, the technical component of economic efficiency is preferred to quantify input and output data in the public sector. Moreover, we adopt the input-oriented estimation

of efficiency with the assumption that local authorities have a relatively greater ability to control the level of inputs (e.g., public money, staff) compared to outputs (e.g., public services provisions to people), which was affirmed by Doumpos and Cohen (2014), "input minimization is more suitable compared to output maximization", and Worthington and Dollery (2001, p. 235), who explain that "local governments take outputs as exogenous and have a larger degree of control over the level of inputs, especially within functional areas". Consequently, this study uses input orientation under the CCR and BCC model for public efficiency estimation.

Local governments in Tokyo are composed of 23 special wards, 26 suburban cities, five towns, one village, and several islands in the Tokyo area. Since the five towns, one village, and several islands account for a relatively insignificant proportion of public finance (as opposed to 23 special wards and 26 cities), we will restrict our analysis to these latter 49 administrative units. Data on these entities was collected these units over a 15-year period (2001–2015) through the Tokyo Metropolitan Annual Report (demographic data, public service provision data) and Tokyo Statistical Yearbook (financial data) published by the Bureau of General Affairs, TMG.

4.4 Empirical results

4.4.1 Descriptive statistics

In general, non-parametric mathematical programming does not deliberately restrict the variables of inputs and outputs when estimating DEA scores; hence, the selection of these variables is contingent on relevant data availability and specific local features. For Tokyo, we select total expenditures excluding personnel expenditures, consistent with prior

literature (e.g., Cordero et al., 2017; Nijkamp & Suzuki, 2009), and personnel expenditures (e.g., Balaguer-Coll et al., 2013; Cordero et al., 2017; Fogarty & Mugera, 2013). These inputs are measured in monetary terms (cash-based accounting data). The average of total expenditure, including operating, investment, and financial components (excluding the personnel expenditure), is roughly six times the average of personnel expenditure (see Table 4). It is important to note that deflated expenditures with the Consumer Price Index (CPI) were an indicator for the Tokyo market for a global DEA analysis.

Concerning variables of production outputs, we select eight elements reflecting the bulk of local service activities: education (schools and students), culture and recreational facilities (libraries), communal services (waste collection), infrastructure (roads), administrative services (number of building permits issued), business development (number of enterprises established), and demographics (percentage of people over 65 per total population) in accordance with earlier literature presented by Narbón-Perpiñá & De Witte (2018). The first two output variables representing educational services are the number of schools at the level of kindergarten, elementary school, and secondary education per 10,000 persons and the number of students at the same level. In fact, there are approximately three educational institutions serving nearly a thousand students in Tokyo local governments. The third variable is the number of public libraries per 1000 people (on average, it was nearly eight libraries per 1000 people). Fourth, garbage collection is defined as the amount of garbage collected in each municipal area per year. Fifth, the number of building approvals represents administrative services and planning activities of the municipality. On average, each Tokyo local government issues one thousand permits per year. Sixth, the length of roads indicates the basic municipal infrastructure level, reflecting the maintenance

requirement for each municipality. Seventh, the number of businesses established per year accounts for the business development in a municipal area. Indeed, nearly 500 companies are approved on average, but with a large standard deviation. The final output is the percentage of people over 65 per total population, which is used to proxy services and facilities primarily targeted towards the elderly (such as welfare and medical care).

Table 4. Inputs and outputs for efficiency model (n = 735)

Variable	Definition	Mean	Standard deviation
Inputs			
Total expenditure	Cash-based expenditure for the fiscal year excluding personnel expenses (thousand yen)	75,543.93	52,634.37
Personnel expenditure	Salary expenditure for personnel (thousand yen)	12,408.5	9,100.063
Outputs			
Schools	Number of schools per population (of 10,000 people)	2.81	0.98
Students	Number of students per population (of 10,000 people) including kindergarten, elementary, and secondary school students	918.65	323.52
Library	Number of public libraries per 1000 people	7.84	3.96
Garbage collection	Amount of garbage collected (tons)	70,503	52,154
New Building Approval	Number of construction works approved	1034.33	851.79
Road	Length of roads (km)	400,762	309,078
Business	Number of business established	496	723
Aged	Percentage of people over 65 per total population (%)	19.21	2.70

To reiterate, this study employed the two inputs and eight outputs mentioned above to estimate efficiency scores. Although we recognize that these measures cannot represent the full range of services provided by Tokyo local governments, due to the data and variable limitations, proxies must be employed. We note our approach is consistent with scholarship in local government efficiency measurement as in Narbón-Perpiñá & De Witte (2016, 2018).

4.4.2 Public efficiency results

The efficiency scores estimated through the input-oriented CRS and VRS technology are exhibited in Figure 7 for special wards and cities. On one hand, Figure 7a indicates that the CRS efficiency scores in most of the special wards (e.g., Chiyoda, Chuo, Minato) have been substantially decreased, whilst others (e.g., Taito, Sumida, Itabashi) have experienced fluctuated efficiency. Only the Edogawa wards seem to have maintained a stable efficiency score. In contrast, Figure 7b shows that the CRS efficiency scores in cities typically increased (e.g., Musashino, Ome, Tama), whilst others attained a stable or slightly upward productive efficiency trend (e.g., Mitaka, Hino). Furthermore, it should be noted that two remote local governments, Hamura and Akiruno, have achieved the "best practice" of efficiency among their peers where they achieved a value of 1. Furthermore, we can observe that Ome, Machida, and Musashimurayama *inter alia* obtained the near "best practice" in the technology frontier besides Hamura and Akirumo. Similarly, Figures 7c and 7d represent the VRS efficiency scores, which exhibited an identical pattern to the CRS efficiency scores in Figures 7a and 7b.
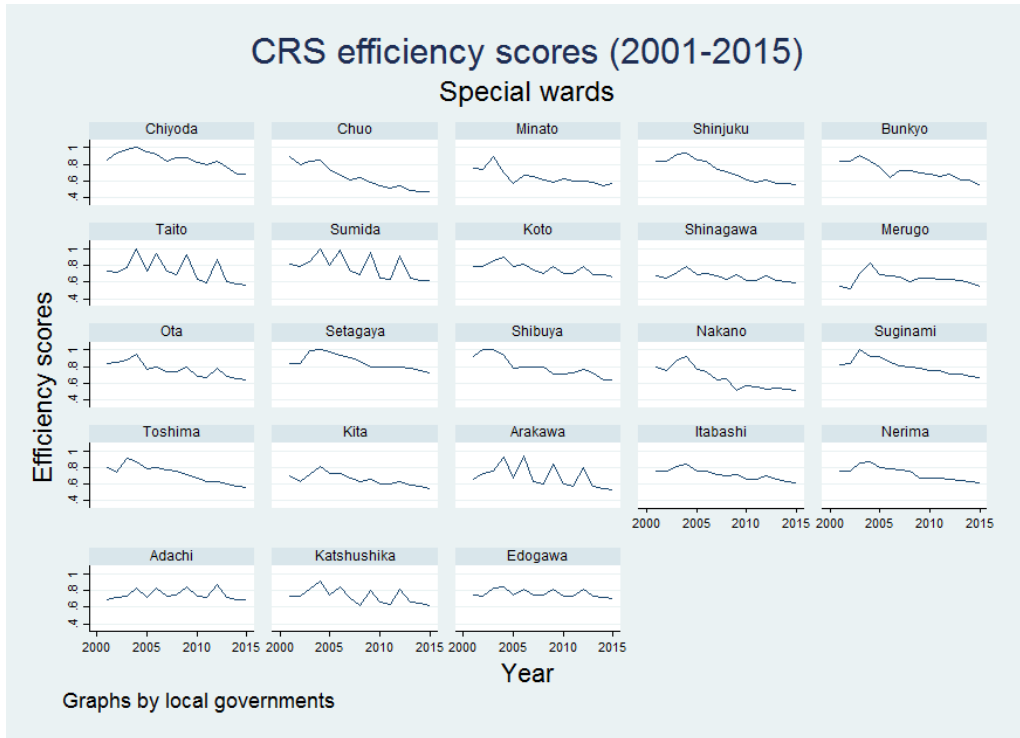
(a)
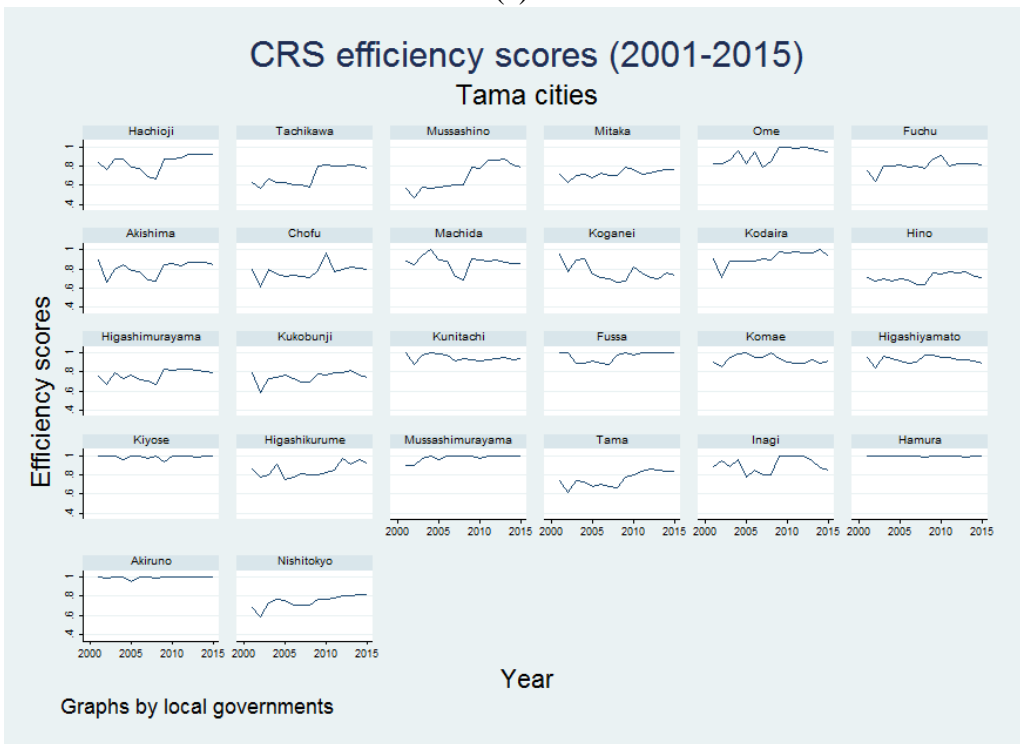
(b)
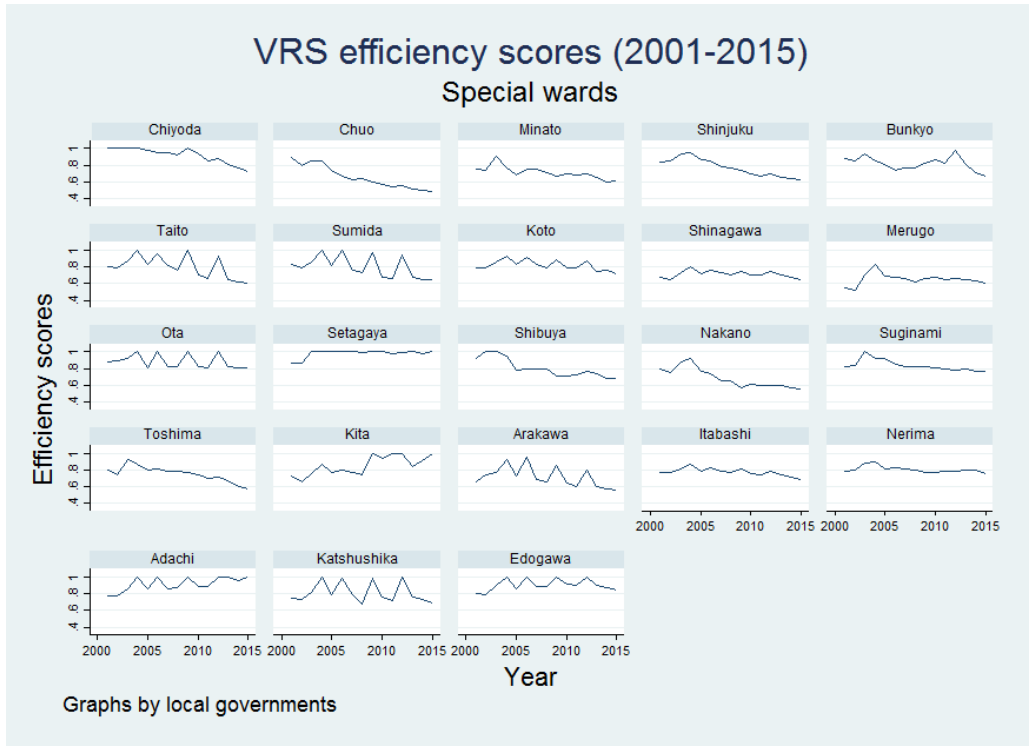
(c)
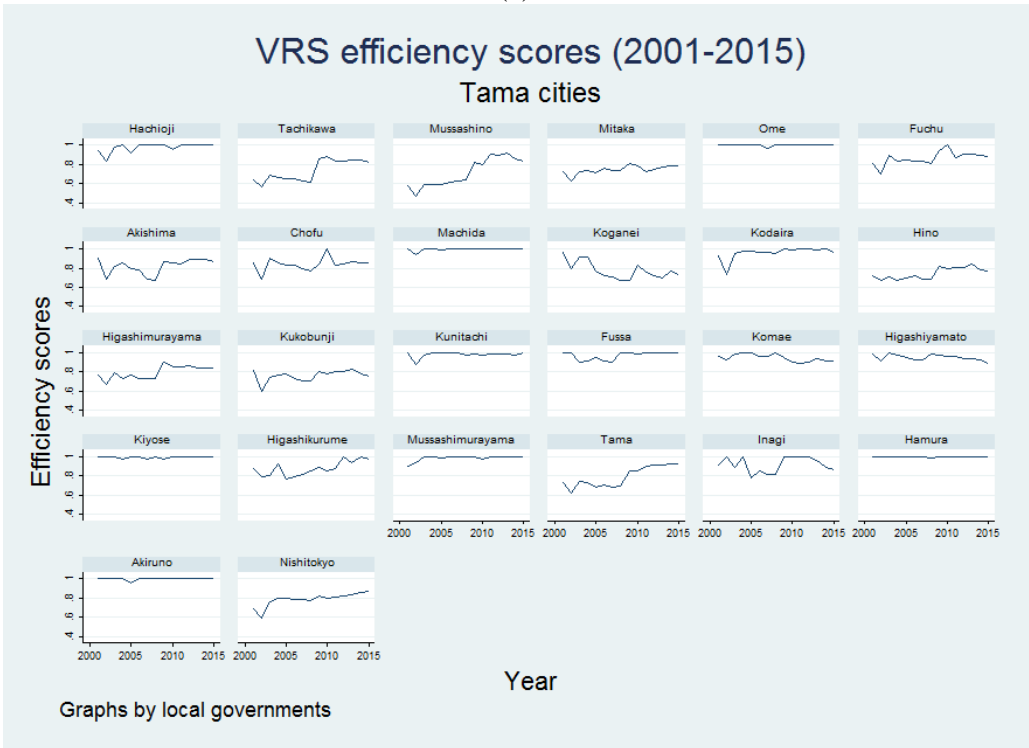
(d)

Figure 7. CRS and VRS efficiency scores of 23 special wards and 26 cities.
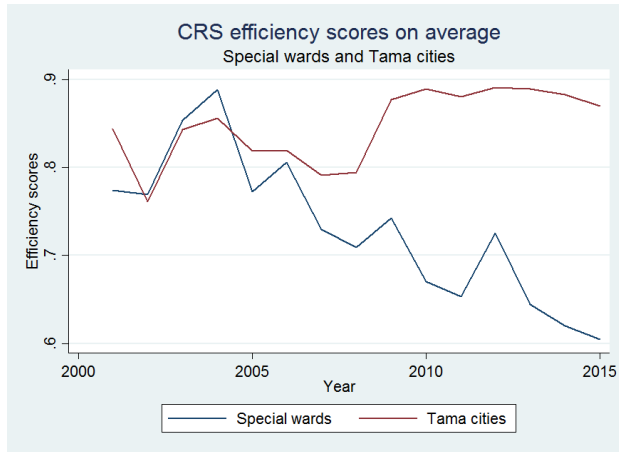
(a)
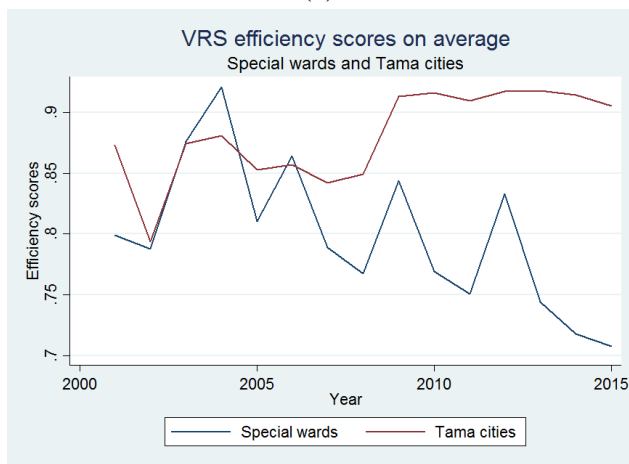
(b)
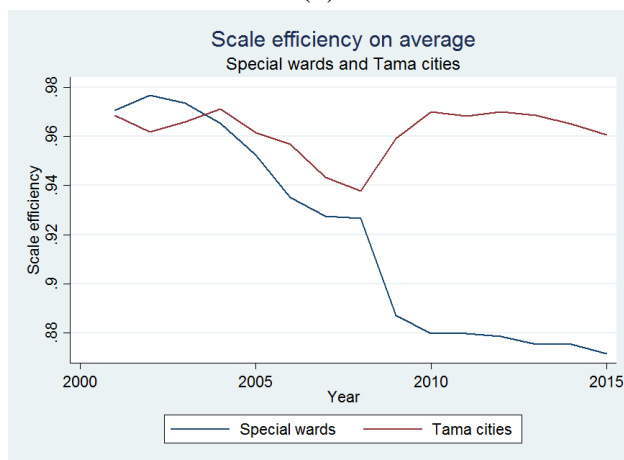
(c)

Figure 8. CRS, VRS, and scale efficiency scores between special wards and Tama cities.

By comparing the overall performance of the two groups, special wards and Tama cities, we take the consideration trends over the time period examined. In general, the consistency of estimated efficiency among CRS and VRS technology is again confirmed, as shown in Figures 8a and 8b respectively. The overall public efficiency of special wards within the period of 2001–2015 tended to fluctuate and typically declined in contrast to the cities, which increased in general. Despite declining and unstable efficiency scores before 2008, Tama cities have attained relatively stable and high efficiency scores, particularly after 2008 (approximately to 0.90). It could be conjectured from our findings that administrative operations and management of cities have been more efficient than those of special wards. It is reasonable to deduce that cities, where less affluent local governments of Tokyo metropolis are mostly found in suburban areas, have a greater tendency to operate and manage their production in an efficient way compared to special wards.

Table 5 exhibits the average value of efficiency for the entire Tokyo municipalities, special wards, and Tama cities over 2001–2015 in terms of CRS and VRS technological and scale efficiency. On average, efficiency scores are in the range of 0.7–0.8 for special wards and 0.8–0.9 for the cities. In greater detail, efficiency scores range from 0.604–0.889 for CRS technology and 0.708–0.921 for VRS technology for the special wards. The lowest and highest estimated efficiency occurred in 2015 and 2004, respectively. In contrast, for cities, efficiency scores range from 0.761–0.889 for CRS technology and 0.794–0.918 for VRS technology. The lowest and highest estimated efficiency occurred in 2002 and 2013, respectively. Accordingly, we can observe an inverse pattern of efficiency scores, in that

special wards' efficiency tended to be higher before 2008 and lower after 2008, whereas cities were the opposite.

Table 5. CRS, VRS, and Scale efficiency scores

Year	Tokyo metropolis			Special wards			Tama cities		
	CRS	*VRS*	*Scale*	*CRS*	*VRS*	*Scale*	*CRS*	*VRS*	*Scale*
2001	0.811	0.838	0.969	0.774	0.799	0.970	0.844	0.873	0.968
2002	0.765	0.791	0.969	0.769	0.788	0.977	0.761	0.794	0.962
2003	0.848	0.875	0.970	0.853	0.876	0.974	0.843	0.874	0.966
2004	0.871	0.900	0.968	0.889	0.921	0.965	0.856	0.881	0.971
2005	0.797	0.833	0.957	0.772	0.811	0.952	0.819	0.853	0.961
2006	0.813	0.860	0.947	0.805	0.864	0.935	0.819	0.857	0.957
2007	0.762	0.817	0.936	0.729	0.789	0.927	0.791	0.842	0.943
2008	0.754	0.811	0.932	0.709	0.767	0.927	0.794	0.849	0.938
2009	0.814	0.881	0.925	0.743	0.844	0.887	0.877	0.914	0.959
2010	0.786	0.847	0.928	0.670	0.769	0.880	0.889	0.916	0.970
2011	0.774	0.835	0.927	0.653	0.751	0.880	0.881	0.910	0.968
2012	0.813	0.878	0.927	0.725	0.833	0.879	0.891	0.918	0.970
2013	0.774	0.837	0.925	0.644	0.744	0.875	0.889	0.918	0.968
2014	0.759	0.822	0.923	0.620	0.718	0.875	0.883	0.914	0.965
2015	0.745	0.813	0.919	0.604	0.708	0.872	0.870	0.906	0.961
Mean	0.792	0.843	0.941	0.731	0.799	0.918	0.847	0.881	0.962

Along with CRS and VRS technology, scale efficiency was also estimated to measure the scale effect of each local unit. Scale efficiency is simply defined as dividing CRS efficiency scores by VRS efficiency scores to measure the scale effects or effect of size (scale) on efficiency. Table 6 indicates that the scale efficiency of special wards (0.918) is generally lower than that of cities (0.962), indicating that special wards are generally

further from the optimal size that would maximize efficiency. The trend of scale efficiency can be seen through Figure 8c, which shows that special wards' scale efficiency substantially declined while cities experienced an abrupt increase in 2008. One potential explanation for the special wards result may be the additional spending on public service provisions since the imposition of the accrual accounting system; however, without additional evidence this claim cannot be proven. Nonetheless, whilst empirical evidence indicates that reliance on an accrual basis can support local administrators to manage resources more efficiently (for example, Lampe et al., 2016), this proposition does not seem to have occurred in the case of the special wards.

Table 6. CRS and VRS efficiency scores before and after 2008

ID	Name (Wards)	CRS	Before 2008	After 2008	Status	VRS	Before 2008	After 2008	Status
1	Chiyoda	0.858	0.921	0.787	▼	0.919	0.976	0.854	▼
2	Chuo	0.647	0.756	0.521	▼	0.655	0.759	0.536	▼
3	Minato	0.647	0.700	0.587	▼	0.710	0.756	0.659	▼
4	Shinjuku	0.726	0.837	0.599	▼	0.772	0.855	0.676	▼
5	Bunkyo	0.720	0.788	0.642	▼	0.820	0.826	0.813	▼
6	Taito	0.738	0.790	0.678	▼	0.799	0.854	0.737	▼
7	Sumida	0.779	0.833	0.716	▼	0.800	0.846	0.747	▼
8	Koto	0.761	0.797	0.719	▼	0.818	0.841	0.792	▼
9	Shinagawa	0.663	0.690	0.632	▼	0.714	0.722	0.704	▼
10	Meguro	0.636	0.654	0.616	▼	0.651	0.657	0.645	▼
11	Ota	0.763	0.821	0.697	▼	0.881	0.891	0.869	▼
12	Setagaya	0.851	0.917	0.774	▼	0.976	0.963	0.991	▲
13	Shibuya	0.796	0.879	0.702	▼	0.803	0.879	0.716	▼
14	Nakano	0.661	0.767	0.539	▼	0.682	0.769	0.583	▼
15	Suginami	0.799	0.868	0.721	▼	0.834	0.873	0.790	▼
16	Toshima	0.722	0.805	0.627	▼	0.754	0.816	0.682	▼
17	Kita	0.655	0.703	0.600	▼	0.852	0.761	0.956	▲
18	Arakawa	0.692	0.739	0.638	▼	0.718	0.764	0.665	▼

19	Itabashi	0.715	0.761	0.663	▼	0.773	0.796	0.747	▼
20	Nerima	0.727	0.793	0.652	▼	0.806	0.827	0.781	▼
21	Adachi	0.753	0.752	0.754	▲	0.916	0.875	0.963	▲
22	Katsushika	0.733	0.767	0.693	▼	0.814	0.819	0.809	▼
23	Edogawa	0.767	0.779	0.753	▼	0.905	0.893	0.918	▲
	Cities								
24	Hachioji	0.841	0.783	0.907	▲	0.976	0.960	0.993	▲
25	Tachikawa	0.702	0.614	0.802	▲	0.735	0.637	0.847	▲
26	Musashino	0.689	0.571	0.825	▲	0.716	0.588	0.863	▲
27	Mitaka	0.723	0.697	0.752	▲	0.745	0.721	0.773	▲
28	Ome	0.916	0.860	0.980	▲	0.998	0.997	1.000	▲
29	Fuchu	0.803	0.771	0.839	▲	0.863	0.821	0.912	▲
30	Akishima	0.808	0.764	0.857	▲	0.821	0.774	0.874	▲
31	Chofu	0.772	0.729	0.821	▲	0.841	0.816	0.870	▲
32	Machida	0.867	0.856	0.880	▲	0.995	0.991	1.000	▲
33	Koganei	0.767	0.793	0.738	▼	0.776	0.807	0.741	▼
34	Kodaira	0.916	0.869	0.969	▲	0.960	0.932	0.992	▲
35	Hino	0.711	0.677	0.750	▲	0.745	0.695	0.802	▲
36	Higashimu-rayama	0.769	0.726	0.817	▲	0.793	0.737	0.856	▲
37	Kokubunji	0.747	0.717	0.781	▲	0.759	0.733	0.790	▲
38	Kunitachi	0.947	0.957	0.935	▼	0.981	0.978	0.984	▲
39	Fussa	0.962	0.932	0.998	▲	0.972	0.949	0.999	▲
40	Komae	0.932	0.952	0.910	▼	0.948	0.976	0.916	▼
41	Higashiyamato	0.930	0.923	0.938	▲	0.948	0.955	0.939	▼
42	Kiyose	0.990	0.991	0.990	▼	0.994	0.993	0.995	▲
43	Higashiku-rume	0.849	0.812	0.891	▲	0.874	0.826	0.929	▲
44	Musashimu-rayama	0.980	0.967	0.995	▲	0.985	0.975	0.996	▲
45	Tama	0.755	0.692	0.828	▲	0.790	0.696	0.896	▲
46	Inagi	0.906	0.864	0.954	▲	0.917	0.880	0.959	▲
47	Hamura	0.998	0.998	0.998	▲	0.998	0.998	0.999	▲
48	Akiruno	0.996	0.993	1.000	▲	0.997	0.995	1.000	▲
49	Nishitokyo	0.748	0.707	0.795	▲	0.788	0.750	0.832	▲

Note: ▲ increasing, ▼ decreasing. From ID 1 to ID 23 are special wards, while from ID 24 to ID 49 are cities.

On an individual local government level, Table 6 displays efficiency scores on average for CRS and VRS technology before and after 2008. First, the results of special wards display the range 0.647–0.858 for CRS technology and 0.651–0.919 for VRS technology, in which Chuo seems to be the highest efficient local administrative unit amongst special wards. Second, efficiency scores for Tama cities are in the range 0.689–0.998 for CRS technology and 0.716–0.998 VRS technology, in which Hamura and Akiruno are likely to be the most efficient local units among the 26 cities.
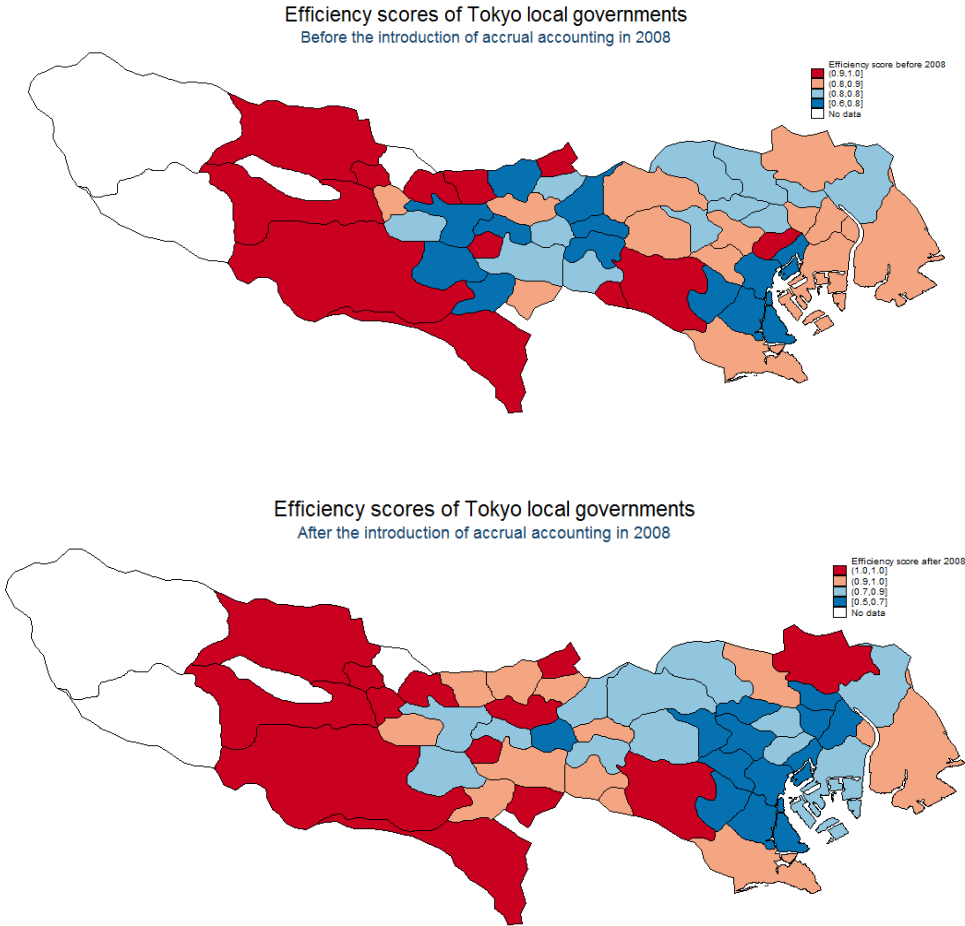
Figure 9. VRS efficiency scores of Tokyo local governments before and after 2008

It is notable in Table 6 that whilst only one special ward has increased its efficiency, 23 cities have increased their efficiency (only three cities have declined) since the introduction of accrual accounting. In other words, efficiency scores move in diverse ways, where most cities have improved their efficiency performance but most special wards have deteriorated. To more clearly detect the change, we model the spatial pattern of efficiency scores with VRS technology before and after 2008, as illustrated in Figure 9. Municipalities shaded in red indicate a higher level of efficiency and those in blue show a lower level of efficiency. Figure 9 reveals that Tama cities on the western side of the Tokyo area display higher and increasing efficiency levels, while special wards on the eastern side display lower efficiency and declining efficiency levels. In a nutshell, special wards tend to record lower efficiency outcomes as contrasted with the efficiency improvements enjoyed by cities.

4.4.3 Further analysis of scale effects

Scale effects are determined in relation to the municipal population. Figure 10 shows the relationship between scale efficiency and the natural logarithm of the population for the entire sample of Tokyo local governments. The scale effect was found to be relatively strong for local governments with large populations (over 160,000 persons corresponding to ln(population) = 12), while optimal scale is indicated in a range of local governments with populations ranging from 60,000 to 160,000. This finding could be important for further interpretation with respect to population in the next chapter.
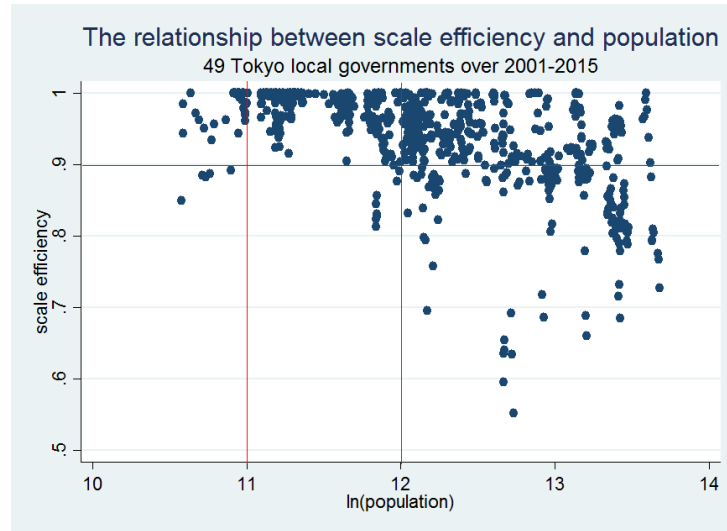
Figure 10. The relationship between scale efficiency and natural log of population

4.4.4 Further analysis of operating expenditure

To eliminate investment and financing activities (expenditure) in total expenditure, we take the operating expenditure into account as an input with the aim of estimating efficiency scores reflected by operational activities (in the data analysis, there are four local governments reported in a general account against 45 local governments reported in a normal account). Due to the data availability, these local governments were pooled together. However, supplementary analysis indicates no significant difference between a general and normal account. Figure 11 illustrates operating expenditure accounts for most of the total expenditure, which is nearly 80%, and increasing annually after 2008. The remaining components of total expenditure are investment and finance expenses of approximately 10%. As shown in Figure 12, the efficiency scores in terms of operating expenditure as an input are consistent with those in terms of total expenditure. Indeed, there remains a substantial distinction in efficiency outcomes between special wards and cities. While efficiency scores in cities are stable and relatively high at 0.9, those in special wards

are volatile, ranging from 0.65 to 0.85 and having declined within 2008–2015. Essentially, consistency was found in efficiency trends regardless of whether total or operating expenditures were used as a production input.
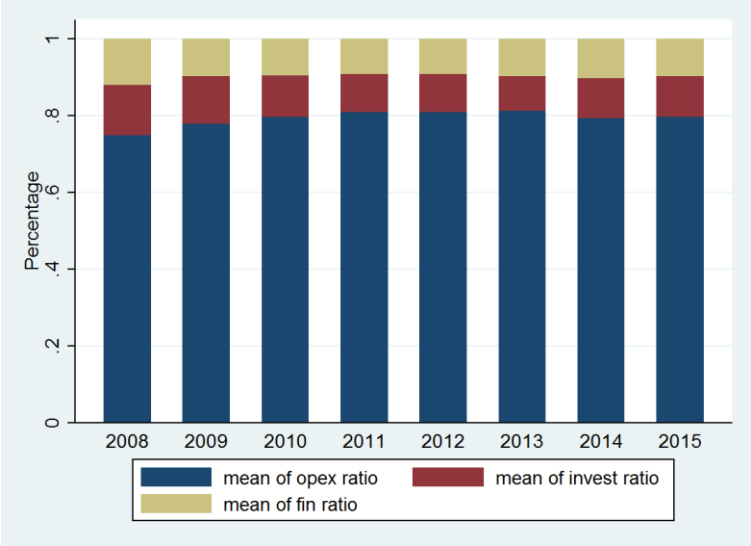
Figure 11. Weighted elements of operating, investment, and finance expenditure
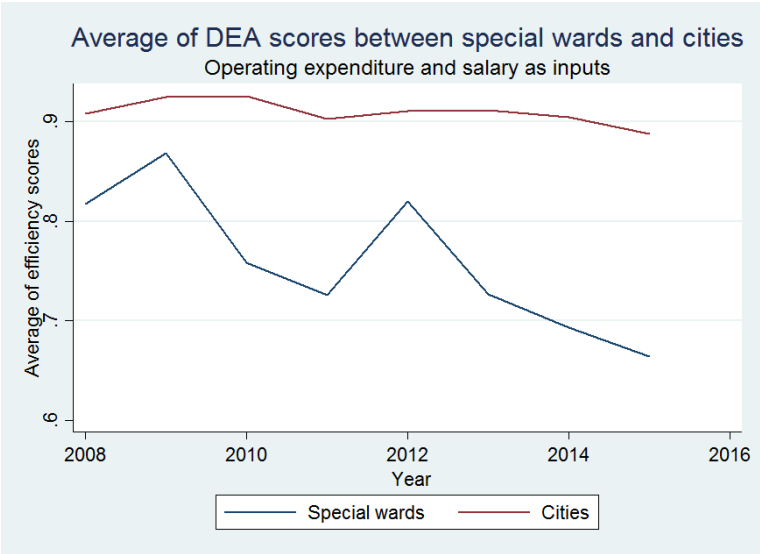
Figure 12. DEA scores in terms of operating expenditure and salary as inputs

4.5 Conclusions

This exploratory study sets out to critically investigate the public efficiency of Tokyo municipalities from 2001 to 2015. The result indicates that the efficiency scores substantially decreased, specifically in the 23 special wards, when the new accounting system was applied in 2008, while a majority of the municipalities in the 26 cities experienced increases in efficiency in public service provision. Lampe et al.'s prior literature (2015) maintains that municipalities have reduced their cost inefficiency since the adoption of accrual basis and found some determinants associated with promoting an increase in public efficiency. Our findings concerning Tokyo should be extended in future studies to investigate the effects of accrual accounting systems and examine the rationale behind why efficiency scores in cities generally increased whereas the same patterns did not appear in special wards. Therefore, the analysis of the next chapter intentionally focuses on delineating determinants affecting estimated efficiency among Tokyo local governments since the important milestone of the introduction of accrual accounting.

CHAPTER 5. PUBLIC EFFICIENCY IN TOKYO LOCAL GOVERNMENTS: THE ROLE OF ASSET UTILIZATION AND BUDGETING

The previous chapter points out the downward trend in public efficiency scores in the 23 special wards and the upward trend in the 26 Tama cities. In this chapter, we further estimate which driving factors may potentially affect efficiency scores through the second-stage regression of bias-corrected efficiency scores on environmental variables. We used the double-bootstrapping truncated regression method, known as the Simar and Wilson (2007) model, and found a significant association between asset utilization and public efficiency. Moreover, an extensive study on two sub-regions of the Tokyo metropolis (special wards and cities separately) and stratification of assets into various categories necessarily shed light on the rational choices of asset investments or public asset reallocations for planning and budgeting.

5.1 Introduction

The advent of NPM since the 1980s has motivated many governments and public agencies to reform public sector management practices, often driven by public sector accounting reforms (Guthrie, 1998; Hood, 1995). The focus of NPM has been on improving public efficiency, effectiveness, and accountability (Hood, 1995). In fact, most public organizations generally lag behind their private counterparts regarding performance measurement and efficiency improvement. This is because public organizations are more bureaucratic, subject to less competitive pressure than their business peers, and benefit from generally reliable sources of public funding; thus, public managers have less incentive to

improve organizational performance (Boyne, 2002). However, in recent years, PSOs have been confronted with budgeting austerity in the face of financial crises and uncertain cash inflows while facing increasing demands for public services that entail considerable expenditures. Consequently, municipalities have started to come under increasing pressure to follow private management practices, improve performance management, and focus on output controls to maximize efficiency (Hood, 1995).

Public efficiency (what economists refer to as technical efficiency) is concerned with the relationship between public sector inputs and outputs. The principal inputs in the production process are capital and staff and the outputs include public goods and services (Drew et al., 2015). Notably, public service delivery is fundamentally supported by public assets controlled by the local governments. These assets are generally underutilized and poorly managed (Kaganova & Nayyar-Stone, 2000; Phelps, 2011). Nevertheless, empirical studies examining the link between efficiency and asset utilization are relatively rare. We address this gap in the extant literature by investigating the association in the context of Japanese municipalities, specifically Tokyo local governments.

There are a number of reasons why Tokyo municipalities are the ideal context for studying the association between technical efficiency and asset utilization. First, the Tokyo metropolis, administrated by the TMG, takes the lead in the economic development of Japan and local units provide the majority of large Japanese corporations and a high proportion of the population of Japan with a diverse range of essential goods and services. Second, the research findings of the previous chapter indicate an unusual phenomenon

whereby the efficiency scores in 23 special wards generally declined while those in 26 cities typically increased. We aim to examine the external factors or covariates (such as utilization of public assets) associated with the trend of efficiency estimated since 2008 to determine the source of this divergence. Third, TMG is composed of 23 special wards (designated as special local entities, in which TMG, the higher tier of government, collects tax, redistributes intergovernmental grants, and provides services on their behalf (Tran et al., 2018)), and 26 cities (classified as ordinary public entities similar to other local governments in Japan). Although a distinction of public finance and administration between the two types of local governments has been recognized, there is little comparative empirical work on this unique arrangement of local governments.

This study should, therefore, be of value to policymakers wishing to explore the public service efficiency of Tokyo local governments, particularly where there is a special interest in the implications of local governments on technical efficiency (special wards and cities). To achieve this objective, we specifically observe the response in efficiency following the introduction of accrual accounting in 2008. In addition, we examine the association between public efficiency and both cash-based budgetary income and expenditure accuracy, with respect to asset utilization, by incorporating appropriate control variables for demographic and socio-economic factors. We draw on the cash and accrual data of expenditures as an input for analysis in order to evaluate the validity of two accounting systems. In doing so, we stratify the asset utilization into two dimensions. First, in terms of cross-sectional regions, we make a comparison of the effects of accrual accounting on public efficiency between 23 special wards and 26 Tama cities. Second, in

terms of assets classification, we stratify assets into various types (e.g., living assets, educational assets, public welfare assets) to investigate the categories of assets which most strongly drive productive efficiency changes in Tokyo local governments. As a result of our analysis, we also derive some implications for the improvement of the budgetary control process for local authorities and public policymakers. To conduct this research, we collected archival data over eight years (2008–2015) and employed two-stage DEA methods to estimate public efficiency and its association with asset utilization (and some other discretionary and non-discretionary variables).

The remainder of this chapter is structured as follows. The next section introduces the research background and formulates hypotheses for municipal efficiency associated with asset utilization and budgetary control systems. The third section delineates the two-stage analysis of DEA adopted and the relevance of data employed in this study. The empirical results and findings are presented in the fourth section. We conclude with some brief remarks regarding the implications of our research for public policy.

5.2 Research background

5.2.1 Public sector accounting regime

Since the 1980s, public sector accounting has been moved forward under the NPM reforms in many advanced economies such as the UK, Australia, and New Zealand, which have transformed from cash-based to accrual-based accounting (Blöndal, 2004; Guthrie, 1998; Schick, 2007). Japan has not been an exception. Indeed, Japan's local governments adopted accrual-based accounting in the public financial reforms undertaken by the

Koizumi Administration in 2001 (Tanaka, 2009). TMG was the vanguard for introducing accrual accounting, officially adopting it for financial management from 2006 onwards (Yamamoto & Noguchi, 2013). Although accrual-based accounting was not required by law, most local governments voluntarily adopted this accounting mode by 2009 (Kobayashi et al., 2016).

The adoption of accrual accounting and reporting has been regarded as a "superior" system (Carlin, 2005) because it can generate a more comprehensive picture of activities than cash-based accounting (Kobayashi et al., 2016). It can thus facilitate more effective cost control and efficiency measurement (Guthrie, 1998). In principle, accrual accounting can identify necessary resources, such as assets controlled for the purposes of providing public services, and therefore articulate full cost information associated with the activities of public service provision (Carlin, 2005; Guthrie, 1998; Kobayashi et al., 2016; Martí, 2013). Succinctly, accrual accounting can help local authorities attain greater accountability and efficiency for stakeholders.

Accrual information generated by financial statements at the local level is expected to be employed for budgeting purposes. According to Carlin (2005), using accrual accounting and budgeting can lead to higher efficiency, better resource allocation, and improved performance. However, in practice, according to a survey of the Ministry of Internal Affairs and Communications Japan (MIC) (2012), accrual accounting information is generally employed to analyze the financial conditions of municipalities in benchmarking cross-sectional units over time (32.88%), to explain the financial conditions to citizens

(24.31%), and for the purposes of informing the legislative assembly (25.16%). Notably, the use of accrual information for budget purposes and asset management was limited to just 2.7% and 2.66%, respectively. The principal reason for the failure to apply accrual accounting information in this regard might be explained that Japanese local government budgets are still prepared on a cash basis rather than an accrual basis. While many of the Anglosphere countries (such as the UK, Australia, and New Zealand) have exercised accrual budgeting since the 1990s (Martí, 2013; Robinson, 2009; Schick, 2007), Japan's local governments have maintained traditional cash budgeting.

Cash-based budgeting differs from accrual-based budgeting in two main ways. First, cash-based budgeting recognizes transactions when cash is received or paid whereas, by contrast, the accrual method records costs and income flows when activity associated with the consumption of assets or the gaining of revenue occurs (Blöndal, 2004; Martí, 2013). Second, in addition to covering cash inflows and outflows, accrual-based budgeting also incorporates other transactions related to revenue (e.g., income tax), expenses (e.g., depreciation expenses), liabilities (e.g., civil service pensions), and assets (e.g., road infrastructure) (Martí, 2013; Robinson, 2009; Schick, 2007). Hence, cash budgeting is restricted to collecting relatively simple data for spending decisions whilst accrual budgeting is responsive to broader functional scopes of municipal activities because it measures the full cost of services (thus facilitating better decision-making, especially with respect to forward planning). Accrual-based budgeting, defined as "the specification of budgetary expenditure authorizations and revenue estimates in terms of accrual accounting measures" (Robinson, 2009, p. 4), could also be beneficial to local governments at the

microeconomic level for a number of reasons. First, the approach can bring about improved efficiency and effectiveness of expenditure controls because the true costs of inputs and outputs are reflected in the service production function, enabling local authorities to better measure the desirability of various input components (given the value of outputs) to ensure allocative efficiency (matching available resources to preferences of citizens for local public goods and services) or to improve decision-making with regards to outsourcing or internal production (which may reduce the cost of service provision and hence enhance technical efficiency) (Robinson, 2009). Second, accrual budgeting can give rise to improved decision-making associated with resource management, as a result of local authorities establishing a detailed plan for the acquisition, disposal, and maintenance of specific assets (Blöndal, 2004; Robinson, 2009). Based on the plan, policymakers can achieve an optimal allocation of available resources in a timely manner. In addition, Blöndal (2004) argues that accrual budgeting could bring about improved discipline for budget execution as a result of illuminating the long-term sustainability of public finances, catalyzing other public sector management reforms, and (most importantly) ensuring compatibility with accrual financial reporting.

5.2.2 Public efficiency

In most countries, local governments are responsible for a wide range of public services delivered to properties and people. Indeed, local governments are required to translate various incomes (e.g., tax, grants, subsidies, and proceeds from bond issues) into an array of expenditures for public services (e.g., schooling, welfare, healthcare services, and infrastructure). Moreover, in recent years, people tend to demand larger quantities and

a higher quality of public services (outputs). As a result, some local governments have suffered from deteriorating financial sustainability, especially in the context of increased austerity (with respect to inputs). Therefore, there is a need for greater emphasis on output controls under the NPM paradigm to measure local government performance, not only under pressure to "do more with less" but also to benchmark local governments (Hood, 1995). Thus, public efficiency has been brought into sharper focus in recent times.

5.2.3 Hypothesis setting

Asset utilization

To supply public services, local governments need to invest heavily in properties and infrastructure (Kaganova & Nayyar-Stone, 2000; Phelps, 2011). Indeed, assets are not only the material base for service provision, but also essential resources for local economic development (Kaganova & Kopanyi, 2014; Phelps, 2011). Therefore, managing assets in a strategic manner is an essential task of local authorities. As defined by Kaganova and Nayyar-Stone (2000), asset management deals with evaluating the performance of each asset in the context of the whole portfolio and providing a rationale for acquiring, holding, or disposing of individual assets in pursuit of the optimal composition for public service provision. Therefore, good practice in asset management can help local authorities optimize local spending. However, public assets are usually unproductively under-utilized, causing substantial opportunity costs for society, which becomes especially evident when public management practice is compared with private asset management (Kaganova & Nayyar-Stone, 2000; Tanzi & Prakash, 2000). Strategically managing and optimizing public asset

utilization can cut down on unnecessary expenditures, leading to enhanced local government efficiency. Therefore, it is reasonable to hypothesize that:

H1: *Public efficiency is positively associated with asset utilization.*

Budgetary expenditure accuracy

Since accrual budgeting can reflect full cost data and exert a measure of expenditure control, the expected result should be higher budgetary accuracy. Moreover, one of the benefits obtained from the implementation of accrual budgeting is the promotion of efficiency (Blöndal, 2004; Robinson, 2009). Thus, it can be conjectured that the higher budgetary accuracy on an accrual basis results in greater public efficiency. However, in practice, Tokyo local governments still employ cash budgeting and (more importantly) strive to ensure that existing budgets are fully spent prior to the end of the financial period[5], despite municipal efficiency being likely to decrease. Therefore, we propose the following hypothesis:

H2*: Public efficiency is negatively associated with budgetary expenditure accuracy.*

Budgetary income accuracy

Tokyo local governments gain their revenues through various kind of taxes, grants, subsidies, and proceeds from bond issues, translating into expenditures (Tran et al., 2018). In contrast to local spending, actual incomes received by these municipalities often exceed

---

[5] Investment are categorized into expenditures in the case of Tokyo local governments' cash basis budgets. Hence, if investments in a specific year are increased, then expenditure could be increased, possibly causing lower efficiency. Note that the actual expenditure must not exceed planned expenditure.

the budgeted levels. Existing empirical evidence suggests that the proportion of own-source revenue is not associated with efficiency; grants delivered to local governments with a lower revenue-raising capacity could improve their efficiency (Fogarty & Mugera, 2013). Indeed, local revenue volatility was observed to be associated with local expenditure volatility in Tokyo (Tran et al., 2018). Hence, higher incomes often translate into higher expenditures. Consequently, spending more on public services and goods provision (greater inputs) without a corresponding increase in service outputs could lead to a reduction in productive efficiency. Furthermore, contemporary budgeting on a cash basis might not provide a comprehensive picture of spending and investments in capital expenditures, so local authorities that employ strategic use of such budgets could make relatively inefficient decisions in terms of spending. In H2, we theorize that higher budgetary accuracy of expenditure has a negative influence on efficiency. Although the link between revenue and efficiency seems weakened, we posit the following hypothesis:

H3: *Public efficiency is negatively associated with budgetary income accuracy.*

5.3 Methodology

In this study, we adopt a two-stage DEA. In the first stage, we estimate the efficiency scores (results provided in Chapter 4) and, in the second, we employ the double-bootstrap truncated regression proposed by Simar and Wilson (2007) to regress DEA efficiency scores against various environmental factors. Although many studies have suggested using the Tobit (censored), Ordinary Least Square (OLS), and other non-parametric regressions, Simar and Wilson (2007) have demonstrated some drawbacks

associated with these approaches due to the potential of efficiency scores being biased and facing serial correlation and boundary problems. Hence, we have adopted truncated regression with the double-bootstrapping algorithm suggested by Simar and Wilson (2007) to alleviate the shortcomings of conventional methods.

Double-bootstrap truncated regression is considered superior to other methodologies for a number of reasons. First, the estimates obtained through this method can provide more meaningful results when compared with others, due to its relaxation of the boundary problem where technical efficient scores are mostly observed in a neighborhood of the technological frontier (Simar & Wilson, 2007). Second, Tobit and OLS regression cannot solve the dependency problems where efficiency score is a relative index (Xue & Harker, 1999), while the super-efficiency estimators have difficulty in interpretation (Fried et al., 2008). Furthermore, the double-bootstrapped procedure performs better than single bootstrapping (Simar & Wilson, 2007), where the source of data is assumed to be processed through the data-generating process (DGP). Thus, we follow the double bootstrap method proposed by Simar and Wilson (2007), in which the bias-corrected efficiency scores ($\hat{\theta}_i^*$) generated in the first stage are regressed on various environmental variables (z_i), as the following specification: $\hat{\theta}_i^* = \alpha + z_i\beta + \varepsilon_i$, $i = 1,2,...,n$, where α is a constant intercept, β is a vector of parameter coefficients, and ε_i is a statistical error term.

We selected six explanatory variables as being representative of financial, economic, and demographic determinants of our efficiency scores in the second-stage analysis. Table 7 describes the statistical summary of six independent variables. These

variables are assumed to be associated with efficiency, but not directly related to input and

output variables in the first-stage procedure (efficiency estimation). The first financial

determinant is asset utilization[6], which can be defined as the inverse value of the total

amount of public assets per population averaged over an accounting period. There are three

categories of total assets in Tokyo including total public assets (living infrastructure,

education, welfare, environmental hygiene, industry promotion, fire-fighting, and general

affairs), total investment assets, and total current assets (see Table 12 for definitions of each

category and examples). Second, budgetary accuracy regarding expenditure and revenue

suggests two additional financial determinants. The budget accuracy indicates the extent to

which actual values at the end of fiscal year deviated from the planned or budgeted values

revised during the same fiscal year. We noted that budget accuracy measured is based on

the general account, which is slightly different from the normal account as expenditure

inputs data (we calculated the rate of its difference at 5.2%). For Tokyo municipalities, a

balanced budget approach for revenue and expenditure is taken on a cash basis and the

fiscal gap is adjusted by nominal reserve funds.

　　　We also selected the mean taxable income as a proxy for the local economic

conditions of each municipality. Fogarty and Mugera (2013) argued that councils with

higher socio-economic disadvantage would have lower efficiency scores; thus, the higher

the socio-economic position, the higher the associated efficiency scores. Finally,

demographic factors include population growth and density. Population growth refers to the

---

[6] The financial data of assets were extracted from balance sheets of each local government. There were 45 municipalities publishing in general account and four municipalities using normal account. We observed that there is no significant difference between two modes.

annual growth rate of inhabitants (an annual percentage increase). Balaguer-Coll et al.

(2013) define population density as the ratio of the population divided by the total relevant

geographic area. A rationale for population density associated with efficiency is that more

widely dispersed populations increase the cost of service delivery and, hence, reduce the

efficiency score obtainable. While the first three variables are discretionary (controllable),

the remainder are non-discretionary or exogenous (beyond managerial control).

Accordingly, the bias-corrected efficiency scores ($\hat{\theta}_i^*$) were regressed on asset

utilization (ASSET_U), budget accuracy of expenditure (BUD_EXP), budget accuracy of

income (BUD_INC), taxable income (TAX_IN)), population growth (POP_GR), and

population density (POP_DEN). Therefore, our specification is defined as:

$$
\begin{aligned}
\hat{\theta}_i^* = \alpha + \beta_1 ASSET\_U_i + \beta_2 BUD\_EXP_i + \beta_3 BUD\_INC_i + \beta_4 TAX\_IN_i \\
+ \beta_5 POP\_GR_i + \beta_6 POP\_DEN_i + \varepsilon_i
\end{aligned}
$$

The bootstrap procedure was generated using 2000 bootstrap replications with 1000

repeats of the boundary. The analysis is performed in STATA 14.0 with a statistical

package developed by Badunenko and Tauchmann (2018) using algorithm 2 of Simar and

Wilson (2007).

For independent variables, we did not use instrumental variables to examine

endogeneity for two reasons. First, the efficiency score is a relative value benchmarking

among DMUs according to time series (Xue & Harker, 1999). Moreover, efficiency

estimated is unknown or unobserved so that reverse causing from efficiency to

environmental variables rarely occurs. Furthermore, efficiency scores vary if one chooses

CRS or VRS technology. Second, by employing the Simar & Wilson (2007) algorithm, the boundary problem and serial correlation are eliminated because the bootstrapping technique reduces error terms and lessens the impact of lag endogenous variables. Furthermore, we provide the correlation matrix of coefficients among independent variables prior to the regression analysis and notice that these coefficient correlations are relatively weak (Table 11 and 12 in Appendix 1).

Table 7. Definitions and descriptive data of environmental variables

Independent Variables	Definition	Mean	Standard deviation
Asset utilization	The inverse value of total amount of assets per population averaged for an accounting period.	0.633	0.28
Budgetary expenditure accuracy	The accuracy of real and planned expenditure budgeting in the general accounts (%)	0.959	0.0311
Budgetary income accuracy	The accuracy of real and planned income budgeting in the general accounts (%)	0.995	0.0326
Municipal economic situation	Taxable income proxy for the local economic condition (in thousands of yen)	548347.7	450804.9
Population growth	Annual growth rate of population (%)	0.97	1.55
Population density	Population per area (people/km^2)	10815	4846

5.4 Empirical results of double-bootstrapping truncated regression analysis

5.4.1 Results of entire Tokyo local governments

The double-bootstrapped truncated regression with algorithm 2 presented by Simar and Wilson (2007) was employed in the second-stage analysis. The bias-corrected efficiency scores of all 49 Tokyo local governments, with a 95% confidence interval, were estimated through two models to cross-check robustness and benchmark the two accounting modes at the local level (see Table 8). Model (1) presents the regression results associated with the efficiency scores when the total expenditures and labor cost are measured using cash-based accounting, while Model (2) uses accrual-based total and labor expenses.

In Model (1), asset utilization has no statistically significant association with bias-corrected efficiency, suggesting that the hypothesis 1 cannot be supported. It posits that cash-based public asset value is unlikely to support improvement of public efficiency. Additionally, budgetary expenditure and revenue accuracy also have no statistically significant impact on efficiency scores. In other words, hypotheses 2 and 3 cannot be supported, implying that the consistency of actual amounts from budget projections using cash-based accounting information cannot significantly affect municipal efficiency levels.

In contrast, population growth is negatively linked with public efficiency at the 1% level of significance, with the coefficient of -0.0385. Productive efficiency could be reduced by 3.8% for every 1% increase in the population growth rate, *ceteris paribus*. Similar results were found in empirical studies by Balaguer-Coll et al. (2013), arguing that a rapid growth rate of population in the short-term imposes higher spending requirements

on local authorities (e.g., on industrial and construction activities), leading to a deterioration of their efficiency.

Finally, population density was also found to have a negative effect on public efficiency at the 1% significant level with the correlation coefficient of -0.0806. This consequently indicates that an increase in one unit of population density could reduce 8.06% of efficiency scores, holding all other factors constant. This is consistent with the findings of Balaguer-Coll et al. (2013) in Spanish local governments, Da Cruz and Marques (2014) in Portugal municipalities, Lampe et al. (2015) in the North Rhine-Westphalia state of Germany, and Forgaty and Mugera (2013) in Western Australia municipalities that municipalities with more concentrated populations can cause growing demand for proportional amounts of public service provision; consequently, cost of services might be high, resulting in lower efficiency. Given the very dense Tokyo local government population, particularly in special wards, further increases in such density might lead to inefficiency of public service delivery as a result of diseconomies of density.

In Model (2)[7], when accrual data for inputs of total expenditures and total labor cost is employed for efficiency estimation, we find a statistically significant and positive association between asset utilization and efficiency scores with the coefficient of 0.0124 at 5% significant level. It could be concluded that asset utilization is considered an important driving factor for public efficiency. As compared with the cash basis model in Model (1), the accrual accounting system provides a better reflection of strategic public asset

---

[7] In the data analysis, there are four local governments reported in general account against 45 local governments reported in normal account. Due to data availability, we pooled them as inputs data because there is no significant difference between general account and normal account.

management. The correlation coefficient of asset utilization under the accrual-based regime is higher than under the cash-based regime (almost double in magnitude). Our results demonstrate that it is desirable to use an accrual-based performance indicator rather than a cash-based one when asset utilization is emphasized as an operational driving factor to improve public efficiency.

Table 8. Bias-corrected efficiency regression results.

VARIABLES	(1) Cash basis	(2) Accrual basis
Asset utilization	0.00768	0.0124**
	(0.00519)	(0.00535)
Expenditure accuracy	0.00584	0.00722
	(0.00747)	(0.00757)
Income accuracy	0.000631	-0.00538
	(0.00744)	(0.00767)
Taxable income	0.00759	0.0197***
	(0.00496)	(0.00513)
Population growth	-0.0385***	-0.0324***
	(0.00497)	(0.00509)
Population density	-0.0806***	-0.0744***
	(0.00538)	(0.00543)
sigma	0.0698***	0.0714***
	(0.00349)	(0.00353)
Constant	0.827***	0.832***
	(0.00483)	(0.00486)
Observations	287	286

Notes: ***, **, * indicate significance at 1%, 5%, and 10% level. Standard errors in parentheses.

Like Model (1), budgetary control of expenditure and income also does not significantly affect municipal efficiency. Noting that Tokyo local governments still practice cash-based budgeting, which results in less information regarding the true cost and hence less control of public management, cash budgeting does not appear to suffice for efficiency

improvement. Robinson (2016, p. 38) states that "to achieve the desired improvements in capital asset budgeting and management, it would be necessary to implement accrual budgeting rather than merely accrual accounting". Furthermore, accrual budgeting seems to be beneficial for effective and efficient utilization of input expenditures (better choices for expenditure priorities and better decisions for production) (Robinson, 2009). Therefore, the introduction of accrual budgeting elements might be necessary for better asset management and promotion of efficiency in Tokyo municipalities, being consistent with similar results found for German local governments (Lampe et al., 2015). Accrual budgeting could undoubtedly be matched with the accrual financial statement and reports, which have received special attention, and then adoption, from some advanced countries such as Australia, New Zealand, the UK, and Germany under the spirit of the NPM.

Next, there is a statistically significant and positive association between productive efficiency and the taxable income per capita at 1% level. This implies that the better the municipality's economic status, the higher the level of public efficiency. This finding is in contrast with earlier studies such as Fogarty and Mugera (2013) in Western Australia municipalities and Doumpos and Cohen (2014) in Greek local governments, where there was no significant relationship between socio-economic conditions and efficiency scores. Likewise, this statistical result contradicts the finding of Storto (2016) and Balaguer-Coll et al. (2013) that higher efficiency is negatively associated with economic conditions.

Lastly, the results for population growth and density in Model (2) were consistent with Model (1). These environmental factors are negatively associated with municipal performance, proving that irrespective of cash or accrual basis, the exogenous variables

(uncontrollable variables) concerning population characteristics have the potential to adversely affect the overall public efficiency.

In sum, drawing on our findings, we suggest that the accrual accounting and budgeting system could be a better management tool for the financial activities of reporting and planning, especially public asset management, although some continuous empirical studies would be conducted to offer more persuasive evidence for this deduction. Nevertheless, due to the differential fiscal mechanism between special wards and Tama cities, a further analysis involving stratification of two sub-groups can provide practitioners and policymakers alike with more detailed explanations.

5.4.2 Stratification of Special wards and Tama cities

Research on asset utilization can be further extended by stratification of the two main groups, special wards and Tama cities, in terms of cash and accrual basis. Figure 13 indicates that a considerable amount of public money has flowed into special wards in the form of investments in assets up to five or six times larger than the amount invested in Tama cities. We seek to explain the difference between the two peer groups in terms of public asset management and budgetary control. Our regression was designed to test to what extent the adoption of accrual accounting affects public efficiency boosting in both regions as compared with the effects of cash accounting. Regression analysis involving special wards is presented in Model 3 (cash basis) and Model 4 (accrual basis), whilst Tama cities are presented in Model 5 (cash basis) and Model 6 (accrual basis) (see Table 9).
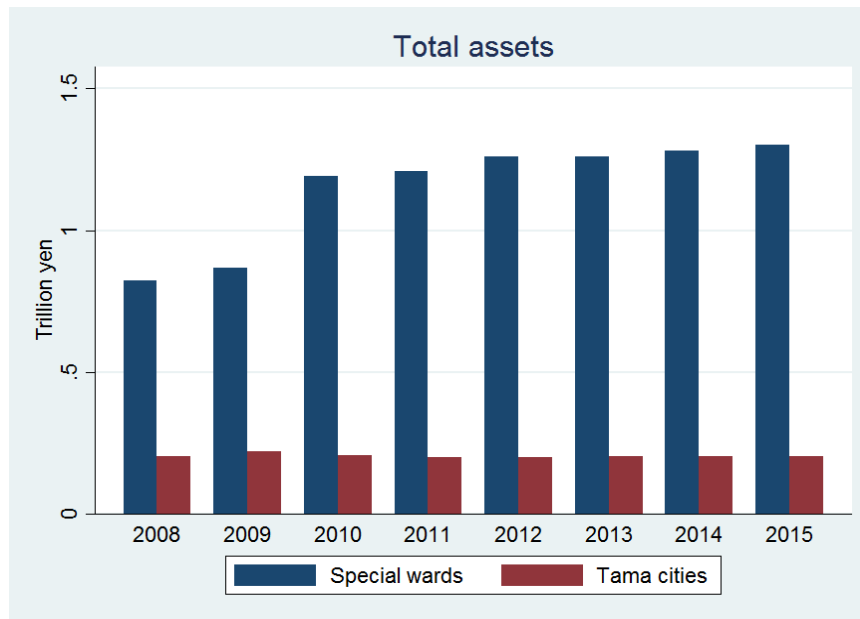
Figure 13. Total assets of Special wards and Tama cities.

Prior to practical interpretation it is essential to get insights into some individual characteristics of special wards and Tama cities with respect to local revenue earning and spending on public services. First, special wards have made an effort to save public money from TMG distributional grants and, thereby, accumulated a large proportion of revenue to cushion unforeseeable events such as the 2008 GFC, whereas Tama cities have collected the income mainly by themselves through taxes, fees, and charges, whilst the amount of savings is relatively limited (See section 3.4.2 in Chapter 3). Second, aside from the mandatory public services and tasks legally assigned by higher tiers of governments (TMG as well as national government), both special wards and cities must be responsive to their constituencies' public service provision. However, there is a difference – while cities work out the service provision *per se*, special wards share some public affairs, such as fire protection and water sewerage, with TMG (Ohsugi, 2011). Since cities must cover a wider

range of public services delivery than special wards, public authorities in Tama cities intentionally prioritize public services or programs that cause high efficiency. Subsequently, low-efficiency activities must be assessed by public managers to determine whether they should be maintained or terminated, where possible, to preserve public money. Accordingly, there is a substantive disparity in own-revenue raising and expenditure spending associated with municipal efficiency across two jurisdictions.

Table 9. Bias-corrected efficiency estimates for Special wards and Tama cities

	Special wards		Tama cities	
	(3)	(4)	(5)	(6)
VARIABLES	Cash basis	Accrual basis	Cash basis	Accrual basis
Asset utilization	0.0102***	0.00983***	-0.0164***	-0.0194***
	(0.00353)	(0.00342)	(0.00615)	(0.00630)
Expenditure accuracy	-0.0246***	-0.0176***	0.0285***	0.0336***
	(0.00558)	(0.00542)	(0.00943)	(0.00988)
Income accuracy	0.0311***	0.0217***	-0.0262***	-0.0149
	(0.00512)	(0.00498)	(0.00964)	(0.00989)
Taxable income	-0.0283***	-0.0175***	-0.123***	-0.469***
	(0.00332)	(0.00323)	(0.0115)	(0.0121)
Population growth	-0.000967	-0.00431	-0.00968	0.000440
	(0.00359)	(0.00355)	(0.00719)	(0.00768)
Population density	-0.00744	-0.0125**	-0.0732***	0.0209***
	(0.00571)	(0.00567)	(0.00879)	(0.00800)
sigma	0.0314***	0.0324***	0.0524***	0.0520***
	(0.00243)	(0.00241)	(0.00338)	(0.00330)
Constant	0.956***	0.949***	0.751***	0.552***
	(0.00861)	(0.00850)	(0.00998)	(0.00986)
Observations	118	118	169	168

Notes: ***, **, * indicate significance at 1%, 5% and 10% level. Standard errors in parentheses.

With regard to the first discretionary factor – asset utilization – for special wards' specifications expressed in Models (3) and (4), the repressors of interest regarding asset utilization are statistically significant and positively associated with bias-corrected efficiency at the 1% level, with correlation coefficients of 0.0102 and 0.00983, respectively. The positive sign – which we would, *a priori*, expect – suggests that growing investments of public assets in special wards lead to a decline in asset utilization and a downward trend of technical efficiency (Figure 14a). Asset investments exceed the need; hence, are far from the optimal scale (Figure 14b). The sign of correlation coefficient of asset utilization in Models (5) and (6) of Tama cities, however, is in the opposite direction. In these models, asset utilization is negatively associated with municipal efficiency at the 1% significance level, with coefficients of -0.0164 and -0.0194, respectively. This finding implies that an increasing amount of asset investment in Tama cities causes lower asset utilization levels, correspondingly raising public efficiency (Figure 14a). A lack of assets invested against the actual need in Tama cities illustrates the fact that an increasing amount of assets can generate higher efficiency (Figure 14b). While an increase of assets in special wards generated inefficiency, an increase in assets in Tama cities resulted in higher efficiency, *ceteris paribus*. Notably, asset utilization has a different effect in two groups on municipal efficiency, leading to a downward trend in special wards but an upward trend in Tama cities (presented in Chapter 4), which also explains research question 4 in this thesis. Finally, it can be implied that the assets in special wards should be cut down but those in Tama cities enlarged to pursue higher efficiency. A further question posed is what types of assets should be prescribed in policies to promote public efficiency. This problem will be solved in the next section.
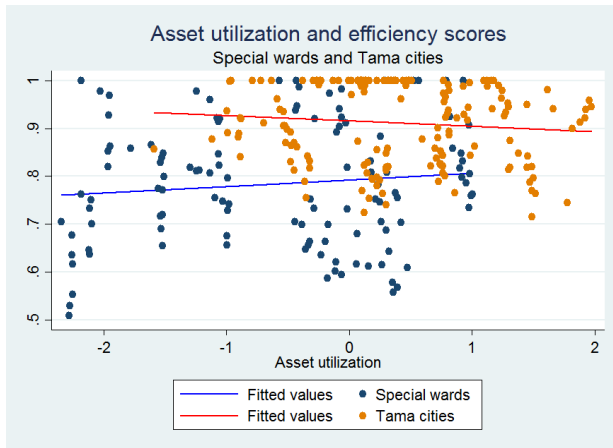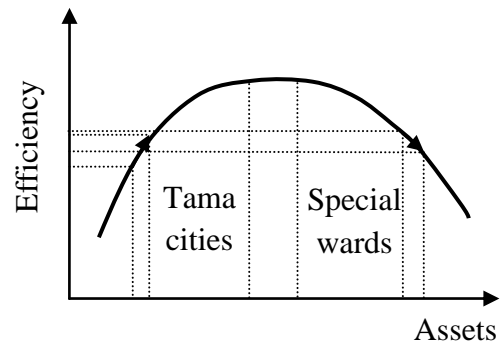
Figure 14. Asset utilization and efficiency scores.

The second discretionary variable of interest is budgetary expenditure accuracy (percentage difference between the amount of actual expenditure and budgeted expenditure for each fiscal year). For Models (3) and (4), the regression indicates that there is a negative association at the 1% significance level. This suggests that the more accurate the budgeting control, the less efficiency can be attained for special wards. The rationale is that the amount of budgeted expenditure principally set out at the beginning of the fiscal year is higher than the amount of actual demand because budget holders believe that annual revenues of special wards are relatively abundant (see Chapter 3) and able to cover an augmented level of spending. Hence, the budgeted expenditures are overestimated, resulting in budgetary slacks. In the presence of larger budgeted expenditures, public managers try to increase spending on public services to reduce the gap between the budgeted and actual expenditure, resulting in increases in budgetary accuracy. In fact, an increase in such expenditure can decrease public efficiency. Therefore, technical efficiency for special wards is reduced in our sample when budgetary accuracy of expenditure is high.

125

It is possible that cash budgeting does not perform well in special wards in driving public efficiency.

In contrast to special wards, the association between budgetary expenditure accuracy and public efficiency in Tama cities (Models 5 and 6) is significantly positive at the 1% level, which implies that the higher the budgetary expenditure accuracy, the higher the level of local efficiency. This justification proves that public managers of Tama cities are aware of the scarcity of their pecuniary resources as opposed to their peers in special wards, so their planning activities are better. Moreover, as mentioned, Tama cities have the responsibility not only of obligatory tasks assigned by higher tiers of governments, irrespective of efficiency, but also their own operations. Therefore, public managers have opted for higher-efficiency rather than lower-efficiency operations. Accordingly, Tama cities generally attain high efficiency scores. Another important outcome to be observed is that the coefficient correlation of budget expenditure accuracy in accrual accounting (Model 6) (0.0336) is higher than that of cash accounting (Model 5) (0.0285), suggesting that pecuniary planning on an accrual basis could foster better results in terms of public efficiency. It should be noted that there is a countersign in these variables between special wards and Tama cities, which may explain why Tama cities have a preference for adopting the accrual basis, but preference for the cash or accrual basis is unclear for special wards.

We provide further evidence on the efficiency change of Tama cities against those of special wards in Table 10 by using the Malmquist Total Factor Productivity Index. Prior literature maintains that if a DMU attaining efficiency and technology change are higher than 1, it has proved efficiency increase (Fried et al., 2008). In fact, Table 9 shows most of

126

DMU in Tama cities exceed to 1 while it does not in special wards. It turns out that the efficiency change of Tama cities was increased in accordance with the higher budgetary accuracy over the 2008-2015 period.

Table 10. Malmquist index, efficiency change, and technology change between Special wards and Tama cities

Year	Malm.wards	Malm.cities	Eff.wards	Eff.cities	Tech.wards	Tech.cities
2008-2009	1.008	1.069	0.979	1.047	1.029	1.024
2009-2010	0.965	1.016	1.026	1.010	0.941	1.0001
2010-2011	0.959	0.945	0.985	1.001	0.973	0.943
2011-2012	1.062	0.993	0.986	0.991	1.077	1.001
2012-2013	0.912	0.991	0.973	1.015	0.936	0.976
2013-2014	0.940	0.983	0.985	0.982	0.955	1.001
2014-2015	0.974	0.967	0.983	1.004	0.991	0.962

The third discretionary variable we are concerned with is budgetary income accuracy, defined as the percentage deviation in the amount of actual and budgeted income. Interestingly, the budget income accuracy is statistically significant at 1% with a positive sign in special wards (Models 3 and 4), while the same association is also statistically significant at 1% with a negative sign in Tama cities (Models 5 and 6). In special wards, the higher the level of budget income accuracy, or the larger the amount of incomes against planned income, the higher the public efficiency scores obtained. This is at odds with most of the extant literature on public revenue and efficiency; however, as we mentioned in Chapter 3, the revenues of the 23 special wards consistently increased (even in the aftermath of the 2008 GFC), as attributed to increases in the amount of intergovernmental grants received from TMG. Despite this, public managers in special wards tend to preserve grants rather than spend them on public provision. Thus, this saving behavior may have

prevented a decline in efficiency. The "flypaper effect" (money sticks where it hits) widely occurs in many local governments across the world; prior scholarship maintains that reliance on grants could lead to low efficiency in Greece (Doumpos & Cohen, 2014), Belgium (De Borger & Kerstens, 1996), and Spain (Balaguer-Coll et al., 2007), but this phenomenon may be not appropriate in special wards circumstances nevertheless.

However, what happened in Tama cities is strikingly different from special wards. The results indicate that the lower the level of actual revenues collected, as compared with budgeted revenues, the higher the level of local efficiency. It can be explained that less affluent municipalities (principally comprised of Tama cities) are inclined to manage and operate their service provision more efficiently because public managers are aware of the scarcity of pecuniary resources. This is supported by prior empirical studies that also indicate that "higher efficiency scores are associated with less developed contexts" in Italian major municipalities (Sorto, 2016, p. 59). Furthermore, public managers in municipalities that rely on modest revenue growth tend to make careful choices in selecting the most efficient service provisions and dropping the less-efficient ones, hence preventing a deterioration of efficiency. If this is the case, it would be prudent to make these deliberate choices by using a cost-benefit analysis of the cost paid for less efficient tasks (but these tasks would bring in benefits for constituents in the near future) and savings of public money (due to cutting down on these tasks).

Regarding the non-discretionary variable of the taxable income that proxies for the municipal economic conditions, a statistically significant and negative association was found with economic efficiency at the 1% level for all special wards (Models 3 and 4) and

Tama cities (Models 5 and 6) was found. This finding was also supported by earlier

scholarly works such as Balaguer-Coll et al. (2013) in Spanish municipalities, Balaguer-

Coll et al. (2018) in Valencia Spanish jurisdictions, and De Borger and Kerstens (1996) in

Belgium, with the critical argument that "at the local level, higher incomes increase the

fiscal capacity of municipalities and may foster featherbedding of politicians and public

managers, thereby increasing the scope for inefficient operation" (Borger & Kerstens 1996,

p. 162). Particularly, it could be plausibly conjectured that special wards with wealthy

incomes pay less attention to efficiency. For the remaining determinants, population density

and growth, the interpretation should be akin to what was presented in Models (1) and (2).

To sum up, the empirical results of the bootstrapped truncated regression associated

across two jurisdictional areas provided mixed results. Most importantly, public assets play

an essential role in both regions and are necessarily expected to invest more in the pursuit

of efficiency gains, although it is clear that both special wards and Tama cities have spent a

sustainable amount of public money on various assets. However, the research question of

what types of assets should be further invested is discussed in the next section. With

regards to budgetary control, our findings indicate that it is still unclear whether cash or

accrual basis can have beneficial consequences of efficiency gain for the public sector in

special wards, but accrual basis adoption is a better choice for Tama cities.

5.4.3 Stratification of assets into various components.

For a supplementary analysis of the results obtained above, we delve into the

decomposition of various assets for all the Tokyo local governments, separating special

wards and Tama cities, in both cash and accrual accounting. In Tokyo local governments,

public assets can be classified into seven different types (living, education, welfare, environmental and hygiene, industry, fire-fighting, and general affairs assets) corresponding to physical and services assets (describing assets in Table 14). We used monetary values derived from financial balance sheets that proxy for magnitude of services delivered to citizens. Figure 15 shows that two main types of public assets – living and education – account for a majority of total assets (more than 80% across all local governments, 89.4% for special wards, and 83% for Tama cities), asserting the importance of these two types in policy-making. The remainder of assets (welfare, environmental and hygiene, industry, fire protection, and general affairs) make up only a limited proportion. Figure 15 shows the comparative amount of each asset in special wards and Tama cities. Although to differing magnitudes in both regions, special wards are more likely to spend more public money on assets than Tama cities.
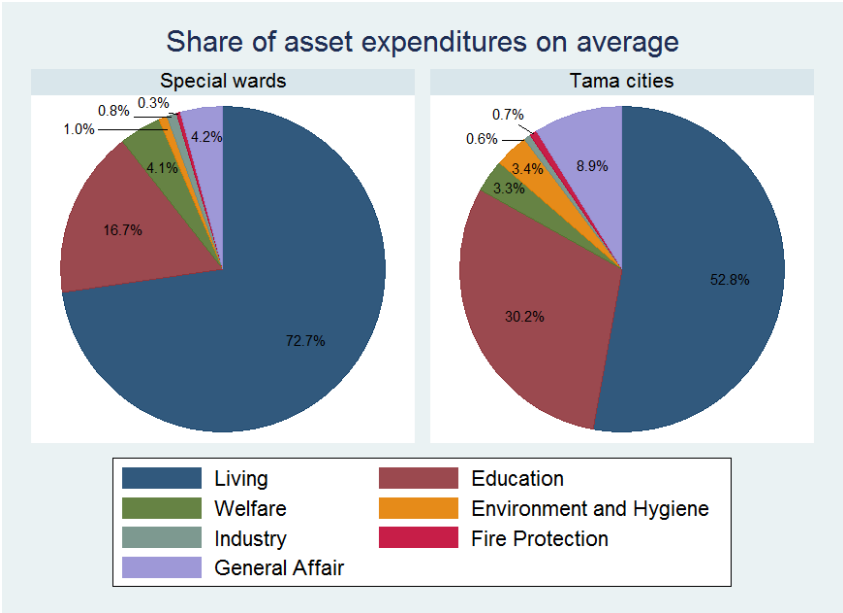
Figure 15. Shares of assets in Special wards and Tama cities.

In Table 11, for all the 49 Tokyo local governments on a cash basis in Model (7), there is a significant positive association between environmental hygiene, industry promotion, and general affairs asset variables and efficiency scores, while the statistically significant negative sign is observed in the association between educational and fire protection variables with public efficiency. Similarly, on an accrual basis in Model (8), a statistically significant association between living, environmental, and industry assets and efficiency is indicated, whereas the same pattern has been shown for education and fire protection (negative sign) as in Model (7). Thus, it can be suggested that Tokyo local governments, in general, should exert effort to invest more in education and fire-fighting and necessarily reduce the amount of living, environment, industry, and general affairs assets in the pursuit of efficiency gains.

When focusing on special wards' specifications of cash and accrual basis in Models (9) and (10) because of the limited number of samples, we alternatively adopt algorithm 1 at the second stage with the VRS technology. In Model (9), we found a statistical negative association between living and fire-fighting assets and public efficiencies and a positive association between education and industry assets and public efficiency. The empirical evidence of living assets suggests that promoting strategic investment in living assets is necessary in special wards to boost their efficiency, although we note this has already occurred over 2008–2015 (See Figure 16). Despite increasing investment in life infrastructure assets in special wards, these assets might not satisfactorily live up to the actual demands of citizens because of the rapid population growth rate in most condensed population areas in the past decades. In terms of education assets, the results suggest that it

is imperative to reduce their amount. However, due to the expansion of population size and, correspondingly, increasing numbers of students, public managers and policymakers alike should consider a public policy of transfer in funding from physical education assets (e.g., facilities) to operational expenses (e.g., labor costs for teachers and instructors, administrative educational works) where these expenses are seemingly insufficient (because of increase in numbers of students). Finally, the results suggest that industry promotion assets and general affairs assets should be reduced with respect to promoting efficiency gains. It is apparent that the magnitude of the coefficient correlation regarding the industry variable is relatively large compared to these other effects (0.251 and 0.262, respectively in Models (9) and (10)), suggesting that local authorities should place more importance on reduction of physical industry assets because of their investment expansion over the past decades.
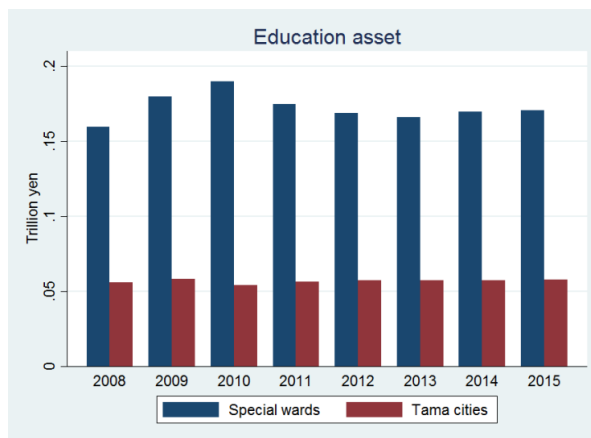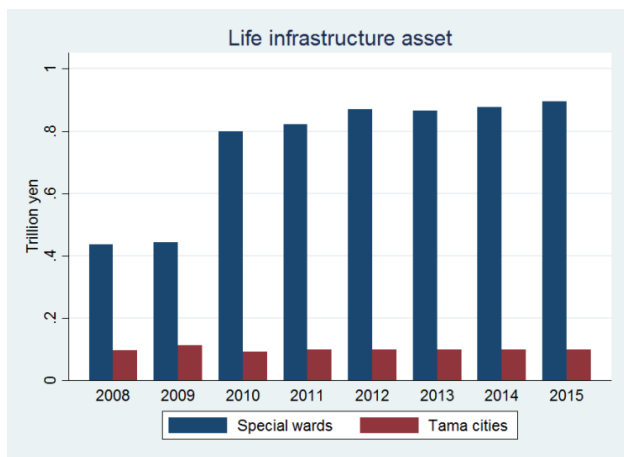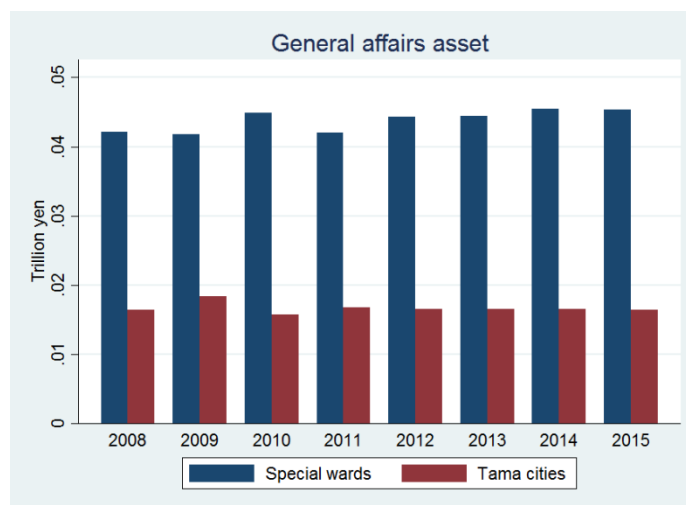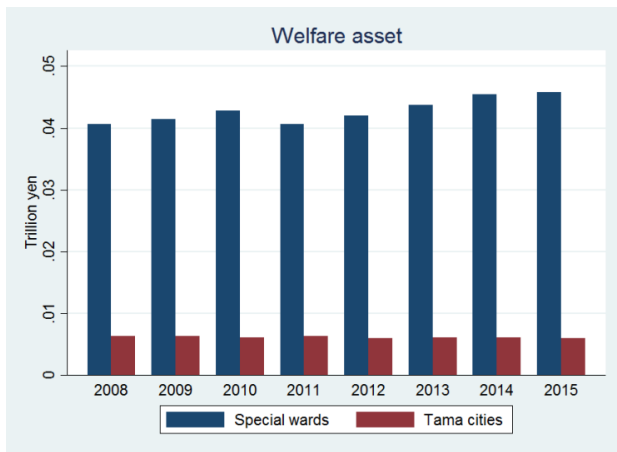
Figure 16. Various types of assets in Special wards and Tama cities.
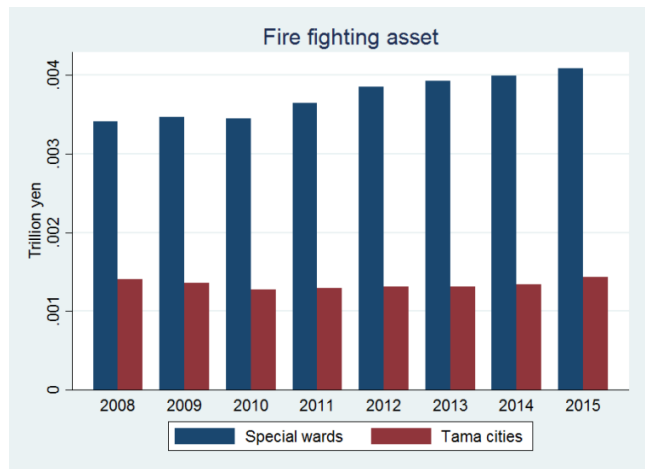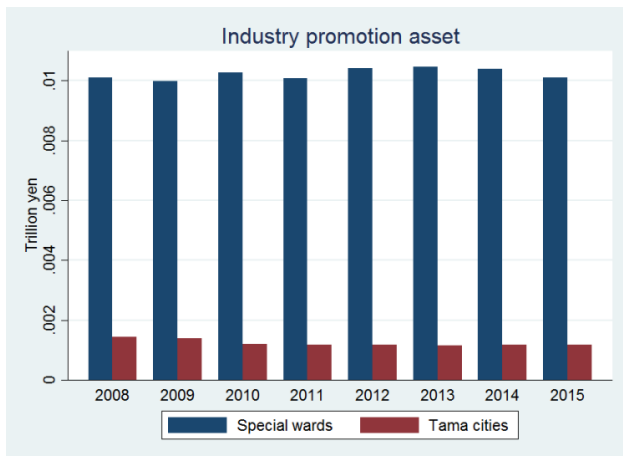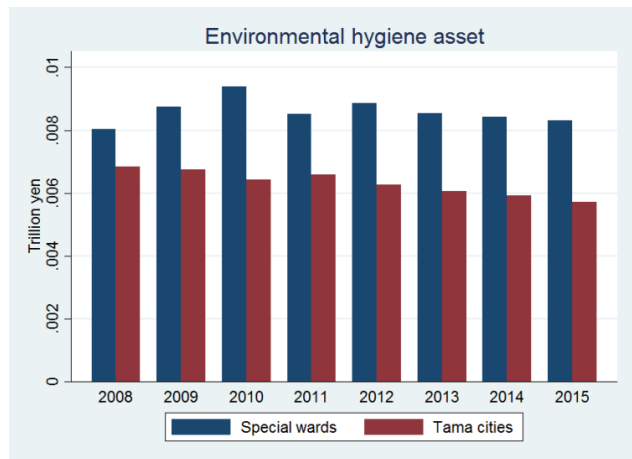
Table 11. Stratification of various type of assets

VARIABLES	Tokyo (alg2)		Special wards (alg1)		Tama cities (alg2)	
	(7)	(8)	(9)	(10)	(11)	(12)
	Cash basis	Accrual basis	Cash basis	Accrual basis	Cash basis	Accrual basis
Living	0.00434	0.0124**	-0.0398***	-0.0233**	0.0223*	0.0691***
	(0.00575)	(0.00596)	(0.00848)	(0.0110)	(0.0125)	(0.00996)
Education	-0.0245***	-0.0348***	0.0499**	0.0274	-0.0556***	-0.0593***
	(0.00721)	(0.00732)	(0.0210)	(0.0268)	(0.00843)	(0.00724)
Welfare	0.00467	0.00906	0.0718	0.0387	0.0178*	-0.00785
	(0.00777)	(0.00767)	(0.0521)	(0.0663)	(0.0107)	(0.00900)
Environmental	0.0614***	0.0626***	-0.00141	0.0147	0.0535***	0.0321***
	(0.00714)	(0.00732)	(0.0139)	(0.0167)	(0.00885)	(0.00602)
Industry	0.0228***	0.0287***	0.251***	0.262***	0.0172***	0.0118**
	(0.00704)	(0.00722)	(0.0508)	(0.0649)	(0.00531)	(0.00572)
Fire fighting	-0.0327***	-0.0256***	-0.0126*	-0.0146	0.000605	-0.0340***
	(0.00540)	(0.00560)	(0.00715)	(0.00959)	(0.00846)	(0.00745)
General affairs	0.0113**	0.00647	0.0991***	0.0600	-0.00599	-0.00204
	(0.00508)	(0.00520)	(0.0352)	(0.0458)	(0.00432)	(0.00365)
Expenditure accuracy	-0.00321	-0.00227	-0.0151	-0.0173	0.0366***	0.0332***
	(0.00741)	(0.00770)	(0.0132)	(0.0160)	(0.00977)	(0.00824)
Income accuracy	0.0141*	0.00889	0.0231**	0.00121	-0.0272***	-0.0127
	(0.00781)	(0.00821)	(0.0106)	(0.0133)	(0.0106)	(0.00903)
Taxable income	-0.0338***	-0.0376***	-0.0963***	-0.0520**	-0.106***	-0.297***
	(0.00936)	(0.00986)	(0.0181)	(0.0230)	(0.0194)	(0.0158)
Population growth	-0.0169***	-0.0163***	-0.0416***	-0.0313***	-0.00486	0.000273
	(0.00550)	(0.00572)	(0.00956)	(0.0119)	(0.00830)	(0.00639)
Population density	-0.0784***	-0.0664***	-0.143***	-0.0889***	-0.0585***	0.00331
	(0.00748)	(0.00793)	(0.0182)	(0.0236)	(0.0159)	(0.0125)
Sigma	0.0646***	0.0661***	0.0585***	0.0733***	0.0489***	0.0405***
	(0.00330)	(0.00339)	(0.00487)	(0.00625)	(0.00325)	(0.00253)
Constant	0.829***	0.826***	1.189***	1.096***	0.802***	0.681***
	(0.00518)	(0.00501)	(0.0463)	(0.0572)	(0.0140)	(0.0109)
Observations	255	254	79	79	164	163

Notes: ***, **, * indicate significance at 1%, 5% and 10% level. Standard errors in parentheses.

For Tama cities in Models (11) and (12), compared to special wards, we found positive associations between living, environmental and hygiene, and industry assets and public efficiency, whilst education assets were significantly negatively associated with economic efficiency scores. These findings suggest an emphasized practical necessity to cut down on living and environmental and hygiene assets and invest more in educational assets in the case of Tama cities. Among discretionary variables, the coefficient correlations of living assets (0.0223 in Model (11) and 0.0691 in Model (12)) and education assets (-0.0556 in Model (11) and -0.0593 in Model (12)) are much larger than others, so these two factors should be taken into particular consideration. For educational factors, on the one hand, local authorities should increase the amount of assets regarding educational activities; on the other, a shift from operational expenses toward physical asset investments might be a feasible scheme to increase municipal efficiency. For the life infrastructure, Tama cities should reduce these living assets in contrast to the situation in the special wards.

Special wards need additional investment of life infrastructure assets, although a significant budget had financed these categories in the past decade, unlike in Tama cities. We attempt to address this issue by examining supplementary evidence on life assets (e.g., roads, bridges, public housing). First, owing to the difference in geographic features and population density between special wards and Tama cities, the purpose of living assets is dissimilar. For example, while special wards particularly invest in coast conservations and seaports, Tama cities do not. Furthermore, concerning both quantity and quality of public constructions, special wards focus more on building high-grade asphalt roads and blocks

135

and maintenance of these roads, whereas Tama cities primarily use concrete technology on road constructions. Special wards prefer assets supporting business-like activities, whereas Tama cities are in favor of life assets targeted towards constituents (i.e. human residents). Moreover, although the number of bridges in Tama cities is double that in special wards, the length and width is far less compared to those in special wards (See Table 15 in Appendix 1).

Second, infrastructure assets in special wards have been built to last for a long time (around five or six decades), so there is a financial demand to cope with renewing and maintaining these living assets, while most habitat communities in Tama cities might be fairly novel (roughly two or three decades). It is, therefore, reasonable to expect higher investments in special wards infrastructure. Third, the growth rate of populations in special wards exceeds the growth rate of living assets investment. The living assets in special wards are thus less responsive to population expansion, contrary to Tama cities where the growth in living assets exceeds the population growth (See Table 10 and Figure 17). Therefore, it is suggested that special wards should necessarily spend more public money on living assets to satisfy their community needs while Tama cities should reduce living infrastructure to pursue efficiency improvement.

All in all, our empirical evidence suggests that special wards should increase investment of public money on life infrastructure assets and cut down investments on education assets (potentially transferring spending on educational assets towards administrative spending). However, Tama cities should decrease investment in living assets

(similar finding to Balaguer-Coll et al., (2013)) and increase investment in educational assets (potentially shifting operational expenses towards acquisition of physical education assets), which is consistent with Storto's extant empirical studies (2016). Overall, both jurisdictions should streamline various assets to enhance public efficiency.

5.5 Concluding remarks

This study sets out to critically examine public efficiency in Tokyo municipalities over 2001–2015 as well as the environmental determinants associated with public efficiency following the accrual accounting reforms implemented at the local government level in 2008. We recognize a substantial decline in efficiency in special wards and an opposite trend for Tama cities, so we aimed to extend our knowledge of what environmental factors have driven these trends by employing the truncated regression with Simar and Wilson's double-bootstrapping approach (2007). The findings from this study make several contributions to the current literature, particularly with respect to public policy implications.

First, our empirical analyses suggest that the utilization of assets has a positive association with municipal efficiency (providing support for H1) for an accrual basis; however, this is not yet conclusive, in the absence of further empirical testing. This implies that Tokyo local governors should properly consider ameliorating the level of asset utilization (by not only reducing but also reallocating the levels of assets) if they wish to enhance public efficiency. In doing so, the more public assets are used strategically, the higher the levels of public efficiency that might be obtained. However, in a contemporary

setting, Tokyo local governments have not yet applied accrual budgeting – by which local authorities can set budgetary control targets in a more comprehensive manner that captures asset utilization – but have instead maintained a cash-based budgeting system (which does not significantly affect the efficiency enhancement). Thus, there is a mismatch between accrual accounting and cash budgeting with regards to asset management and, therefore, good reason for Tokyo local governors to rationally consider incorporating elements of accrual budgeting in practice. If this is the case, accrual budgeting could not only support accrual accounting systems already in place but also support the current performance budgeting systems in Tokyo local governments.

Second, this inference is additionally supported by the evidence that there is no statistically significant association between cash-based budgetary expenditure and income control and public efficiency; similarly, it cannot be regarded as conclusive. In other words, it argues that the current situation of cash-based budgeting is not adequate for strategic planning towards the improvement of public efficiency. Cash-based budget tightening appears to be merely used for financial stability rather than efficiency improvement. Consequently, it may be reasonable to replace the cash budget with the accrual budget for Tokyo local governments to improve efficiency.

Third, with respect to the separation across the two jurisdictions of Tokyo metropolis, we draw plausible deductions on the use of assets, particularly that special wards, in general, should streamline their public assets because of escalated capital expenditures during the past decades, while Tama cities needs to be enlarged nevertheless.

Our further empirical analysis, involving stratification of various assets, cannot provide definitive conclusions without additional supplementary research but the results should provide helpful insights as potential policy prescriptions in relation to the reallocation of public assets. Special wards should ideally invest more in living assets and reduce educational assets, whereas Tama cities' local authorities should consider the opposite (i.e., reduction in living assets and augmentation in educational assets). Our study presents some important policy recommendations associated with public assets and budgeting regimes for local administrators and policymakers alike. Thus, it should necessarily be strengthened with more compelling evidence from academics and practitioners alike.

Appendix 1.

Table 12. Correlation matrix of coefficients in cash basis

	Asset utilization	Expenditure accuracy	Income accuracy	Taxable income	Population growth	Population density
Asset utilization	1.0000					
Expenditure accuracy	-0.0802	1.0000				
Income accuracy	0.0582	-0.7887	1.0000			
Taxable income	0.2464	0.1147	-0.2739	1.0000		
Population growth	0.2908	0.3030	-0.0981	-0.0526	1.0000	
Population density	0.0293	- 0.0562	0.1258	-0.3528	0.0780	1.0000

Table 13. Correlation matrix of coefficients in accrual basis

	Asset utilization	Expenditure accuracy	Income accuracy	Taxable income	Population growth	Population density
Asset utilization	1.0000					
Expenditure accuracy	-0.0981	1.0000				
Income accuracy	0.0361	0.8149	1.0000			
Taxable income	0.1425	0.1678	-0.2926	1.0000		
Population growth	0.2669	0.3268	-0.1512	-0.0574	1.0000	
Population density	0.0661	-0.2403	0.2975	-0.3804	- 0.0181	1.0000

Table 14. Categories of public assets in Tokyo local governments

	Public assets	*Examples*	*Description*
1	Lifestyle infrastructure and land conservation	Roads, bridges, parks, public housing, etc.	Assets classified as civil engineering management expenses, road bridging expenses, river costs, port expenses, urban planning fees, and housing expenses.
2	Education	Kindergarten, elementary, and junior high school, high school, library, public hall, etc.	Assets classified as educational administrative expenses, elementary school expenses, junior high school expenses, high school expenses, kindergartens expenses, social education expenses, health and physical education expenses, special school expenses, and university fees.
3	Welfare	Nursery school, children's hall, etc.	Assets classified as social welfare expenses, elderly welfare expenses, child welfare expenses, welfare expenses, and disaster relief expenses
4	Environment hygiene	The health center, garbage disposal site, garbage collection vehicle, etc.	Assets to be classified as health and sanitation expenses, tuberculosis control expenses, public health care expenses.
5	Industry promotion	Work welfare hall, etc.	Assets classified as agricultural expenses, livestock industry expenses, agricultural land costs, forestry expenses, fishery industry expenses, unemployment countermeasure expenses, labor various expenses, and commerce and construction expenses.
6	Fire	The fire department, fire-fighting tank, fire-fighting vehicles, etc.	Assets classified as fire-fighting expenses
7	General affairs	City hall government building, community center, etc.	Administrative and administrative expenses, tax collection expenses, family register, resident registration card fee, election expenses, the statistical survey fee, audit committee fee, and other assets.

Table 15. Differential living assets between Special wards and Tama cities over 2008-2016

		Special wards	*Tama cities*
Bridges	Number (unit)	163	338
	Length (km)	14,251	9,660
	Area (km^2)	275,291	110,415
Improvement of roads	Length (km)	10,176	6,878
	Area (km^2)	95,432	60,769
Un-improvement of roads	Length (km)	1,696	3,360
	Area (km^2)	6,146	9,143
Concrete roads	Length (km)	80,517	129,151
	Area (km^2)	3,112,213	3,370,007
Blocks	Length (km)	128,964	42,139
	Area (km^2)	4,238,174	1,550,206
High-class asphalt roads	Length (km)	8,809,723	5,181,447
	Area (km^2)	70,167,393	37,209,076
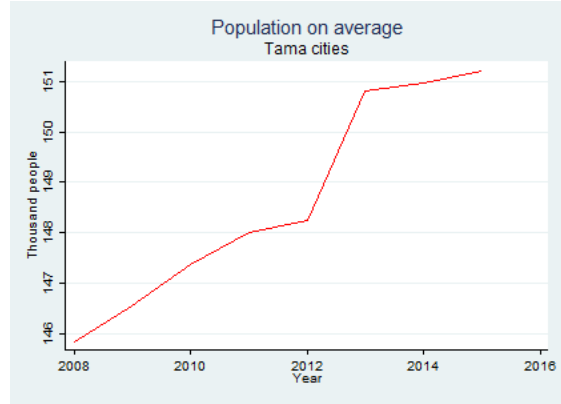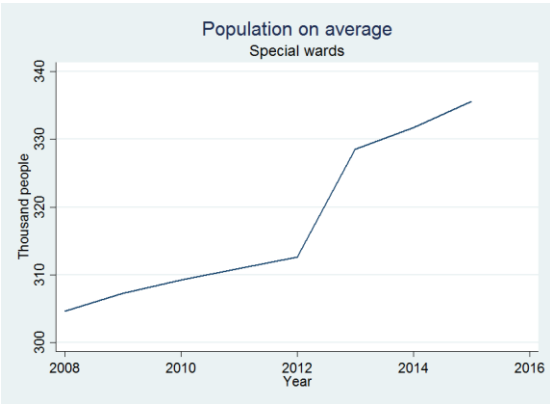
(a) The population on average between Special wards and Tama cities
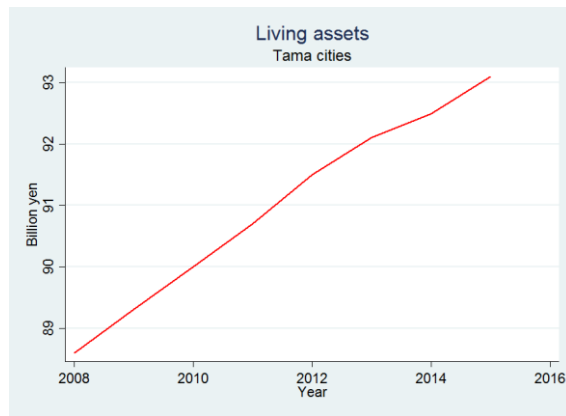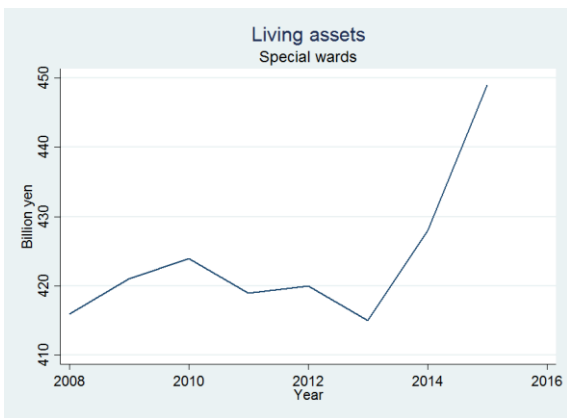
(b) Living assets between Special wards and Tama cities
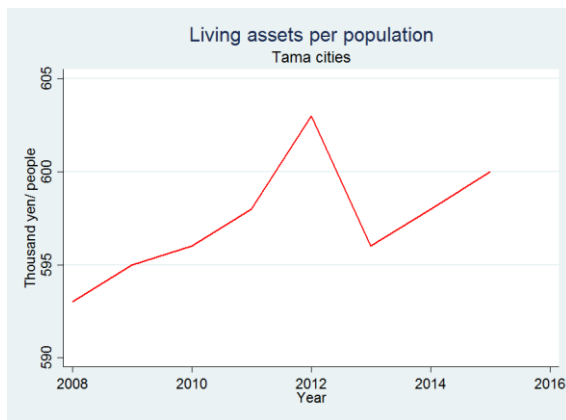
(c) Living assets per population between Special wards and Tama cities

Figure 17. Interregional comparison of living assets per population.

CHAPTER 6. ARE PERFORMANCE MEASUREMENT SYSTEM USED FOR INCENTIVE OR EXPLORATORY PURPOSES?

Budgeting activities in the public sector serve not only to plan expenditures in the year ahead but also to measure and control public organizational performance. In the previous chapter, we suggested that to innovate public financial management in Tokyo local governments, accrual-based budgeting should be adopted. However, designing accrual budgeting without complementary performance measurement cannot provide a comprehensive picture of cost and resource allocation to local authorities and policymakers (Martí, 2013; Schick, 2007). Therefore, performance measurement may be necessary to reinforce and supplement accrual budgeting in the design and implementation of budgetary activities. This chapter attempts to investigate the contemporary use of PMS in TMG and local governments, and to draw some implications to effectively support the introduction of accrual budgeting.

6.1 Introduction

Performance measurement (PM) has increasingly garnered much attention in terms of public policy developments and practices under the umbrella of NPM doctrines since the 1980s (Cavalluzzo & Ittner, 2004; Verbeeten, 2008; Verbeeten & Speklé, 2015). Focal points of this paradigm are to highlight output controls by emphasizing results-oriented management rather than processes-oriented management and to establish a regime of PM and explicit standards for accountability and efficiency improvement (Hood, 1991). The seminal definition of PM is the process of quantifying the efficiency and effectiveness[8] of action (Neely et al., 1995).

---

[8] Neely et al. (1995) define effectiveness as the extent to which customer requirements are met, and efficiency as a measure of how economically the organization's resources are utilized when providing a given level of customer satisfaction.

Measurement systems are adopted and implemented in the public sector to reduce the complexity – in nature, public provisions are multiple-values and many purposes – into a single dimension, facilitating organizational management (De Bruijn, 2007). PM thus becomes a viable management tool for examining organizational efficiency (Johnsen, 2005), challenging legitimacy (De Bruijn, 2007; Johnsen, 2005), enhancing accountability, and improving the quality of decision-making (Cavalluzzo & Ittner, 2004).

Performance measurement systems (PMS), considered "vital catalysts for performance improvement" in many OECD countries, play a pivotal role in public sector reforms in conjunction with the results-based culture (Verbeeten & Speklé, 2015). As a core element of management control systems (MCS), PMS may enable authorities to control and monitor the feedback of internal and external practices (through the bundle of PM data collected from the PMS) so as to adjust deviations and promote organizational change, innovation, and learning (Henri, 2006). Hence, the use of PMS is expected to facilitate strategy implementation and promote organizational performance (Franco-Santos et al., 2012; Speklé & Verbeeten, 2014; Verbeeten & Speklé, 2015).

Following the wave of NPM originating from Anglosphere countries in the 1980s (e.g. UK, Australia, New Zealand), Japan's democracy has integrated NPM principles into public governance since the late 1990s (Yamamoto, 2003). Due to the differential institutional settings and cultural aspects, NPM practices in Japan, an industrialized Asian country, are slightly distinct from several Western nations. Regarding PM, the National Diet (assembly) of Japan enacted the Government Policy Evaluation Act (GPEA) at the central government in April 2002, which was deemed similar to the Government Performance and Results Act (GPRA) promulgated by the US federal government in 1993 (Yamamoto, 2003). However, the

145

performance evaluation movement at the local level has surpassed the central level program of performance. For instance, the Mie prefecture – the precursor employing administrative evaluation – initially introduced far-reaching public management reforms, providing a precursor for other prefectures and cities such as Saitama prefecture, Sapporo city, and Kawasaki city (Eshima et al., 2001; Tanaka, 2009). In tandem with these administrative units, TMG has dedicated itself to applying PM in its organization in departments, division, and sections, but whether the system has become effective remains to be seen, presenting an avenue for research.

As far as we understand, the body of knowledge associated with PM is nascent in Tokyo metropolis. Thus, we aim to investigate the effectiveness of PMS in TMG. We examine what factors drive the use of PMS in TMG. Particularly, we wish to investigate whether the adoption and implementation of PMS is for incentive or exploratory use. Since a dozen departments, divisions, and sections of TMG have exercised some management tools (for example, BSC, KPI) to measure performance comprised of financial and nonfinancial measures, we intentionally study the role of financial and nonfinancial metrics in the mediating relationship between PMS use and several of its determinants.

To achieve these purposes, this study used a quantitative research method to examine PMS. First, a research framework and questionnaire survey were designed and analyzed and a follow-up questionnaire survey carried out. Second, we conducted semi-structured interviews with several departments in local governments for complementary quantitative findings and to expand our insight into the PMS. Based on our results and findings, some policy prescriptions would be implied and suggested for public managers and policymakers alike.

The remainder of this chapter is structured as follows. In the next section, we introduce the theoretical perspectives on PMS, the use of PMS, and the determinants affecting the use of PMS in TMG. Section 3 presents the strategic research method. Section 4 demonstrates the results and findings of the data analysis. Finally, we conclude with discussions, policy prescriptions, and remarking conclusions.

6.2 Literature Review

6.2.1 Performance Measurement System in the Public Sector

A plethora of PM and PMS notions have been introduced in the corpus of literature, varying between fields and individual perspectives in past decades. The most commonly accepted concept was proposed by Van Dooren (2015) from the viewpoint of logic production. Van Dooren (2015, p. 20) defines performance as the result of a production process or outputs and outcomes and PM as a bundle of activities[9] to achieve information on performance. PMS is a set of mutually dependent elements (inputs, processes, outputs, and feedbacks), in which the outputs provide the performance information (Bourkaert & Van Dooren, 2002).

The rationale behind the design and use of PMS in PSOs in the wave of NPM relies on its potential substantial benefits. First, PMS is expected to promote accountability and transparency in public management (Cavalluzzo & Ittner, 2004; Craig, 1999; De Bruijn, 2002, 2007; Dooren, Bouckaert, & Halligan, 2015; Pidd, 2012). This is because performance information used for planning and budgeting is made explicit for various kinds of stakeholders (e.g. citizens, public servants, enterprises, central government). Consequently, stakeholders can hold budget practitioners accountable for financial performance. Second, PMS can urge local authorities to

---

[9] The activities are composed of defining a measurement object, the formulation of indicators, data collection, data analysis, and reporting (Van Dooren, 2015, p.32)

enhance the quality of decision-making (Cavalluzzo & Ittner, 2004; De Bruijn, 2007). Focusing on performance-based results for management purposes enables public authorities to gauge the deficiency gaps between expected and actual results and attempt to negate them in order to achieve strategic objectives and goals. Moreover, PMS can trace the translation of inputs (e.g. expenditure, human capital) into outputs (e.g. public services provisions), assisting decision-makers in improving allocative efficiency. Third, PMS can stimulate learning and innovation (De Bruijn, 2007; Dooren et al., 2015; Hernandez, 2002). Indeed, performance-based results utilizing time-series and cross-sectional data enable authorities to make a comparison and benchmark performance within and between organizations, prompting organizations with lower performance to learn from the best practices of those with higher performance. Fourth, PMS can measure individual performance and rewards (or sanctions) of public servants who attained (did not attain) the strategic goals and targets (Craig, 1999; De Bruijn, 2007; Pidd, 2012). Finally, linking PMS with budgeting efficiently and effectively facilitates planning and control (Craig, 1999; Dooren et al., 2015; Hernandez, 2002; Pidd, 2012). Thus, PMS is regarded as a vital management tool for results orientation.

However, despite its merits in the public sector, there have been detrimental effects. At first glance, adoption of measurement systems has been subject to strategic manipulation and gaming in the execution and evaluation of performance reports (Cuganesan et al., 2014; De Bruijn, 2007). Such incidents of ratcheting effects, cherry-picking service delivery, data manipulation, and misrepresentation of data distort the actual performance insight of organizations (Cuganesan et al., 2014). This leads to a distortion of performance information and outputs (Dooren et al., 2015). Second, although PMS encourages organizational learning, PMS adoption may result in imitation or copying of best practices instead of learning (De Bruijn,

2007). Consequently, this blocks innovations, decreasing ambitions (De Bruijn, 2002, 2007) and

restricting flexibility (Cuganesan et al., 2014). Hence, PSOs become more bureaucratic

(Cuganesan et al., 2014; De Bruijn, 2007). The most important limitation may arise when PMS

and strategies and targets are misaligned (Cuganesan et al., 2014). This can arise because local

authorities concentrate on short-term or operational performance measurement (e.g.

measurement myopia) irrespective of established long-term initiatives or objectives.

Subsequently, many believe that the design and implementation of PMS in the public sector is

akin to a double-edged sword. PMS can be efficacious if utilized in the right manner; otherwise,

it could have adverse effects on management.

6.2.2 The use of PMS: Incentive-oriented and exploratory use.

The growing body of literature on performance measurement and management has

presented multiple uses of a PMS in management control systems. Pidd (2012, p. 31) synthesized

prior studies and summarized them into six categories: planning and improvement, monitoring

and control, evaluation and comparison, accountability, financial budgeting and planning, and

individual performance management. From the other perspectives, Hansen & Van der Stede

(2004) proposed four roles: operational planning, performance evaluation, communication of

goals, and strategy formation, in which the first two roles are short-term and operational in

nature and the last two are long-term and strategic. Henri (2006) classified the use of PMS by top

management into monitoring (feedback provision), attention focusing (signal of performance

from the top managers to lower levels), strategic decision-making (decision-facilitating process),

and legitimizing (decision or actions adjustment). Similarly, Franco-Santos et al. (2007)

proposed five different categories for the role of PMS: measurement performance, strategy

management, communication, influencing behaviors, learning, and improvements. Though the

use of PMS can vary according to the perspective employed, the use of PMS can essentially be

stratified into two levels: operational and strategic.

	Incentive-oriented use	**Exploratory-oriented use**
Hansen & Van der Stede (2004)	• Short-term and operational in nature • Performance evaluation	• Communication of goals • Strategy formation
Henri (2006)	• Monitoring	• Attention focusing • Strategic decision making
De Bruijn (2007)	• Appraising • Sanctioning	• Creating transparency • Learning
Speklé & Verbeeten (2014)	• Target setting • Incentives • Rewards	• Priority setting • Double loop learning • Policy development
Van Dooren (2015)	• Hard use of performance information • Account giving	• Soft use of performance information • Learning

Figure 18. Dimensions of incentive and exploratory use.

Empirical studies have provided some insights and dimensions into two classifications of

PM (Figure 18). For instance, Hansen & Van der Stede (2004) consider incentive use as being

short-term and operational in nature, primarily used in evaluating performance. It also has

exploratory use as a function in the communication of goals and strategy formation. Similarly,

Henri (2006) regards monitoring to be conducted for incentive use and attention-focusing and

strategic decision-making to be instruments for exploratory use. De Bruijn (2007) categorizes

150

rewards (appraising) and punishments (sanctioning) as incentive use but creating transparency

and learning as exploratory use. Likewise, incentive-oriented uses are target-setting, incentives,

and rewards, while exploratory-oriented use is priority setting, double-loop learning, and policy

development (Speklé & Verbeeten, 2014). Van Dooren (2015) addresses the concept of hard and

soft use of performance information. PMS is used for exploratory purposes in the case of

ambiguous goals, immeasurable outputs, varied activities, and unknown intervention (Verbeeten,

2008). In sum, performance information for incentive use encourages managers and employees

to make decisions for efficiency and accountability, while performance information for

exploratory use can serve to learn strategy and develop policy.
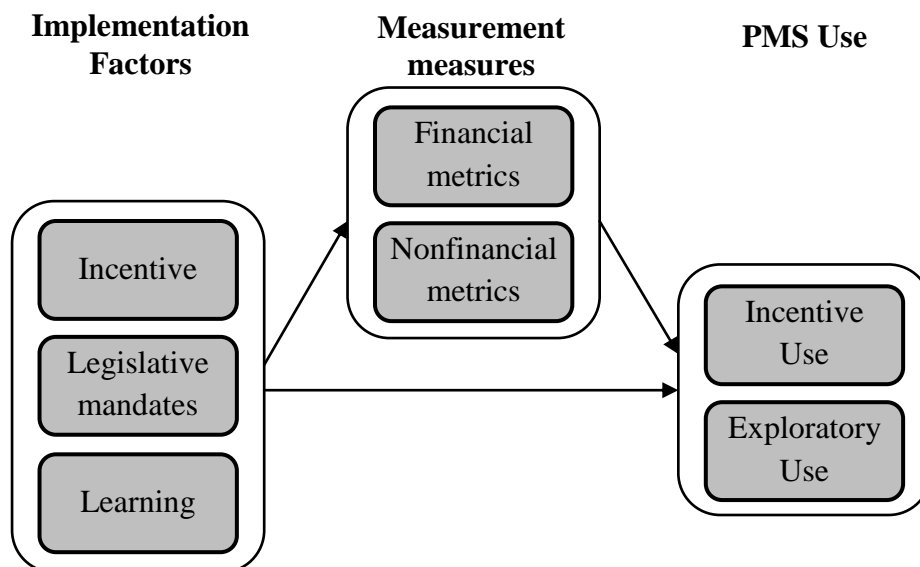
Figure 19. A conceptual framework linking among implementation factors, measurement measures, and PMS use.

In this study, we tend to examine the various driving factors determining both intensive-

oriented and exploratory use of PMS in a Japanese context (Figure 19). Prior literature indicates

that relying on one theory could not be sufficient to explain contemporary PM (Franco-Santos et

al., 2012; Verbeeten, 2008); thus, we employ integrated theoretical frameworks, including

agency theory, contingency theory, and institutional theory (these theories are complementary) representing three constructs: incentive, learning, and legislative mandate, respectively, presumed to affect the use of PMS. In addition, we will examine financial and nonfinancial indicators theorized to correlate with the use of PMS. Therefore, we attempt to construct an exploratory research framework, as follows.

Financial and nonfinancial measures

In traditional control systems, financial management techniques (e.g. budgetary control, Economic Value Added) have been frequently used in practice, yet financial measures appear to be insufficient for organizational control (Otley, 1999). Thus, in addition to more traditional performance measures, nonfinancial measures – performance information in non-monetary terms – might be useful to complement the inherent drawbacks of financial measures (Verbeeten, 2005). For instance, solely using financial indicators for motivating managers and employees can distort their performance because they may attempt to game (deliberate manipulation) within short-term activities to maximize their compensation irrespective of long-term objectives; therefore, nonfinancial indicators in addition to financial ones may enable a more accurate reflection of their performance (Ittner et al., 2003). Ittner et al. (2003, p. 729) also suggest that "... financial measures alone are unlikely to be the most efficient means to motivate employees,...incentives based on nonfinancial measures can improve contracting by incorporating information..." in the context of the US banking industry. Thus, we similarly assume that financial measures as well as nonfinancial measures[10] can affect an incentive use of PMS. In our study, while some financial indicators are extracted from the financial statements such as the ratio of income to expenses, the ratio of income to assets, some nonfinancial

---

[10]

measures such as citizen satisfaction, quality scores, office workplace satisfaction indicators were also used. Hence, we pose the hypothesis:

H1: The *incentive use* is positively affected by (a) *financial measures* and (b) *nonfinancial measures.*

PMS may empower and support strategy formulation and implementation (Cavalluzzo & Ittner, 2004; Hansen & Van der Stede, 2004; Henri, 2006), resulting in the necessity to use both financial and nonfinancial measures linked to strategic planning and objectives (Franco-Santos et al., 2012). The main function of PMS exploratory use is to formulate, implement, and make strategic decisions (Speklé & Verbeeten, 2014). Therefore, a hypothesis is posed that financial and nonfinancial metrics are also used for the exploratory purpose of a PMS.

H2: The *exploratory use* is positively affected by (a) *financial measures* and (b) *nonfinancial measures.*

Incentive

The NPM approach has heavily relied on output-related performance measures including pay, bonuses, and career prospects; hence, the concept of pay-for-performance has become prevalent in the public sector over the last several years (Frey, Homberg, & Osterloh, 2013; Verbeeten, 2008). Output-based payment is considered an incentive or extrinsic motivation, where individual achievements (e.g., rewards, promotion) are linked to performance (Bonner, Hastie, Sprinkle, & Young, 2000). Thus, individuals are presumed to prefer increasing their utility by achieving the strategic goals of the organization, because "extrinsic incentives in the form of monetary rewards motivate individuals' additional or marginal effort" (Frey et al., 2013).

From the perspectives of principal-agent theory in the public sector, agents (e.g., public servants) are delegated by one or groups of individuals as principals (e.g., citizens, the high tiers of government) (Baiman, 1990). Agents have a responsibility to respond to the demands of and to be accountable; however, there is a preference to disclose less than the full amount of information in their own self-interest to principals who offer the resources. Nonetheless, in the design and implementation of PMS, public managers explicitly exhibit multiple-dimensional performance measures in the form of reports with the aim of reducing the information asymmetry between agents and principals and enhancing organizational performance (Franco-Santos et al., 2012). Consequently, this enables principals to advocate incentives or extrinsic motivation (e.g., pay-for-performance, compensation) according to agents' performance: on the one hand, motivating agents to contribute to their organizations' success instead of solely maximizing their utility; on the other, aligning their incentives with the principals' objectives. Therefore, we assume in this study that incentive or extrinsic motivation links to the role of incentive PMS use through the instruments of target setting, incentive provision, and rewards (Speklé & Verbeeten, 2014). The hypothesis is posed as follows:

H3: The *incentive* is positively associated with *incentive use*.

The importance of financial and nonfinancial performance indicators for managers to control outputs of employees is relevant to individual incentives since these measures provide managers with helpful information for the purposes of rewarding and sanctioning. Simons (2000, p. 207) indicates that "to achieve financial data and nonfinancial goals, managers must rely on the efforts and initiative of employees". In other words, individuals are motivated to achieve financial and nonfinancial goals as target setting. Prior literature also argues that both intrinsic and extrinsic motivation to target setting is associated with both financial and nonfinancial

154

measures (Lau & Roopnarain, 2014). In this study, we assume that the incentives of individuals are linked to financial and nonfinancial metrics.

H4: The *incentive* is positively associated with **(a)** financial measures and **(b)** nonfinancial measures.

Learning

According to De Bruijn (2007, p. 12), PM promotes learning in cross-sectional organizations, even within the organization. This is because PMS increases organizational transparency for individuals and organizations to learn and innovate with the aim of improving performance. Evidence of this can be seen through the feedback of information that can facilitate double-loop learning (Franco-Santos et al., 2007). Particularly, in the uncertainty that can adversely affect the clear goal and target setting, it is preferable to use PMS in an exploratory manner to reduce the prevalence of gaming effects associated with monetary rewards (Frey et al., 2013). In other words, a PMS should be contingent on the change in goals and targets resulting from environmental fluctuation to prioritize policies. This argument is based on the contingency theory that an organization's characteristics (e.g. PMS use) and organizational performance are dependent on specific contingencies in a specific period (Franco-Santos et al., 2012). Each organization might be operated in accordance with environmental uncertainty. Thus, there is a necessity for individuals and managers' double-loop learning for continuous performance improvement and loss performance avoidance. Frey et al.'s prior literature (2013) also confirms that when performance information is used to spur double-loop learning, performance measures can perform an exploratory role. Hence, we assume that learning has an influential impact on the exploratory use of PMS:

H5: The *learning* is positively associated with *PMS exploratory use*.

Financial and nonfinancial data extracted from the PMS could be valuable for individuals and managers to learn and innovate. Ittner et al.'s findings (2003) indicate that financial and nonfinancial results affect future financial performance through the learning process. Accordingly, we pose a hypothesis that (individual) learning is associated with both financial and nonfinancial metrics.

H6: The *learning* is positively associated with **(a)** *financial measures* and **(b)** *nonfinancial measures*.

Legislative mandate

PM implemented by Japanese local governments has transitioned further into the mainstream of performance evaluation predominantly used by the US federal government, incorporating features such as "separation of powers" and active legislation, rather than a reliance on executive dominance and reactive legislation typically associated with UK performance regimes (Talbot, 2006). Japan, by contrast, lies down somewhere in the middle between two extremes (Talbot, 2006). In the case of TMG, some departments, divisions, and sections have been undertaking measurement systems to gauge performance. As expected, financial information is very important for decision-making, but we underestimate the importance of nonfinancial information, because both serve as input information for managerial actions. In fact, several agencies in TMG have adopted business-like management tools (e.g., BSC, KPI) composed of financial and nonfinancial measures.

Institutional theory points out that "organizations compete not just for resources and customers, but for political power and institutional legitimacy, for social as well as economic fitness" (DiMaggio & Powell, 1983, p. 150). Therefore, organizations generally make an effort to gain legitimacy using adopted management control systems aligned with environmental requirements and external expectations (Cavalluzzo & Ittner, 2004; De Bruijn, 2007). In fact, under the pressures of regulations and institutional constraints, the organizations' PMS provides various external institutions with multi-dimensional performance information in order to maintain their legitimacy (Modell, 2004). However, Cavalluzzo and Ittner (2004) cast doubt on the PM innovation implementation being responsive to legislative mandates or external requirements. They find that performance measures are designed to meet the GPRA requirement, but performance information may be not used to some extent because PM development under the pressure of outside actors is likely to be symbolic and will have little impact on internal control. Hence, we have assumed that legislative mandates do affect PMS incentive use in the case of TMG. Conversely, according to the Local Government Finance Law, the higher tier of government requests local government to report on financial indicators (e.g., surplus/deficit ratio, consolidated loss ratio, real public debt, future burden ratio) (Local Government Finance in Japan, 2014, p. 198). Therefore, a hypothesis is posed that institutional significance and legislative mandates restrict the design and implementation of performance measures regarding financial and nonfinancial indicators.

H7: The *legislative mandate* is positively associated with **(a)** *financial measures* and **(b)** *nonfinancial measures*.

H8: The *legislative mandate* is positively associated with *incentive use.*

6.3 Research methods

The explanatory sequential mixed methods employed in this study analyze the impact of incentive and exploratory use of performance measures on the PMS in the Tokyo local governments. Creswell (2013) defined the explanatory sequential mixed methods as the research procedure, first utilizing the quantitative analysis to address the critical research questions and subsequently conducting the qualitative analysis for a further explanation of the quantitative results with qualitative data in one study. It was, therefore, apparent that a two-stage process was required. In the first stage, we established and tested the hypotheses using a statistical approach, namely the Partial Least Square Structural Equation Model (PLS-SEM). In the second stage, we administered semi-structured questionnaires to collect cognitive information from respondents. Despite being more complicated and time-consuming than single design research, mixed-methods can provide a more comprehensive picture and better understanding of research issues compared to either quantitative or qualitative approaches in isolation (Creswell, 2013; Ivankova, Creswell, & Stick, 2006; Tashakkori & Teddlie, 2010).

In phase 1, we established the research framework and sent out the questionnaire survey to authorities at the TMG. The purpose of this investigation was to predict which of the aforementioned determinants were associated with PMS use and whether this link was also controlled by the legislative mandate factor. In phase 2, we consulted with some selected key participants at specific administrative departments of Tokyo local governments for further exploration of PM in use and the situation of contemporary budgeting procedures (Figure 20).

Phase 1: The quantitative method

Sampling and data collection

Non-probability samples were collected from the public servants at the departments, divisions, and sections of TMG. We dispersed the questionnaire form through e-mail and participants were able to access the URL and respond with their answers online. The active email addresses of potential respondents were retrieved through the TMG's accounting department. We disseminated the questionnaire under the guidelines and regulations of the Ministry of Education, Culture, Sport, Science, and Technology (Japan). Data collection for phase 1 was conducted from November 10th to December 9th, 2017. We also expanded our study by the follow-up questionnaires.

Measurement

We devised the original questionnaire survey in English and translated it into Japanese for the respondents. A Japanese public accounting scholar in academia was in charge of English–Japanese translation in consultation with several experienced and senior public officials so that the questions reflected the Japanese organizational context and were understandable for participants. The responses in the Japanese version were retranslated into English. All constructs in this study, including management commitment, incentive, training, clear and measurable goals, strategic decision-making authorities, and learning are the first-layer latent variables and the construct of PMS use is the dependent variable. Each construct was reflected by ordinary variables with a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The construct of incentive (INCENTIVE) was extracted from the scholarly work of Verbeeten (2008), while the learning (LEARNING) and legislative mandate (LEG_MAN) were established by authors based on the TMG's actions. The mediating constructs were financial (FIN) and nonfinancial (NONFIN) measures/metrics, whereby we adopted the seminal study of Verbeeten

(2005). Finally, we employed performance measures used for incentive purposes (INC_USE) and exploratory use (EXP_USE) previously employed by Verbeeten & Speklé (2015).

Data Analysis

To assess the proposed theoretical framework, we use the Partial Least Squares (PLS) path modeling, a variance-based structural equation modeling (SEM), which estimates both measurement (outer) and structural (inner) models with multi-item variables. As second-generation techniques of multivariate data analysis, there are two types of SEM: covariance-based SEM (CB-SEM) and Partial Least Squares SEM (PLS-SEM) (Fornell & Larcker, 1981; Hair et al., 2017). As compared with CB-SEM, the PLS path model offers a more powerful tool; because of it, this approach has been widely applied in social sciences as it demands minimal criteria on measurement scales (working with various kind of metric data, quasi-metric scaled data, and binary data), sample size (a small one is acceptable), and residual distribution (soft distribution assumptions) (Chin, 1998, 2011; Hair et al., 2017; Lee et al., 2011). Furthermore, reporting the global Goodness-of-fit (GOF) indices and Chi-square (χ^2) statistics in the PLS reports is not prominently expected, while the evaluation of the same in CB-SEM is compulsorily required (Hair et al., 2017; Chin, 2011). Indeed, Hair et al. (2017) state that "CB-SEM is primarily used to confirm (or reject) theories, whereas PLS-SEM is primarily used to develop theories in exploratory research". The most important reason for the adoption of PLS path modeling rests on the fact that, in this study, the investigation of PM essentially incorporates and intertwines a variety of theories such as behavioral, economic, and institutional theory. Consequently, PLS-SEM seems to be fitting for our explanatory and predictive purposes of critical factors affecting PMS use.

160

For the measurement model, we evaluate the quality of the model (reliability and validity) in the following procedures. First, the internal consistency reliability is displayed by the **Cronbach's alpha** (representing the lower bound) and **composite reliability** (representing the upper bound). Second, we test the convergent validity through the **average variance extracted (AVE).** Finally, the **heterotrait-monorail ratio (HTMT)** of the correlation indicates discriminant validity.

For the structural model, we assess the performance of the research model including the significance of path coefficients and R^2 and delineate the results of hypothesis testing. Based on the result of the regression analysis, we further aim to examine the correlation among the constructs and further variables participants responded to on budgeting. Finally, we represent the global PLS-SEM-based model GOF measures (even these indicators seems not to be strictly required in PLS path modelling as opposed to CB-SEM).

Phase 2: The qualitative method

Participant selection

In the second phase, we exploited the multiple case study approach (Yin, 2009) by conducting semi-structured interviews with the aim of enriching our understanding of public officers' perceptions on the reasons why some determinants significantly predict PMS use, but others do not. We also posed the research question of whether contemporary budgeting is affected by the performance measures and attempt to further explore hindrances of budgeting procedures in Tokyo local governments. We intentionally selected several departments from the local governments for further responses to the interview. We designed the open-ended and semi-

structured questions based on the first stage to explore more reasons behind the results of statistical analysis.

It is important to note in our research procedure that at the outset, we aimed for the participants of Tokyo local governments for our research for phase 1 and 2. However, participants at local governments were willing to participate in the interviews, but unwilling to answer the questionnaire survey, while those in TMG were favorable with our questionnaire at phase 1, but hesitant to have the interview at phase 2. Then, we conducted the questionnaires at TMG for phase 1 and interviewed with public managers at departments of some Tokyo local governments for phase 2 alternatively. Even though the objectives of two focusing groups existed discrepancy, these two objectives still work closely at Tokyo prefecture. Specially, they share the promotion of Public Sector Accounting Innovation Association in the pursuit of efficiency and performance improvement.

Phase 1   **Phase 2**

| Quantitative approach | Qualitative approach |

- Data collection (Online survey N = 108)
- Demographic characteristics
- Evaluation of measurement model.
- Assessment of structural model
- Follow-up questionnaire survey

- Interview protocol
- Individual interview (N = 6).
- Coding and thematic analysis.
- Interpretation and discussions of quantitative and qualitative analysis.

Figure 20. Research procedure for sequential explanatory quantitative and qualitative mixed methods.

6.4 Research results

6.4.1 Phase 1: The quantitative method.

6.4.1.1 Demographic profile.

The questionnaire was returned by 108 respondents over the survey period; however, of these, 64 were eliminated because their organizations had not yet adopted and implemented the PMS (we tried to avoid biased responses by using the filter or contingency question). The remaining 44 samples were used for further analysis. Most of the responses were collected from public officials working in the department of spatial urban planning and housing (16.7%), finances and accounting (13.9%), water supply and sewerage (13%), environment and disaster countermeasure (12%), education, science, and technology (9.3%), health and welfare (7.4%), and so on. The outputs of their works were essentially to provide public goods (28.7%), public service (39.8%), and mixed products of public goods and services (18.5%). In short, the demographic characteristics of respondents imply that most public servants approached were from departments, divisions, and sections associated with public service provision.

6.4.1.2 Measurement model assessment

Internal consistency reliability

Internal consistency, defined as the degree to which instrument items are homogeneous and reflect the same underlying constructs (Cooper & Schindler, 2014), was calculated using Cronbach's alpha (representing the lower bound) and composite reliability (representing the upper bound) (Fornell & Larcker, 1981). The minimal cutoff score of all variables for Cronbach reliability was higher than 0.6 (Nunnally & Bernstein, 1994) and 0.7 (Cronbach & Meehl, 1955;

Hair et al., 2010). Most of the constructs in our analysis meet the basic demands of Cronbach reliability, being higher than 0.60. These values range from 0.610 to 1.00.

Conversely, composite reliability should be greater than the benchmark of 0.7 (Fornell & Larcker, 1981). The results in Table 16 indicate that all composite reliability of constructs have a value greater than 0.7, meeting the sufficient condition of internal consistency reliability. It is important to note that the Cronbach's alpha and composite reliability of EXP-USE is 1.00 because only one indicator fully reflects this construct.

Convergent validity (AVE)

Convergent validity, by definition, is the extent to which a measure correlates positively with alternative measures of the same construct (Hair et al., 2017). The outer loadings of the indicators and the average variance extracted (AVE) measure the nature of convergent validity. According to Hair et al. (2017), indicators with outer loadings between 0.4 and 0.7 should be considered for removal from the measurement scale. We have excluded some items in this range. Then, we again conducted the analysis and obtained indicator loadings higher than the criterion of 0.7 (Hair et al., 2010), showing the adequacy of convergent validity (see Appendix 1). Additionally, the AVE value should be higher than the cutoff of 0.5 (Chin, 1998; Fornell & Larcker, 1981; Hair et al., 2010), which was exceeded by all constructs under analysis (see Table 12), suggesting that all latent variables could explain a substantial part of each indicator's variance of at least 50%.

Discriminant validity (Heterotrait - Monotrait Ratio - HTMT)

Discriminant validity is the extent to which a construct is truly distinct from other constructs by empirical standards. Researchers have commonly relied on the traditional approach

164

to evaluate discriminant validity (the cross-loadings and Fornel-Larcker criterion), yet we have employed novel measures (the **heterotrait-monotrait ratio** abbreviated by **HTMT**) in this study. An HTMT value greater than 0.90 implies a lack of discriminant validity (Hair et al., 2017). Our results of HTMT values range from 0.117 to 0.709 (with the confidence interval at 95%), being smaller than the threshold of 0.90 and consequently implying that all constructs in our model seem to be unique. In turn, these constructs do not overlap in their representation of each other.

Table 16. Summary of measurement model assessment

Construct	Mean	S.D.	Cronbach's Alpha	Composite reliability[11]	AVE[12]	Discriminant Validity
			(0.60-0.90)	>0.7	(>0.5)	HTMT[13]
INCENTIVE	3.193	0.909	0.804	0.883	0.718	Yes
LEG_MAN	2.477	1.128	0.790	0.865	0.683	Yes
LEARNING	2.505	1.220	0.895	0.928	0.765	Yes
FIN	1.878	0.988	0.782	0.872	0.695	Yes
NONFIN	2.057	0.992	0.783	0.858	0.604	Yes
INC_USE	2.223	1.094	0.619	0.840	0.724	Yes
EXP_USE	3.570	1.180	1.000	1.000	1.000	Yes

Note: HTMT ranging from 0.117 to 0.709 (<0.90)

---

[11] Composite Reliability $\rho_c = \dfrac{(\sum \lambda_i)^2 \operatorname{var} F}{(\sum \lambda_i)^2 \operatorname{var} F + \sum \Theta_{ii}}$

[12] Average Variance Extracted $AVE = \dfrac{(\sum \lambda_i^2) \operatorname{var} F}{(\sum \lambda_i^2) \operatorname{var} F + \sum \Theta_{ii}}$

[13] HTMT confidence interval does not include 1.

6.4.1.3 *Structural model assessment*

The path coefficients and R^2 indicate how well the model is performing in the structural model assessment (Chin, 1998; Hair et al., 2017). Regarding PLS-SEM path estimation in Figure 21, Table 17 summarizes the results of the estimation with regards to the hypothesized relationship between the constructs. It was found that INCENTIVE had a statistically significant and positive effect on INC_USE ($\beta = 0.320$, $p = 0.009$), congruent with agency theory. While INCENTIVE is associated with NONFIN ($\beta = 0.331$, $p = 0.040$), it does not significantly affect FIN. This argument is consistent with the statement that, as opposed to nonfinancial measures, financial measures are considered as historical and backwards-looking so they lack the predictive ability to explain future performance and rewards (Henri, 2006). In fact, TMG public servants probably focus on nonfinancial measures (e.g. citizen satisfaction, workplace satisfaction, quality scores) to achieve goals or objectives regarding nonfinancial indicators, rather than financial ones. Yet, both FIN and NONFIN are not significantly associated with INC_USE, suggesting that the establishment of financial and nonfinancial indicators do not affect the incentive-oriented use of PMS.

For the LEG_MAN, FIN measures are significantly and strongly influenced by LEG_MAN ($\beta = 0.503$, $p = 0.000$), but NONFIN measures are not. It can be explained that the external norm (e.g. law at the national level) determines and motivates the use of financial measures/metrics for performance measurements rather than internal regulations (e.g. TMG orders). The institutional restrictions probably provide local governments with instructions in the establishment of specific financial indicators. However, nonfinancial measures are discretionarily devised and operated unilaterally by the public institution. There is no significant association in the direct relationship between LEG_MAN and INC_USE. This is consistent with

166

the regulations from the central government coercing local governments to undertake financial

indicators in issuing a bond and controlling debt (Local Government Finance in Japan, 2014), but

does not impact incentive-oriented use. Akin to Cavalluzzo and Ittner (2004), this result is

consistent with the institutional theory that externally mandatory acts appear to be symbolic

because the financial and nonfinancial indicators were not explicitly employed for incentive-

oriented use of PMS at the organizational level.
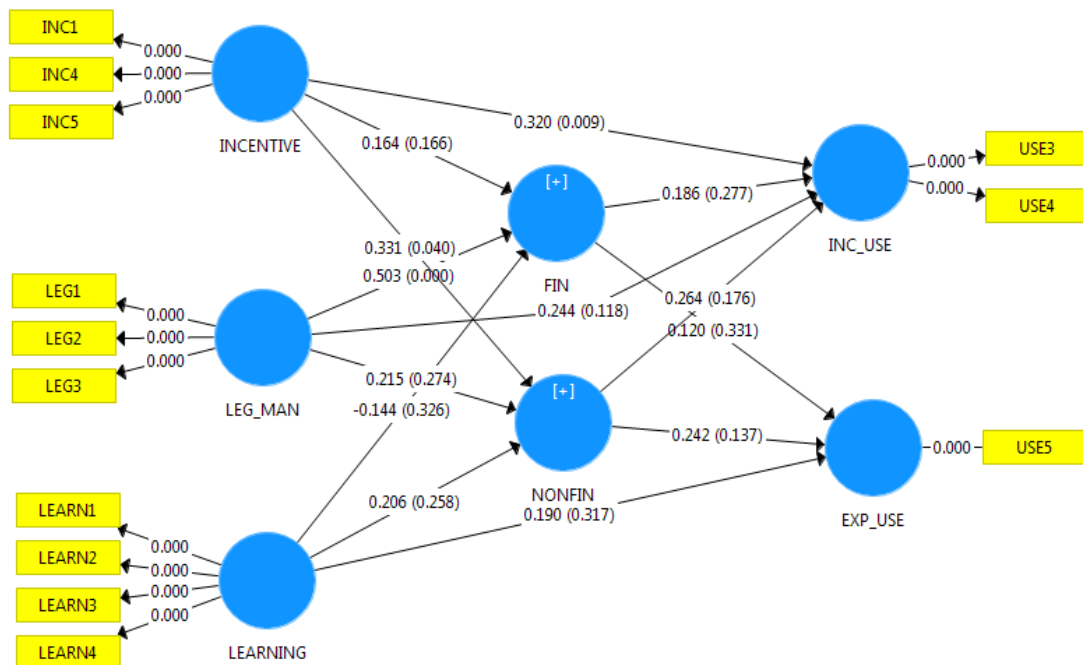
Figure 21. Results of path analysis.

For the dependent constructs (INC_USE and EXP_USE), while there are no significant

associations between EXP_USE and its antecedents, only INC_USE is affected by INCENTIVE.

The findings suggest that PMS is used for incentive rather than exploratory purposes. In other

words, TMG establishes the PMS to evaluate performance based on whether public servants

attain their targets and objectives for the sake of rewards and promotions at the individual level

and to benchmark performance in the previous year for budgetary control at the organizational

level. Although it cannot be conclusive, TMG does not seem to use PMS for priority setting in a strategic manner, long-term plan development, policy making, communication goals, and double-loop learning.

Table 17. The summary of path coefficients, t-value, and p-value.

H	Hypothesis	Path Coefficient[14]	t-value	p-value	Support
H3	*INCENTIVE → INC_USE*	0.320	2.670	0.009	Yes***
H4a	*INCENTIVE → FIN*	0.164	1.387	n.s.	No
H4b	*INCENTIVE → NONFIN*	0.331	2.169	0.040	Yes**
H7a	*LEG_MAN → FIN*	0.503	4.101	0.000	Yes***
H7b	*LEG_MAN → NONFIN*	0.206	1.101	n.s.	No
H8	*LEG_MAN → INC_USE*	0.244	1.562	n.s	No
H6a	*LEARNING → FIN*	-0.144	0.955	n.s.	No
H6b	*LEARNING → NONFIN*	0.206	1.095	n.s.	No
H5	*LEARNING → EXP_USE*	0.190	0.940	n.s.	No
H2a	*FIN → INC_USE*	0.186	1.159	n.s.	No
H2b	*FIN → EXP_USE*	0.120	0.980	n.s.	No
H1a	*NONFIN → INC_USE*	0.264	1.396	n.s.	No
H1b	*NONFIN → EXP_USE*	0.242	1.563	n.s.	No

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

While the path coefficients indicate the significance of path estimates, the R^2 value or coefficient of determination shows the variance of constructs and explains the power of the model. There is no generalized statement about the acceptable threshold value of R^2. Notwithstanding, the greater the R^2, the stronger the explanatory power of variance in the model. Moreover, while a consensus has not yet been reached, Hair et al. (2011) proposed a rule of

---

[14] Path coefficients have been standardized with the range between -1 and +1. Estimated path coefficients approach to +1 means strong positive relationship and vice versa for negative values. Value closing to 0 represents weak relationships (Hair et al., 2017).

thumb in the field of marketing that R^2 values of 0.75, 0.50, and 0.25 could be respectively substantial, moderate, and weak. In our structural model, the R^2 value for mediating constructs (FIN and NONFIN) is 0.258 and 0.248, reflecting that the model provides an acceptable explanation of variance. Additionally, the R^2 value for dependent construct (INC_USE) is moderate (0.474), but the construct (EXP_USE) is weak (0.159). Finally, the overall fit of the structural model ($\chi^2 = 280.098$, $\chi^2/df = 1.87$, SRMR = 0.11) showed an acceptable fit to the data (Hair et al., 2017), while other indicators (NFI = 0.540 and RMS = 0.249) suggested a low explanatory power, explained mostly by the study's limited sample size.

6.4.2 Follow-up questionnaire survey.

6.4.2.1 Semi-structured designs and implementation.

In this phase, we aim at extending our knowledge and understanding of the impact of driving factors on incentive-oriented and exploratory PMS use in TMG by means of semi-structured questions presented to participants involved in the first phase, for the sake of unveiling some nuanced arguments. In doing so, we designed a follow-up questionnaire protocol in terms of open-ended inquiries and semi-structured questionnaires (See Table 18) and sent it to antecedent participants.

We received 13 responses from the 44 respondents in the previous sample. Most of these respondents (Q1) were middle-range managers (seven, accounting for 54%), followed by employees (five staff, accounting for 38%), and one chief of department (accounting for 8%). In their department/division/section, PM instruments (Q2) used for performance evaluation in the last several years were comprised of accounting reports for budget settlement (69.2%), financial reports and indicators concerning inputs and outputs (30.8%), business process indicators (i.e., productivity, number of works per hours, time consumed) (23.1%), and the staff satisfaction

index (23.1%) or nonfinancial indicators. In fact, financial indicators are likely to prevail over nonfinancial indicators in case of TMG. The new public accounting regime (accrual basis) was initially employed in TMG from 2006 and the performance evaluation was established following this reform. However, most of the respondents did not recall the exact time when the performance evaluation was established in their department/division/section. Notwithstanding, the performance evaluation was set up in 2008 or 2009, which was contingent on the individual department/division/section (Q3).

6.4.2.2 Findings of follow-up questionnaires.

First, we observed that the PMS in department/division/section projects and programs undertaken by TMG public officials were mainly utilized for budgeting (84.6%), monitoring the achievements of goals (69.2%), setting up the targets for future projects (61.5%), and coordinating their projects and programs *vis-à-vis* other external ones outside of TMG (15.4%) (Q4). This suggests that performance evaluation roles are subject to budgeting (84.6%), monitoring (69.2%), setting targets (61.5%), and coordinating the goals (15.4%), illuminating that incentive-oriented and exploratory uses are recognizable at TMG despite the greater propensity for incentive-oriented uses.

Second, the outcomes of PMS use are generally financial and nonfinancial information. In the same vein as PMS roles, the information reported in financial statements by means of the new public accounting system associated with the functions of budgeting (61.5%), setting up the targets for future projects (23.1%), monitoring (15.4%) and coordinating their projects with other agencies (15.4%) was used in their department/division/sections (Q5). Additionally, this information was mainly utilized for resource allocation implementation (53.8%) and control of business activities (46.2%) (Q6). The performance information on resource allocations is aligned

with budgetary activities. Indeed, in the departments where PMs are actively conducted, there is a tendency to utilize the accrual information obtained for budgetary purposes. This suggests that these departments at TMG have the potential to introduce an element of accrual budgeting.

We further asked to what extent the performance evaluation impacts budgeting. Here, 38.5% of respondents indicated that they have no idea while 46.2% argued that there is an effect to some extent (Q7). Apparently, performance information including financial and nonfinancial metrics is used for budgetary purposes to some extent, though this link is relatively weakened. In fact, under the auspices of the national government, when devising a financial plan for the year ahead related to some national works and programs, the TMG has been generally influenced to use the previous year's budget rather than the performance information from PMS (Tanaka, 2009; Cabinet, 2017). Hence, even the performance evaluation is used for budgeting at the outset; in reality, this use is limited to some extent. Therefore, this obstruction might lead to the rejection of hypotheses 1 and 2, that financial and nonfinancial metrics do not affect incentive-oriented and exploratory-oriented uses.

Third, respondents reflect their perspective on incentive-oriented use of PMS (Q8) that they were motivated to contribute to TMG's overall goals and mission (76.9%), contribute to society (their departments as well as external organizations) (69.2%), and gain their self-interest (rewarding, increasing salaries) (15.4%). According to these data, we can infer that public servants' motivation is prone to organizational commitment and missions and goals achievement because they have an understanding of the success of their department in tandem with their own personal success (hence individuals can achieve the rewards and promotions). Thus, they are less likely to prioritize their own personal interests.

Table 18. Semi-structured question protocol

Dimensions	Questions
Background information (Warm-up stage)	▪ Q1. Could you please introduce some information about your background profiles (which department/division/section do you belong to; what are your position and responsibility,...)? ▪ Q2. What kind of performance management tools (e.g. BSC, KPI) has your department/division/section adopted? ▪ Q3. When was it introduced and applied in your department/division/section?
PMS use	▪ Q4. What purposes do you think that the current PMS is used for?
Financial and Nonfinancial measures	▪ Q5. What do you think about the role of financial and nonfinancial metrics in PMS? ▪ Q6. How were the financial metrics used in your department/division/section? ▪ Q7. How did the performance evaluation affect the annual budgeting?
Incentive	▪ Q8. For what purposes do you think public servants try to commit and achieve strategic goals? (rewards, promotions,...?) ▪ Q9. Do you think the individual incentive and motivation affect target settings and PMS utilization at an organizational level?
Learning	▪ Q10. Do you seek external information from other sources and other organizations to establish and determine performance indicators? To set targets for individuals in your department?
Legislative Mandates	▪ Q11. When designing performance indicators, are you imposed upon to conform to external norms (e.g. Government Policy Evaluation Act (GPEA) enacted in 2002 by central government) and institutional restrictions (e.g. law)? ▪ Q12. Do external norms constrain financial and nonfinancial measures?
Closing stage	▪ Q13. Do you have any comment or recommendations for us?

Next, for the question asking about individual motivations affecting overall achievement or performance (Q9), 69.2% respondents supported this argument, indicating a high correlation between INC and INC_USE (approximately 70%). This could reaffirm hypothesis 3 in the first stage. However, it would be prudent to draw a distinction between the results found in the first (PLS-SEM) and second stage (PLS-SEM is a cause-effect model while follow-up questionnaire is a correlational statistics model). In the first stage, the individual incentives tend to personal interests (e.g., bonus, rewards, promotions), but follow-up surveys show that most incentives serve society and the community.

Fourth, some participants had no ideas on the validity and appropriateness of performance evaluation instruments (46.2%) while some perceived that there exists a value of performance evaluation to some extent (46.2%) (Q10). It could be explained that, in practice, some public servants use the indicator of budget execution rate[15] as the main performance indicator, while others notice the budget execution rate to be inappropriate as a performance indicator. Moreover, some public managers seem to take the budget execution rate as an indicator for inputs rather than outputs and outcomes. Hence, we can observe that budget execution rate is a vague performance indicator in practice. In this case, we reiterate the support for the introduction of accrual budgeting to support performance measurement in TMG.

Fifth, the respondents had no knowledge or experience of situations whereby external information influenced the setting of performance evaluation indicators and targets (46.2%) (Q11). This means that public servants might place less emphasis on learning and not use the external information to set the performance indicator. Thus, learning seems not lead to exploratory use that is consistent with hypothesis 5. Additionally, they did not recognize the

---

[15] Budget execution rate means budgeting consistency between planned and actual amount.

influence of institutional regulations on designing and implementing performance evaluation (46.2%), while some advocated for legislative requirements (52.9%) (Q12). This could be congruent with hypothesis 7a, where legislative mandates affect the design and implementation of financial metrics.

To sum up, the follow-up survey further fortifies our finding from the SEM path analysis that the personal incentives had a direct impact on PMS incentive use and designing nonfinancial indicators. Furthermore, legislative mandates are positively associated with designing financial indicators. In general, the driving factors are prone to affect PMS incentive-oriented use rather than PMS exploratory use in TMG.

6.4.3 Phase 2: The qualitative method.

6.4.3.1 Demographic profile.

We intended to conduct the semi-structured interview (phase 2) at TMG along with the questionnaire done in phase 1, but could not because we were not allowed to make public servants' perception explicit as a representative voice of TMG. Hence, we conducted the interviews at the local level, where we selected some advanced local governments of Tokyo metropolis in terms of operation of the public accounting system. Even though the interviewees' objectives were slightly different, we designed the subjects by closely working with TMG's workable accrual accounting system development in the presence of the New Public Accounting System Promotion Liaison Conference.

Table 19. Interview background information

No.	Municipal	No. of participants	Department	Place	Duration	Time
1	M city	3	Public accounting	On-site	1 hour 42 min.	January 22nd 2018
2	F city	1	N/A	On-site	34 min.	March 22nd 2018
3	A ward	1	N/A	On-site	45 min.	March 27th 2018
4	E ward	1	Public accounting	Telephone	19 min.	March 21st 2018
5	F city	1	Public accounting	Telephone	25 min.	March 21st 2018
6	H city	1	Public finance	Telephone	30 min.	March 21st 2018

Note: In F city, we conducted the interview by both on-site and telephone call.

Eight interviewees from five local governments in Tokyo participated in phase 2 (see Table 19). Most were middle managers in public accounting departments and finance divisions. We conducted the semi-structured interview on-site or face-to-face meetings (we used anonymous names for municipals participating in our semi-structured interview as M city, F city, and A ward) as well as by telephone call (anonymous names for E ward, F city, and H city) in March 2018. We conducted the interview in groups of three in M city, whilst individual interviews were undertaken at other local governments.

6.4.3.2 Interview findings

We conducted the semi-structured interview at five local governments of Tokyo metropolis, which replaced our participant objectives (aiming at public managers and servants at TMG) at the outset. Though the interview content was less likely to reflect our expectation, being congruent with the PLS-SEM model findings, we at least gained some essential sensory perceptions of public managers on PM and budgeting at the local level, which offered some addenda to our reasoning deduction. By the content analysis technique, we found the most keywords in the conversations. For example, the most frequently used keywords were *budget (275), financial (154), plan (137), evaluation (135), administrative (128),* etc and most frequent
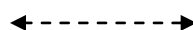
two keywords are *financial statements (107), administrative evaluation (79), implementation plan (75), public accounting (74)*, etc. (Figure 22). Next, we selected two keywords and created a valid connection between them (Figure 23).
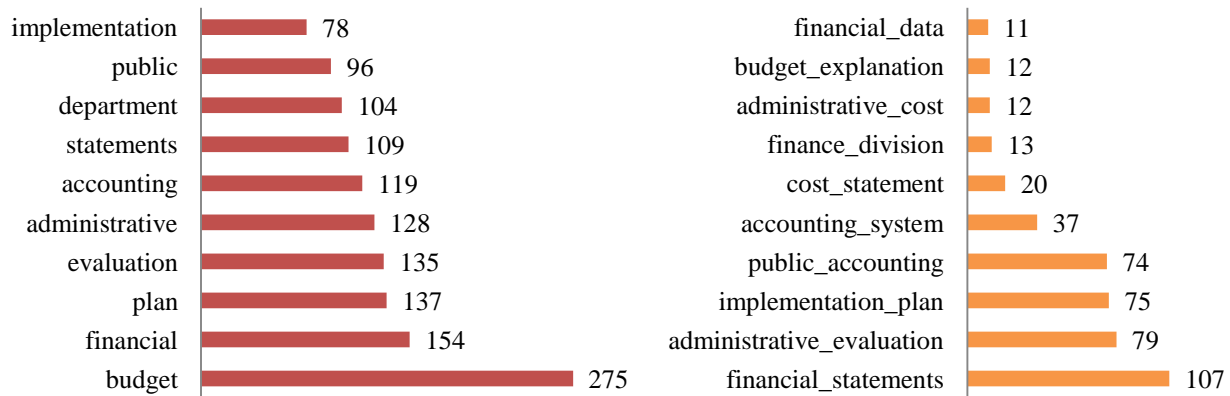
Figure 22. Content analysis for top-ten key one word and two words
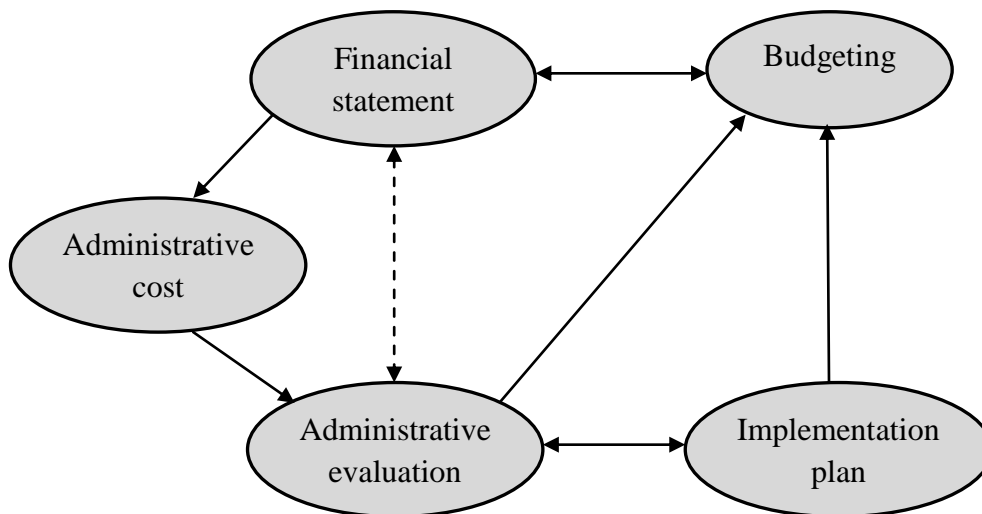
Figure 23. Map highlighting the connection of themes for individuals

Two-sides effect

The effect is unclear

One-side effect

First, we determined whether financial statements at the local level supporting the administrative evaluation[16] was unclear; however, some municipals (e.g., M city, E ward, F city, H city) asserted that there had been an indirect and partial link between financial statements and administrative evaluation through the administrative cost (Figure 23). In fact, an interviewee in E ward said: "…*We are preparing financial statements published, but there is no direct bidding on the part of direct administrative evaluation now….*". Hence, it could not be conclusive, but we could observe hypotheses 2a and 2b that financial information derived from financial statements seemed not to support the use of PMS.

Second, regarding employee incentives, M city had a management tool (e.g., process chart), aligning the personal goals with department goals. In essence, the goals were set by employees (or bottom-up method) and the goal achievement was gauged by the personnel evaluation system in the case of M city. As a result of personnel measurement, performance information would be published to constituencies and one could be *naming and shaming* if not achieving final goals or targets (public manager uses the A, B, C level to point out personal achievement of goals at the end of every fiscal year). To avoid negative consequences or even obtain positive achievement such as promotion and rewards, one must be self-motivated and inevitably obtain department achievement as well. In other words, one uses PMS for his/her own incentives, supporting hypothesis 2. Furthermore, the individual incentive is generally provoked by such a management tool as ranking A, B, C level (that is a non-financial indicator rather than financial indicator), consistent with hypothesis 4a.

---

[16] The term "administrative evaluation" used at the local level is similar to "performance evaluation" at TMG. However, administrative evaluation represents whole local government performance, but performance evaluation at TMG expresses the performance of a department/division.

Third, concerning the learning construct, although the training cannot be a full reflection of learning, there is a somewhat partial nexus between two concepts that training can offer external information such as market information, governmental regulations, local-national financial relations, and demographical information to individuals in strategically setting goals and devising programs. Indeed, some municipals (e.g., F city, A ward, E ward, and H city) provide public servants with several hands-on training courses and workshops annually, which were instructed by certified accounting experts. However, the effects of training in the use of PMS in setting goals and strategies were indeterminate. Even though it is a very common belief that training and education lets learners and practitioners be positively impacted in the long term, learning seems to be less supportable for exploratory use of PMS in this case (hypothesis 1b) and moreover, learning seems fragmented in using non-financial indicators for personal evaluation (hypothesis 1a).

Fourth, relating to legislative mandate, most local governments must prepare the financial statements in compliance with the policy evaluation implementation guidelines issued by the Ministry of Internal Affairs and Communications. It is apparent that the guidelines affects which public managers at financial and accounting department record financial data (hypothesis 6a is supported). Among local governments, M city is exposed to be "superior" in financial reporting and statement by dividing into the smaller units of departments to make daily entries in the ledger. Thus, financial statements are always ready for submission to higher tiers of government immediately if requested.

Fifth, we found some interaction among budgeting, financial statements, administrative evaluation, and implementation plan at the local level (Figure 23). Firstly, it is argued that the financial reporting and statement reiteratively interact with budgeting activities in some local

governments such as M city and A ward. Particularly, these municipals frequently adopt the

PDCA (Plan – Do – Check – Act) cycle in operation and stress on C (Check) daily and monthly.

The manager of M city said: *"…this effort of financial statements is positioned within the PDCA*

cycle, and in the case of settlement, C. We grasp the issue in the financial statement. If we find

any problem, we can fix the problem before submitting to the Congress and budgeting for the

next fiscal year…".

Sixth, we also found a positive effect of administrative evaluation on the budgetary

activities. It may maintain an argument that outputs of performance measurement system link to

budget or funding mechanisms to some extent, as defined by Robinson (2013) as performance

budgeting. However, in some cases, this link seems ambiguous. In fact, some departments have

embraced the incremental budgeting. For example, the E ward manager argued: *"…I think it is*

better to make budgeting basically based on previous year than to create it on a zero basis…"

and H city maintained: *"…Evaluation result of fiscal year 29 should be based on the fiscal year*

28…". Lastly, each local unit basically must formulate the implementation plan, which is usually

a 10-year term. Some divide 4/3/3 years for medium terms (also known as Medium Term

Budgeting Framework – MTBF), while others undertake 5/5 years term. Although most

municipals devise the implementation plan for long-term purposes, there is a considerable need

for adjustment in a single year relating to budget because of some external factor changes (e.g.,

aging ratio) influencing pecuniary planning. It could be maintained that the annual budget and

implementation plan are not linked well and tend to be operated independently and the current

system has problems in this respect.

All in all, we found some evidence that cities seem more "superior" than special wards in

accounting and financial affairs, though there needs to be more empirical supporting evidence.

Based on the results of the PLS-SEM model at TMG, it can be inferred that PM aligning with budgeting at TMG seems akin to what financial and accounting affairs are fundamentally exercised in cities.

6.5 Concluding Remarks and Policy Implications

6.5.1 Concluding remarks

This research set out to predict determinants related to incentive, legislative mandates, and learning, which affect PMS incentive-used use and/or exploratory use in the case of TMG since the introduction of performance evaluation initiatives. We conducted the research by employing the mixed research methods, including path modeling analysis and follow-up questionnaire survey at the first stage and semi-structured interviews at the second stage. Our research targets were public officials who have had an engagement in the performance evaluation system at their department/division/section of TMG and were interviewed at Tokyo local governments.

The findings clearly indicate that personal incentives to contribute to organizations' goals and missions and achieve rewards and promotion are positively associated with the organizational PMS incentive-oriented use. Furthermore, the personal incentives also affect the use of nonfinancial metrics, meaning that public officials are motivated for organizational commitment through nonfinancial indicators such as citizen satisfaction and productivity rather than financial indicators prepared by the new public accounting system.

Under the auspices of legislative determinants, guidelines on performance evaluation are positively associated with design and implementation of the financial metrics, while the nonfinancial metrics are more likely to be discretionary, varying between individual prefectural

governments. Moreover, the empirical evidence has also shown that legislative requirements do not directly influence PMS incentive use, which is similar to the US, where "implementation of externally-mandated control systems is likely to be symbolic, with little influence on internal operations" (Cavalluzzo & Ittner, 2004, p. 265).

Regarding the role of learning, it seems that the PMS exploratory use is not associated with learning functions. It cannot be conclusive, but we maintain that the factors that encourage public servants to set strategic targets and expenditure priorities and develop public policy seems to be more closely linked to external information rather than organizational learning. In sum, it can be reasonably inferred from our evidence that performance evaluation initiatives in TMG are used for incentive purposes, which outweigh exploratory purposes. The results were also supported by the semi-structured interviews, although it cannot be clearly confirmed.

6.5.2 Policy implications

Performance budgeting is supported by the accrual budgeting system (implications in chapter 5) (Martí, 2013; Robinson, 2009a). Schick (2007) affirms that for better performance management, it would be necessary to adopt accrual budgeting. The foundation of performance budgeting is performance evaluation adoption and implementation. Hence, the accrual accounting system and performance measurement system should go hand-in-hand in order to better support public authorities in setting up strategic goals, allocating resources, and budgeting budgets for a long-term period (Robinson, 2009, 2016). In reality, departments at TMG have actively engaged in performance evaluation and performance information has been used for budget purposes at TMG. There is a tendency to utilize the accrual information for the budget in conjunction with strategic and operational planning (Hansen and Van der Stede, 2004). This

suggests that these departments have a good potential to introduce the elements of accrual budgeting, which makes performance budgeting more substantive.

6.5.3 Limitations and future research

This study inevitably faces the drawback of limited numbers of survey participants. We notice that the number of respondents at the first stage was limited at 44 samples with a lower response rate (under 20%). Even if it was not large enough for our analysis, we accepted the reality and used the bootstrapping technique in PLS-SEM analysis. Furthermore, at the second phase, interviews with TMG public officials would have been ideal for our purposes, but they did not have enough time for face-to-face conversations, so we designed a follow-up survey with semi-structured and open-ended questions for them. Furthermore, we conducted interviews at local governments, although the perceptions on PMs of local managers were slightly different from those of TMG public officials. For future research, we should expand our analysis into study on the impact of PMS incentive and exploratory use on overall organizational performance perceived by public officials at TMG.

Appendix 2: Survey questions, description of items, factor analysis results and indicator loadings

Incentive (INCENTIVE)

Please indicate the extent to which you agree with the following statements: (1 = strongly disagree, 5 = strongly agree).

	Mean	S.D.	Indicator Loadings
Achieving the goals of the project is important as an indicator of the degree of contribution to the goals or mission of the organization.	3.80	1.18	0.735
The achievement of strategic goals is an important signal to reward for department members.	2.55	1.18	0.903
The achievement of strategic goals is important for the promotion of department members.	2.86	1.06	0.894

Legislative mandate (LEG_MAN)

Please indicate the extent to which you recognize: (1 = strongly disagree, 5 = strongly agree).

	Mean	S.D.	Indicator Loadings
External norm (e.g., laws) rather than regulations set by TMG affect the establishment, operation, and use of specific indicators.	2.68	1.50	0.878
Besides the question above (e.g., comparability with other local government bodies and requests from other important organizations), institutional restrictions affect the establishment of specific indicators.	2.93	1.43	0.725
Your department/division/section is consulted to create and report performance information by external professional organizations.	1.81	1.04	0.868

Learning (LEARNING)

If you are (or were) a department/division/section's manager, please indicate the extent to which you: (1 = strongly disagree, 5 = strongly agree).

	Mean	S.D.	Indicator Loadings
Seek external information sources (e.g., market information, statistical information, demographic information, etc.) to establish your department/division/section's strategic programs.	2.80	1.47	0.906
Seek external information related to other government organizations (e.g., information about the strategic programs adopted by the other local governments or private enterprises, etc.) to determine indicators used in programs/projects.	2.52	1.36	0.913
Seek external information related to other government organizations (e.g., information about the strategic programs adopted by the other local governments or private enterprises, etc.) to monitor the programs/projects' achievement.	2.55	1.39	0.924
Seek information related to other government organizations (e.g., information about the operational process adopted by the other local governments, etc.) to establish individual operational targets for department members	2.16	1.30	0.743

Financial measures/metrics (FIN)

To what extent your department/division/section uses the following performance measures: (1 = strongly disagree, 5 = strongly agree).

	Mean	S.D.	Indicator Loadings
We use the indicators extracted from the financial statements.	2.05	1.21	0.851
We use the ratio of income to expenses (Percentage of specified resources to expenditure in general accounts), indicating how much of the expenses required for a specific project are covered by the income.	2.05	1.43	0.802
We use the ratio of the income generated from the	1.55	0.89	0.847

project to the assets (Percentage of specified resources to assets in general accounting), indicating how much the assets invested in the business generate incomes.

Nonfinancial measures/metrics (NON_FIN)

To what extent your department/division/section uses the following performance measures: (1 = strongly disagree, 5 = strongly agree).

	Mean	S.D.	Indicator Loadings
We use the ratio of outputs to inputs.	2.16	1.35	0.701
We use the indicator of customer satisfaction measures (e.g., citizen satisfaction, etc.)	2.50	1.42	0.712
We use the indicator of work quality (e.g., quality scores, number of defects, etc.)	1.88	1.18	0.883
We use the indicator of employee satisfaction (e.g. office workplace satisfaction, job turnover rate)	1.68	1.15	0.799

Incentive Use of PMS (INC_USE)

Please indicate the extent to which your department/division/section uses the performance measures: (1 = strongly disagree, 5 = strongly agree).

	Mean	S.D.	Indicator Loadings
To coordinate with other departments/organizations' programs (including inside and outside the Tokyo Local Governments)	2.73	1.30	0.854
To set up the individual targets and monitor the degree of achievements.	2.32	1.24	0.848

Exploratory Use of PMS (EXP_USE)

Please indicate the extent to what your department/division/section uses the performance measures: (1 = strongly disagree, 5 = strongly agree).

	Mean	S.D.	Indicator Loadings
To select and decide programs/projects.	3.57	1.18	1.000

Appendix 3: Questionnaire survey in Japanese language

PMS仕事に関するアンケート

主催：首都大学東京大学院社会科学研究科経営学専攻

アンケート調査についてのお願い

拝啓　平素は格別のご高配を賜り、厚く御礼申し上げます。

さて、私ども首都大学東京大学院社会科学研究科経営学専攻野口研究室では、
かねてより予算・決算情報の相互フィードバックに関する研究に従事してまいりました。

現在、私たちは、自治体における、予算・決算を含む「業績評価システム」に関する日豪比較研究プロジェクトに従事しています。
今回は、東京都庁とニュー・サウス・ウェールズ州政府で利用・活用されている「業績評価システム」に関する比較研究を企図しております。

そこで、東京都庁職員の皆様に、下記アンケート調査について、ぜひご協力を頂ければと思い、
失礼とは思いましたが、このようなお願いをさせて頂きました。

ご回答頂いた内容は、両自治体の予算・決算システムを含む「業績評価システム」の運用方法に関する重要な示唆として
活用していきたいと考えております。
自治体公会計制度の利活用研究の発展のため、趣旨をご理解いただければ幸いでございます。

なお、ご回答頂いた内容については、統計的な解析を施し、学術論文の作成・公表のためのデータとして利用させていただきます。
当該論文の作成・公表に際しては、所属部署等に関する個別具体的な名称は一切用いません。
また、確認をご希望される場合、公表前に記述内容をご確認いただき、公表後に作成した論文を寄贈させていただきます。
なお、ご回答頂いた内容は、学術的な目的のみに使用させて頂き、それ以外の目的には使用いたしません。
文部科学省「文部科学省所管事業分野における個人情報保護に関するガイドライン」に沿って管理いたします。

回答期間は平成29年　11月　10日（金）から同年　11月　24日（金）までの2週間です。

ご多忙中に誠に恐縮ではございますが、ご協力いただきますよう、よろしくお願い申し上げます。

敬具

質問票に関するお問い合わせ

東京側：
首都大学東京大学院社会科学研究科経営学専攻
野口昌良研究室
〒192-0397
東京都八王子市南大沢1-1

Tel+81 (0)42 677 2331 Fax+81 (0)42 677 2298
E-mail　mynoguchi@tmu.ac.jp

NSW側：
Dr Joseph Drew
Institute for Public Policy and Governance
University of Technology Sydney
PO Box 123
Broadway NSW 2007
Australia
Tel +61 2 6778 4843
E-mail　Joseph.Drew@uts.edu.au

以下の質問事項については、ご所属の部署（課）の状況を回答してください。

<u>**パートI:**</u> 所属部署についてお尋ねします。

Q1．ご所属の部署（課）の担当部門についてお教えください。（複数選択可）

☐ 総務部門　　　　　　　　　　　☐ 通信・メディア部門

☐ 財務・会計部門　　　　　　　　☐ 環境・災害対策部門

☐ 教育・科学技術部門　　　　　　☐ 健康・福祉部門

☐ 農林水産部門　　　　　　　　　☐ 水道・下水道部門

☐ 都市計画・住宅部門　　　　　　☐ その他（　　　　　　　　　）

☐ 文化・スポーツ・観光部門

Q2ご所属の部署（課）の主たる業務内容（公共施設の建設・維持の有無）についてお教えください。

☐ 公共施設の建設・維持が中心

☐ 公共サービスの提供が中心

☐ 公共施設の建設・維持と公共サービス提供の双方を行っている

☐ その他

Q3.
ご所属の部署（課）において、「所管する事業の実績を評価し、事業運営や予算編成等に反映・フィードバックさせる取組」（以下、「業績評価の取組」とします）を行っていますか

☐ はい　　　　　　　　　☐ いいえ

Q4. ご所属の部署（課）において、近い将来、「業績評価の取組」を行う予定がありますか？

　　　　□ はい　　　　　　　　□ いいえ

5ポイントリッカートスケール

<u>1</u>　　該当 しない。
<u>2</u>　　あまり 該当 しない。
<u>3</u>　　どちらか 判断 で きない。
<u>4</u>　　ある 程度該当 する。
<u>5</u>　　該当 する。

Q5. パフォーマンス指標/指標

業績評価にあたり現在どのような指標を用いていますか。【 その他の指標以外必須 】

1　　決算の予算に対する執行率を用いている。　　　　　　　　　　　　　　　　　　1 2 3 4 5

2　　財務諸表から得られる指標を用いている。　　　　　　　　　　　　　　　　　　1 2 3 4 5

3　　費用に対する収入の割合等の指標を用いている。（一般会計においては歳出に対す　1 2 3 4 5
　　　る特定財源の割合）特定の事業に要した費用をその事業の収入でどの程度賄ってい
　　　るか、その割合を示すものです。事業別財務諸表を作成した場合、「事業別の収入
　　　／事業別の費用」で求められます。用いていない場合は1を選択してください。

4　　資産に対する当該事業から生じる収入の割合等の指標を用いている。（一般会計に　1 2 3 4 5
　　　おいては資産に対する特定財源の割合）事業に投資した資産がどれだけ収入をあげ
　　　たかを表す指標です。事業別財務諸表を作成した場合、「事業別の収入／事業別の
　　　資産」で求められます。用いていない場合は1を選択してください。

5　　資源投入量（インプット）に対する産出量（アウトプット）の割合等の指標が用い
　　　られている。アウトプットに入院患者数、インプットに病床数を用いて利用度を計
　　　測する、また、アウトプットに年間利用者（来場者）数、インプットに施設設備の
　　　規模を用いて利用度を計測するなどの方法で用いられている例があります。用いて
　　　いない場合は1を選択してください。

6　　業務プロセスに関する指標　　　　　　　　　　　　　　　　　　　　　　　　　1 2 3 4 5
　　　（たとえば、時間当たり処理量や所要時間等の指標）が用いられている。

7　　利用者満足度に関する指標（たとえば、住民満足度調査の結果等の指標）が用いら　1 2 3 4 5
　　　れている。７と８の違いは、７は利用者の主観的な指標で、８は数値などで表すこ
　　　とができる客観的な指標のことです。

8　　業務品質に関する指標（たとえば、欠陥数・割合やクレーム数・割合等の指標）が　1 2 3 4 5
　　　用いられている。７と８の違いは、７は利用者の主観的な指標で、８は数値などで
　　　表すことができる客観的な指標のことです。

9　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　1 2 3 4 5
　　　所属職員の満足度に関する指標（たとえば、職員の職場満足度調査結果や転職率等

の指標）が用いられている。

10　上記以外で用いている指標がありましたら、下欄に具体的な指標の名称をご記入願　1　2　3　4　5
います。

Q6. 性能測定用途

前問の指標の利用目的についてお伺いします。

1　事業の目標達成度をモニターするのに用いている。　　　　　　　　　　　　　　　1　2　3　4　5

2　事業に関連して、翌年度の予算を策定（要求）するのに用いている。　　　　　　　1　2　3　4　5

3　事業について他組織（東京都庁内外を含む。）の関連事業との調整を行うのに用い　1　2　3　4　5
ている。

4　事業に関連して、その所属職員の個別目標を設定し、その達成度をモニターするの　1　2　3　4　5
に用いている。

5　今後取り組むべき事業を選定・決定するのに用いている。　　　　　　　　　　　　1　2　3　4　5

Q7. 明確な目標

事業の目標設定についてお伺いします。事業目標は

1　所属部署の分掌事務と整合している。　　　　　　　　　　　　　　　　　　　　　1　2　3　4　5

2　明確に記述され、所属部署内で共有されている。　　　　　　　　　　　　　　　　1　2　3　4　5

3　明確に記述され、対外的に公表されている。　　　　　　　　　　　　　　　　　　1　2　3　4　5

4　その達成度が、事後的に検証可能な程度に十分に特定され、詳細である。　　　　　1　2　3　4　5

5　その達成度が、事後的に検証可能なように数値化されている。　　　　　　　　　　1　2　3　4　5

Q8. インセンティブと動機

事業の目標達成の位置づけ、取組についてお伺いします。問8の中で該当するものがな
い設問は1を選択してください。

1　事業の目標達成は、所属部署（あるいはその上位部署）の、組織全体の目標（ある　1　2　3　4　5
いは使命）への貢献度を示す指標として重要である。

2　事業の目標達成は、所属部署（あるいはその上位部署）の社会貢献の指標として重　1　2　3　4　5
要である。

3　所属職員は、総じて事業の目標達成に対して積極的に取り組んでいる。　　　　　　1　2　3　4　5

4　事業の目標達成は、所属職員の報酬（給与・手当等）にとって重要である。　　　　1　2　3　4　5

5　事業の目標達成は、所属職員の処遇や昇進にとって重要である。　　　　　1　2　3　4　5

6　事業の目標達成は、所属職員個人の自己実現（達成感や自己成長の実感等）にとっ　1　2　3　4　5
　　て重要である。

7　事業の目標達成は、所属職員個人のスキル・アップ（特定の知識・技術の修得）に　1　2　3　4　5
　　とって重要である。

Q9.　戦略的意思決定権限

所管する事業に関する「課長級の管理者」の意思決定権限についてお伺いします。問９
の中で該当するものがない設問は１を選択してください。

1　新規事業の選定、事業の終了等を評価・決定する権限を有している。　　　　　1　2　3　4　5

2　新規事業の選定、事業の終了等の評価・決定に参画する（会議体に出席して意見を　1　2　3　4　5
　　述べ、会議体の意思決定に参画する）権限を有している。

3　事業の評価に用いる指標を個人で選定・評価・決定する権限を有している。　　1　2　3　4　5

4　事業の評価に用いる指標の選定・評価・決定に参画する（会議体に出席して意見を　1　2　3　4　5
　　述べ、会議体の意思決定に参画する）権限を有している。

5　事業の目標達成度をモニターする権限と責任を有している。　　　　　　　　1　2　3　4　5

6　事業との関連で、所属職員の個別目標を個人で決定する権限を有している。

7　事業との関連で、所属職員の個別目標の決定に参画する（会議体に出席して、意見
　　を述べ、会議体の意思決定に参画する）権限を有している。

8　事業との関連で、所属職員の個別目標の達成度をモニターする権限と責任を有して
　　いる。

Q10.　管理コミットメント

「課長級の管理者」よりも上位の管理者の事業の評価への関与についてお伺いします。
（ご所属の部署（課レベル）から見た主観で結構です）。問１０の中で該当するものが
ない設問は１を選択してください。

1　事業の評価に用いる指標の選定・運用に対して、その権限を十分に行使している。　1　2　3　4　5

2　事業の評価に用いる「発生主義決算（財務諸表）」から得られる指標の選定・運用　1　2　3　4　5
　　に対して、その権限を十分に行使している。

3　実施中の事業に関して適用できると考えられる指標の「選定・評価・決定」に向け　1　2　3　4　5
　　て、部下と効果的な意思疎通を行い、あるいはその参画を求めている。

4　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　1　2　3　4　5
　　実施中の事業に関して現在用いている指標の「運用・利用」に対して、部下と効果

的な意思疎通を行い、あるいはその参画を求めている。

Q11. 養成

事業の実績の評価に関し過去３年間の研修の状況についてお伺いします。問１１の中で該当するものがない設問はNoを選択してください。

		Yes	No
1	局内研修もしくは中央研修の機会が用意されていた。	1	0
2	上記1のうち、とくに「発生主義決算（財務諸表）」の意味と機能について理解する研修の機会があった。	1	0
3	外部で開催される研修もしくは外部委託による研修の機会が用意されていた。	1	0
4	事業の実績を評価するために、所属職員自らが主体的に指標を確立・運用・利用するための発展的な研修の機会が用意されていた。	1	0

Q12. 修業

事業の実績の評価にあたり外部組織から取得した情報の利用状況についてお伺いします。

1	新規事業の選定、事業の終了等を評価・決定するため、「外部情報」（たとえば、マーケット情報、経済統計、地域情報、他の自治体組織で採用されている業績評価対象事業に関する情報等）を利用している。	1	2	3	4	5
2	事業の評価に用いている指標を選定・評価・決定する（あるいは選定・評価・決定に参画する）のに、「外部情報」（たとえば、民間企業や他の自治体等の外部組織で用いられている業績評価指標に関する情報等）を利用している。	1	2	3	4	5
3	事業の目標達成度をモニターするのに、「外部情報」（たとえば、外部組織に委託した利用者満足度調査に関する情報等）を利用している。	1	2	3	4	5
4	事業との関連で、所属職員の個別目標を決定する（あるいは決定に参画する）のに、「外部情報」（たとえば、民間企業や他の自治体等の外部組織で用いられている業務プロセス関連の情報等）を利用している。	1	2	3	4	5

Q13. 立法委任

事業に関して用いている指標（「発生主義決算（財務諸表）」を除く）についてお伺いします。財務諸表から得られる指標以外のすべての指標についての設問です。所属部署では、

1	東京都が定める規則・規定以外の、何らかの外部規範（たとえば、法令・規則・規定等）により、特定の指標の確立・運用・利用が求められている。	1	2	3	4	5
2	上記1以外の制度的制約（たとえば、東京都以外の他の自治体組織との比較可能性の維持あるいは他の重要組織（国等）からの要請）により、特定の指標の確立・運用	1	2	3	4	5

・利用が求められている。

3   東京都からみて外部の専門家組織（たとえば、コンサルタント会社等）からの勧告　　1　2　3　4　5
や推奨に基づき、指標を確立・運用・利用している。

Q14.
事業に関して用いている指標（「発生主義決算（財務諸表）」を除く）についてお伺いします。財務諸表から得られる指標以外のすべての指標についての設問です。所属部署において、現在用いている指標以外の指標の確立・運用・利用が、東京都からみて外部の規範により、「合理的でない」と判断する場合がありましたら、以下に具体例をご記入ください。（たとえば、外部の法令等により、特定の形式以外は使用できない、あるいは使用できても効率的・効果的でない等）

Q15.

「発生主義決算（財務諸表）」についてお伺いします。所属部署では、

1   会計管理局や財務局が定める決算参考書財務諸表や事業別財務諸表に　　1　2　3　4　5
関する規定・規則・ガイドライン（公営企業会計においては地方公営
企業法などの法令等）とは別に、独自の工夫により「発生主義決算（
財務諸表）」を運用・利用している

Q16.
「発生主義決算（財務諸表）」についてお伺いします。前問において、会計管理局や財務局が定める規定・規則・ガイドラインとは別に、独自の工夫により「発生主義決算（財務諸表）」を運用・利用しているとお答えの方にお伺いいたします。上記について、以下に具体例をご記入ください。

Q17.　発生主義予算

実施中の事業に関する予算要求の状況についてお伺いします

1   所属部署（事業執行部署を意味し、管理部門（財務局主計部や各局の経理部門）を　　1　2　3　4　5
除く。以下のすべての設問において同じ。）において、要求する予算等を策定して
いる。

2   所属部署において、予算要求手順・手続は、あらかじめ定められている。　　1　2　3　4　5

3   所属部署において、要求する予算は、前年度の実績をベースに策定される。　　1　2　3　4　5

4   所属部署において、要求する予算は、ゼロ・ベースで策定される。　　1　2　3　4　5

5 　所属部署において、事業の予算編成にあたり、中長期（5〜10年先）の将来予測を取　　1　2　3　4　5
り入れる。

Q18. 発生主義予算

実施中の事業に関する予算要求の状況についてお伺いします

1 　所属部署における予算要求にあたり、中長期的（5〜10年先）の将来予測を取り入れ　　1　2　3　4　5
るのが難しい。

2 　所属部署における予算要求にあたり、事業目標との結びつきが弱い。　　　　　　　　1　2　3　4　5

3 　所属部署において、事業の目標を予算要求に反映させている。　　　　　　　　　　　1　2　3　4　5

4 　所属部署における予算要求にあたり、固定的な要素が大きいため、環境変化に対応　　1　2　3　4　5
できていない。

Q19. 追加的なご意見やコメントがございましたら、以下にご記入をお願いいたします。

Follow-up questionnaires results

Q1. What is your position and responsibility?		13	100.0%
1	Chief of department/division/section	1	7.7%
2	Middle manager	7	53.8%
3	Employee	5	38.5%
4	Other	0	0.0%
Q2. What kind of performance management tools does your department/division/section adopt?		13	100.0%
1	Accounting report for budget settlement.	9	69.2%
2	Indicators extracted from the financial statement reports (i.e., assets, liabilities, administrative costs) prepared by the new public accounting system.	2	15.4%
3	Indicators associated with inputs and outputs.	4	30.8%
4	Business process index (i.e., productivity, number of works per hour, time consumed)	3	23.1%
5	Satisfaction indicators	3	23.1%

6	Indicators associated with failures.	1	7.7%
7	Staff members satisfaction indicators.	0	0.0%
8	Other.	2	15.4%
Q3. When did your department introduce the current performance evaluation?		13	100.0%
1	Unknown	9	69%
2	2008	1	7.6%
3	2009	1	7.6%
4	Before 2010	1	7.6%
Q4. What is the purpose of your own project assessment implemented in the department /department /section you belong to?		13	100.0%
1	Budgeting	11	84.6%
2	Explaining the efficiency and effectiveness of the projects carried out by the TMG.	0	0.0%
3	Monitoring the goals achievement of the projects.	9	69.2%
4	Coordinating our projects with other related organizations	2	15.4%
5	Setting up the targets for the future projects.	8	61.5%
6	Evaluating the personnel or staff.	0	0.0%
7	Other.	1	7.7%
Q5. What information about the financial statements prepared by the new public accounting system is used for the purpose of your department?		13	100.0%
1	Budgeting	8	61.5%
2	Explaining the efficiency and effectiveness of the projects carried out by the TMG.	0	0.0%
3	Monitoring the goals achievement of the projects.	2	15.4%
4	Coordinating our projects with other related organizations	2	15.4%
5	Setting up the targets for the future projects.	3	23.1%
6	Evaluating the personnel or staff.	0	0.0%
7	Other.	4	30.8%
Q6. Do you attach importance to the preparation of the annual budget		13	100.0%

	(original draft before assessment) in the departments in charge?		
1	To evaluate performance	5	38.5%
2	To control business activities	6	46.2%
3	To implement the resource allocation	7	53.8%
4	Other	1	7.7%
Q7. To what extent do you think that the performance evaluation used affect the annual budget in your department/division/section?		13	100.0%
1	It does not affect	0	0.0%
2	No effect	1	7.7%
3	I do not know	5	38.5%
4	Affect to some extent	6	46.2%
5	Have a big impact	1	7.7%
Q8. What is the purposes of public servants to commit and achieve strategic goals?		13	100.0%
1	Contribute to the overall goal and mission of the Tokyo Metropolitan Government Organization	10	76.9%
2	To contribute to the social contribution of the departments (or higher-level departments) in which they are affiliated.	9	69.2%
3	Increase the salary of employees themselves (salaries and benefits)	2	15.4%
4	To make the staff's own treatment and promotion decisions favourable	1	7.7%
5	Self-actualization of achievement and self-growth	5	38.5%
6	To acquire specific knowledge and skills of the staff themselves	4	30.8%
7	Other	1	7.7%
Q9. How much do you think the motivation of each of the members who answered in Q8 will affect the degree of achievement of the project goals or the performance evaluation of the departments?		13	100.0%
1	It does not affect	0	0.0%
2	No affect	1	7.7%
3	I do not know	3	23.1%
4	Affect to some extent	9	69.2%

5	Have a big impact	0	0.0%
Q10. Do you think that the indicators used in performance evaluation are valid and appropriate?		13	100.0%
1	There is room for substantial improvement	0	0.0%
2	There is room for some improvement	0	0.0%
3	I do not know	6	46.2%
4	Some are valid	6	46.2%
5	Very effective and reasonable	1	7.7%
Q11. To what extent do you think that the external information influence to set performance evaluation indicators and set targets?		13	100.0%
1	Do not use it all the time	1	7.7%
2	Do not use it	1	7.7%
3	I do not know	6	46.2%
4	Use some degree	4	30.8%
5	Use for large	1	7.7%
Q12. To what extent do you think that external norms (policy evaluation laws enacted by the government in 2002) and institutional regulations affect to how to design and implement performance evaluation indicators.		13	100.0%
1	Do not use it all the time	0	0.0%
2	Do not use it	0	0.0%
3	I do not know	6	46.2%
4	Use some degree	3	23.1%
5	Use for large	4	30.8%

(SA)Q1b. 回答者ご自身の現在の役職についてお答えください。

		実数	%
	全体	13	100.0
1	課長級	1	7.7
2	統括課長代理級又は課長代理級	7	53.8
3	担当者	5	38.5
4	その他	0	0.0

(MA)Q2. 所属する部署で実施されている「所管する事業の業績を評価し、事業運営や予算編成等に反映・フィードバックさせる取組」（以下「業績評価」という。）は、どのような目的のために実施されていますか？（複数回答可）

		実数	%
	全体	13	100.0
1	次年度予算編成のため	11	84.6
2	都庁全体で実施される事業評価（財務局主管）に向けて利用するため（たとえば、実施した事業の効率性・有効性等について説明するため）	0	0.0
3	所属部署独自で実施している事業の目標達成度をモニターするため	9	69.2
4	所属部署において実施している事業について、他組織（東京都庁内外を含む）の関連事業との調整を行うため	2	15.4
5	所属部署において将来的に取り組むべき目標（あるいは事業そのもの）を選定・策定するため	8	61.5
6	所属職員の人事評価のため	0	0.0
7	その他	1	7.7

(MA)Q3. 業績評価に当たり、どのような種類の指標が所属する部署で採用されていますか？（複数回答可）

		実数	%
	全体	13	100.0
1	決算の予算に対する執行率	9	69.2
2	新公会計制度で作成される財務諸表（資産・負債・純資産・行政コスト・一般財源（税収・補助金）等）のデータから得られる指標	2	15.4
3	資源投入量（インプット）（たとえば施設設備の規模）に対する産出量（アウトプット）（年間利用者数）の割合（施設利用度）等の指標	4	30.8
4	業務プロセス（時間当たり処理量や所要時間等）に関する指標	3	23.1
5	利用者満足度に関する指標（たとえば、住民満足度調査の結果等の指標）	3	23.1
6	業務品質に関する指標（たとえば、欠陥数・割合やクレーム数・割合等の指標）	1	7.7
7	所属職員の満足度に関する指標（たとえば、職員の職場満足度調査結果等の指標）	0	0.0
8	その他	2	15.4

(FA)Q4_1. 所属する部署が現在の業績評価指標を導入したのはいつ頃ですか？

	実数	%
不明	9	69
平成29年度	1	7.6
平成28年度	1	7.6
平成22年度以前から	1	7.6

(MA)Q5. 新公会計制度で作成される財務諸表の情報は、所属する部署において、どのような目的のために使用されていますか？（複数回答可）

		実数	%
	全体	13	100.0
1	次年度予算編成のため	8	61.5
2	都庁全体で実施される事業評価（財務局主管）に向けて利用するため（たとえば、実施した事業の効率性・有効性等について説明するため）	0	0.0
3	所属部署独自で実施している業績評価に向けて利用するため（たとえば、事業目標の達成度をモニターするため）	2	15.4
4	所属部署において実施している事業について、他組織（東京都庁内外を含む）の関連事業との調整を行うため	2	15.4
5	所属部署において将来的に取り組むべき目標（あるいは事業そのもの）を選定・策定するため	3	23.1
6	所属職員の人事評価のため	0	0.0
7	その他	4	30.8

(MA)Q6. 所属する部署で年度予算（査定前の当初原案）を作成する際に重視することはなんですか？（複数回答可）

		実数	%
	全体	13	100.0
1	業績評価との関連性	5	38.5
2	事業活動をコントロールすること	6	46.2
3	事業（あるいは部署）間での優先順位を確定し、適切な資源配分を実施すること	7	53.8
4	その他	1	7.7

(SA)Q7. 所属部署独自で実施している業績評価の結果が、所属する部署の年度予算にどの程度影響しますか（業績評価の結果と予算編成手続の関連性の程度についてお伺いしています。）

		実数	%
	全体	13	100.0
1	まったく影響しない	0	0.0
2	あまり影響しない	1	7.7
3	どちらともいえない	5	38.5
4	ある程度影響する	6	46.2
5	大きく影響する	1	7.7

(MA)Q8. 所属する部署の各職員が、事業目標を達成しようとする動機、あるいは所属部署独自の業績評価に取り組もうとする動機はどのようなものであるとお考えですか？（複数回答可）

		実数	%
	全体	13	100.0
1	都庁組織全体の目標や使命に貢献すること	10	76.9
2	所属部署（あるいはその上位部署）の社会貢献につなげていくこと	9	69.2
3	職員自身の報酬（給与・手当等）を高めていくこと	2	15.4
4	職員自身の処遇や昇進決定を有利なものにしていくこと	1	7.7
5	職員自身の自己実現（達成感や自己成長の実感等）を図ること	5	38.5
6	職員自身のスキル・アップ（特定の知識・技術の修得）を図ること	4	30.8
7	その他	1	7.7

(SA)Q9. Q8.で回答した各構成員の動機が、部署での事業目標の達成度、あるいは業績評価の取り組みにどの程度影響を与えると思いますか？

		実数	%
	全体	13	100.0
1	まったく影響しない	0	0.0
2	あまり影響しない	1	7.7
3	どちらともいえない	3	23.1
4	ある程度影響する	9	69.2
5	大きく影響する	0	0.0

(SA)Q10. 業績評価で用いられる指標は、有効・妥当なものであると思いますか？

		実数	%
	全体	13	100.0
1	大幅な改善の余地がある	0	0.0
2	多少の改善の余地がある	0	0.0
3	どちらともいえない	6	46.2
4	ある程度は有効・妥当である	6	46.2
5	非常に有効・妥当である	1	7.7

(SA)Q11. 業績評価で用いられる指標の設定やその目標値の設定にあたり、外部の情報をどの程度利用していますか？

		実数	%
	全体	13	100.0
1	まったく利用しない	1	7.7
2	あまり利用しない	1	7.7
3	どちらともいえない	6	46.2
4	ある程度利用する	4	30.8
5	大いに利用する	1	7.7

(SA)Q12. 業績評価で用いられる指標の設定に当たり、法令・規則といった制度上の規制や、国からの通知・要請といった外部の基準がどの程度影響すると思いますか？

		実数	%
	全体	13	100.0
1	まったく影響しない	0	0.0
2	あまり影響しない	0	0.0
3	どちらともいえない	6	46.2
4	ある程度影響する	3	23.1
5	大きく影響する	4	30.8

CHAPTER 7. CONCLUDING REMARKS AND POLICY IMPLICATIONS

This chapter encapsulates the study. The first section summarizes the research results and findings. In the second, we provide concluding remarks and implications for public policy, which may aid decision-makers and local authorities of Tokyo local governments alike. Finally, we discuss some limitations of this thesis and outline a future research agenda.

7.1 Summary of research results

Regarding the public finance topic in Chapter 3, we investigate the association between the volatility of various incomes – local tax, intergovernmental grants, subsidies, local bonds – and the volatility of local expenditure in 49 Tokyo local governments within the 2008–2015 period by FE panel regression. This study has four main findings. First, due to the difference in financial mechanisms for special wards and Tama cities, there was a significant dissimilarity in the volatility of revenues and expenditures. Second, there was a significantly positive association between local tax volatility and spending volatility, which is congruent with Sacchi and Salotti's findings (2017). Third, ordinary grants had a significantly negative association with the local expenditure volatility. There might be a good reason for smoothing out the volatile spending in the face of the financial crisis. Finally, local bonds have a positive impact on local spending volatility, though the correlation coefficient is relatively limited.

In Chapter 4, we examined the public efficiency (what is known as technical efficiency from the economic viewpoint) of 49 Tokyo local governments from 2001 to 2015. We adopted the global DEA approach in terms of the CCR and BCC models to calculate the efficiency scores of local units. We observed that the efficiency scores in special wards had declined over the period under analysis, whilst those in Tama cities had improved substantively. At the local

government level, as the adoption of accrual accounting was subjective to the voluntary decision among local governors, most local governments had been implementing and preparing accrual-based financial statements and reports since 2008. The findings concluded that there was a decline in efficiency scores in special wards (19 administrative units decreased and four administrative units increased), while those in Tama cities had increased (four local units decreased and 22 local units increased). Chapter 5 presents an explanation of the driving factors affecting the decline of efficiency scores in special wards, but an increase in Tama cities.

In Chapter 5, we delve into which factors affect efficiency for 49 municipalities, for which we used a truncated regression with double bootstrapping presented by Simar and Wilson (2007). To produce bias-corrected efficiency scores, we utilized algorithm 2. The regressors selected were non-discretionary (asset utilization, expenditure, and revenue budget accuracy) and discretionary variables (taxable income, population growth, and density). Findings indicate that asset utilization has significant positive associations with efficiency scores while the budgetary accuracy of revenues and expenditures was not significantly associated with efficiency scores. This suggests that asset utilization plays an important role in improving efficiency scores but budget control has limited use in improving efficiency (instead maintaining financial stability) in Tokyo local governments.

Next, we found evidence of different effects of public assets on efficiency in special wards and Tama cities. While the sign of the effect is positive in special wards, it is negative in Tama cities. This suggests that there was a large amount of public assets invested in special wards, leading to inefficiency over 2008–2015, while asset investments in Tama cities were likely insufficient to satisfy the increasing demands created by population growth. Hence, cutting

down on assets is necessary to improve efficiency in special wards, while marginal asset investment may be required to improve efficiency in Tama cities.

In terms of the various assets utilized by local governments, we deconstructed the total assets into key categories: living infrastructure, education, welfare, environment and hygiene, industry, firefighting, and general affairs. For all 49 local governments, the results show that while utilization of education and firefighting assets was negatively associated with estimated efficiency, the living, environment, and industry assets were positively associated with efficiency. Subsequently, to enhance efficiency, policymakers and local authorities should take into consideration an increase in the amount of educational and firefighting assets and/or a reduction in the amount of living, environment, and industry assets. There is also a distinction between asset utilization among local units in special wards and Tama cities. There is a positive association between living assets and efficiency and a negative association between education, industry, and general affairs assets and efficiency in special wards. Educational and firefighting assets are negatively associated with efficiency and living, environmental, and industry assets are positively associated with efficiency scores. It is reasonable to deduce that special wards tend to spend public money on industry and infrastructure whereas Tama cities have the propensity to cover human-oriented service assets.

PMS plays an essential supporting role in measuring the performance of individuals and organizations in association with budgeting activities in the public sector. The design and implementation of PMS are aligned with strategic planning and budgeting. In Chapter 6, we explored the use of PMS for incentive-oriented and exploratory purposes, which are influenced by the latent variables of individual incentives, legislative mandates, and learning. These

relations are mediated by financial and non-financial indicators. We used mixed methods composed of PLS-SEM and semi-structured interviews for this study.

In the quantitative research (Phase 1), the results indicated that TMG has a tendency to use PMS for incentive-oriented purposes at the operational level rather than exploratory purposes at the strategic level. Non-financial metrics act as mediators in the relationship between incentives and PMS incentive use. Legislative mandates and learning do not influence the use of PMS. However, evidence suggests that performance evaluation guidelines (legislative mandates) affect the design and use of financial indicators. In the qualitative research (Phase 2), due to the inability to conduct semi-structured interviews at TMG, we alternatively interviewed some local officials at five typical Tokyo municipalities. We found that performance measurement was aligned with budgeting activities. However, the findings in Phase 1 indicated that PMS use is prone to incentive-oriented (operational level) rather than exploratory use (strategic level), whereby the cash-based budgeting system supports operational use. This suggests that in order to use performance information for exploratory purposes (it is better for allocative resources), a shift from cash budgeting to accrual budgeting is good and reasonable. The recommendation for the adoption of accrual budgeting is consistent with the proposal in Chapter 5.

7.2 Policy implications

In the study of budgetary volatility, we found that the volatility of local tax positively affects the volatility of expenditure. Hence, to smooth out the volatility of local spending, we suggest reducing the volatility of local tax income by focusing on less volatile tax, such as property tax rather than income tax, in the face of financial uncertainty (Oates, 2011; Afonso, 2017). Second, the volatility of intergovernmental grants in special wards has a negative effect on local spending volatility. This means that varied intergovernmental grants could lead to less

volatile spending. Thus, it may be inferred that local authorities should inter-transfer the number of ordinary grants to special grants (e.g., disaster counteraction), contingent on the yearly financial condition. The current rate of ordinary and special grants is 95% and 5%, respectively. If the amount of special grants allocated is higher than 5%, local governments could relieve the financial austerity in specific spending. Furthermore, the amount of ordinary grants should be set lower than 95% because these grants are used to save for unforeseen events instead of spending on service provisions. Therefore, it is reasonable to transfer the amount of ordinary grants into special grants to smooth out the volatility of local spending. More importantly, grants distributed to special wards are collected by the TMG through tax collection. Controlling grants and amending the amount of grants in pursuit of lessening the volatility of local spending justifies the raison d'etre of TMG's tax collection on behalf of the special wards.

In the study of estimated public efficiency scores, we found that the efficiency scores have tended to decrease since 2001, particularly after the introduction of accrual accounting in the public sector in Tokyo local governments. We noticed that an unnecessary amount of assets in Tokyo could be a reason for the decline in efficiency scores. Therefore, we aim to investigate assets and various determinants such as budgetary expenditure accuracy, budgetary income accuracy, and their impact on efficiency scores. We found that asset utilization positively affects public efficiency, meaning that a reduction of public assets can result in improvement in efficiency scores. Hence, public assets become an important driving factor of public efficiency. The second set of findings indicated that budgetary expenditure and income accuracy have no significant impact on public efficiency. It can be concluded that the contemporary regime of cash-based budgeting does function as a controlling and tightening mechanism for the municipality's financial condition, rather than promoting public efficiency. Our study

recommends the introduction of accrual budgeting matching with extant accrual accounting regimes in Tokyo local governments.

Asset utilization becomes an important element for improving public efficiency, but its effects are different in special wards and Tama cities. To pursue the higher efficiency, special wards should decrease their public assets but Tama cities should increase theirs in response to population growth. For special wards, it is necessary to reduce educational and industry assets and increase living assets; for Tama cities, diminishing living, environmental, and industry assets and increasing investment in educational assets can lead to higher efficiency.

In the research of performance measurement, we examined how the incentive-oriented and exploratory use of PMS was affected by individual incentives, legislative mandates, and learning, mediated by financial and non-financial metrics. The research results indicate that PMS in TMG is primarily used for incentive-oriented rather than exploratory purposes. Furthermore, non-financial metrics play a more important role than financial metrics. This suggests that TMG public servants are likely to rely on non-financial metrics to motivate them in achieving organizational objectives and goals, but legislative mandates (performance evaluation guidelines by MIC) determine the financial metrics in performance evaluation. Further semi-structured interviews indicated the important role of performance measurement in budgeting activities. The findings in phase 1 show that the incentive-oriented use of PMS at the operational level can be attributed to current cash budgeting system, which functions as planning public money within a year ahead. Therefore, if local authorities expect the efficiency improvement in strategic manner, the introduction of accrual budgeting is necessarily important. First, accrual budgeting not only matches with accrual accounting mode but also facilitates public money planned in multiple years, particularly public asset management. If doing so, PMS use becomes exploratory at the

strategic level rather than incentive purpose. Figure 24 exhibits the link between current accounting and budgeting system and PMS use in the case of accrual budgeting introduction.
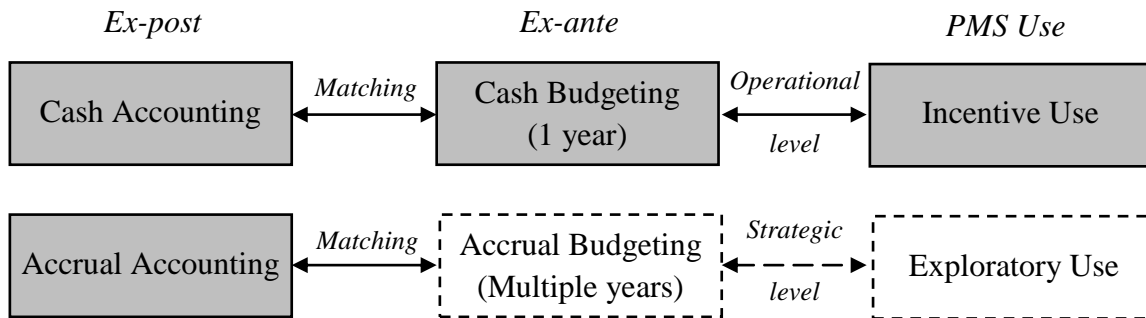
Figure 24 Research results on the nexus between budgeting system and PMS use

7.3 Limitations and future research works

Although the research provides valuable insights, it is also constrained by several limitations. In the study of budgetary volatility (Chapter 3), we examined the data frame from 2008 to 2015; however, further extension using more recent data is necessary to improve the relevance and quality and provide further insights for the research undertaken. Second, as we only focused on the FE model in Chapter 3, possibly examination of the same model with RE, mixed effects, and feasible Generalized Least Squares may be required to cross-check and improve the robustness of the results obtained.

In Chapter 4, we utilize pure technical efficiency estimation to measure the ratio of outputs to inputs, but largely neglect investigating the outcomes of public service provisions. To evaluate the outcomes, some prior literature (see Narbón-Perpiñá & De Witte, 2018) suggests using citizen satisfaction questionnaires. Furthermore, Zhu and Peyrache's study (2017) recommends incorporating quality factors as an input element to measure adjusted efficiency. These suggestions provide a clear direction for future research.

In Chapter 5, we outlined public policy implications with regard to decomposition of assets. It is essential to identify the amount of assets that should be decreased or increased to pursue higher efficiency. In earlier literature, Tone (2001) proposed the Slack-Based Model (SBM) to access the slack amount, which is potential for our research on public assets. In future research, we can extend our study by adopting the SBM in measuring the slack of certain public assets in special wards and Tama cities.

In Chapter 6, we faced difficulties in data collection. While TMG was willing to aid us in the questionnaire survey, it hesitated to participate in the interviews; public officials at the local level were ready to take part in our semi-structured interviews but were reluctant to answer questionnaires. In the future, to make our data equivalent and more consistent in research interpretations, we will need to increase our efforts in data collection from either interviews at TMG or questionnaire surveys from the local governments. Doing so could increase the reliability of our research results.

In conclusion, our research results and public policy implications could be beneficial and valuable to both local authorities and policymakers alike. However, these findings and suggestions should be carefully supplemented by further confirmation through extensive empirical studies. In addition, we recognize the drawbacks and, by resolving them, we could provide a more holistic view of improving TMG's efficiency and performance. Finally, our recommendations based on these findings for TMG and its local governments also provide an important indicator of the future of public sector management. However, our research is only a starting point for the journey in public sector reforms.

REFERENCE

Afonso, A., & Fernandes, S. (2006). Measuring local government spending efficiency: Evidence for the Lisbon region. *Regional Studies*, *40*(1), 39-53.

Afonso, A., & Furceri, D. (2010). Government Size, Composition, Volatility and Economic Growth. *European Journal of Political Econ*omy, *26*, 517-532.

Afonso, W. B. (2017). Revenue Portfolio and Expenditures: An Examination of the Volatility of Tax Revenue and Expenditure Patterns during the Great Recession. *International Journal of Public Administration*, *40*(10), 896-905.

Akizuki, K. (2001). Partnership in Controlled Decentralization: Local Governments and the Ministry of Home Affairs. In Michio Muramatsu, Farrukh Iqbal, & Ikuo Kume (Eds.), *Local Government Development in Post-War Japan*, (63-68). Oxford University Press.

Albuquerque, B. (2011). Fiscal Institutions and Public Spending Volatility in Europe. *Economic Modelling*, *28*, 2544-2559.

Andrews, R., & Entwitsle, T. (2014). *Public Service Efficiency: Reframing the debate*. Routledge Critical Studies in Public Management.

Aoki, I. (2008). Decentralization and Intergovernmental Finance in Japan. *PRI Discussion Paper Series*, 04(8).

Badunenko, O., & Tauchmann, H. (2018). Simar and Wilson two-stage effieciency analysis for Stata. *FAU Discussion Papers in Economics, No. 08/2018*, Friedrich Alexander-Universität Erlangen-Nürnberg, Institute for Economics, Erlangen.

Baiman, S. (1990). Agency research in managerial accounting: A second look. Accounting. *Organizations and Society*, 15(4), 341–371.

Baily, S. J. & Connolly, S. (1998). The flypaper effect: Identifying areas for future research. *Public Choice*, 95, 335-361.

Balaguer-Coll, M. T., Prior, D., & Tortosa-Ausina, E. (2013). Output complexity, environmental conditions, and the efficiency of municipalities. *Journal of Productivity Analysis*, *39*(3), 303–324.

Balaguer-Coll, M. T., Prior, D. & Tortosa-Ausina, E. (2007). On the determinants of local government performance: A two-stage nonparametric approach. *European Economic Review*, 51, 425-451.

Balaguer-Coll, M. T. & Prior, D. (2009). Short- and long-term evaluation of efficiency and quality. An application to Spanish municipalities. *Applied Economics,* 41, 2991-3002.

Balaguer-Coll, M. T., Brun-Martos, M. I., Márquez-Ramos, & Prior, D. (2018). Local government efficiency: determinants and spartial interdependence. *Applied Economics,*

51(14), 1478-1494.

Baltagi, B. H. (2005). *Econometric analysis of panel data*. John Wiley & Sons, New Jersey, USA.

Bandy, G. (2015). *Financial Management and Accounting in the Public Sector*. Routledge Masters in Public Managers, Routledge.

Banker, R. D., & Natarajan, R. (2008). Evaluating Contextual Variables Affecting Productivity Using Data Envelopment Analysis. *Operations Research*, 56(1), 48–58.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30(9), 1078–1092.

Bartolini, D., Sacchi, A., Salotti, S., & Santolini, R. (2017). Fiscal Decentralization in Times of Financial Crises. *CESifo Economic Studies*.

Bessho, S., & Ogawa, H. (2015). Fiscal Adjustment in Japanese Municipalities. *Journal of Comparative Economics*, 43(4), 1053-1068.

Bird, R. M., & Vaillancourt, F. (2009). Fiscal decentralization in developing countries: an overview. In Bird, R.M. & Vaillancourt, F. (Eds.), *Fiscal Decentralization in Developing Countries*, Cambridge University Press.

Blöndal, J. (2004). Issues in accrual budgeting. *OECD Journal on Budgeting*, 7(4), 1–12.

Bonner, S. E., & Sprinkle, G. B. (2002). The effects of monetary incentives on effort and task performance: Theories, evidence, and a framework for research. *Accounting, Organizations and Society*, 27(4–5), 303–345.

Bonner, S. E., Hastie, R., Sprinkle, G. B., & Young, S. M. (2000). A Review of the Effects of Financial Incentives on Performance in Laboratory Tasks: Implications for Management Accounting. *Journal of Management Accounting Research*, 12(1), 19–64.

Bourkaert, G., & Van Dooren, W. (2002). Progress/regress in performance measurement systems. *IRSPM VI Edinburgh*, VI, 1–28.

Boyne, G. A. (2002). Public and private management: what's the difference?. *Journal of Management Studies*, 39(1), 97–122.

Brignall, S., & Modell, S. (2000). An institutional perspective on performance measurement and management in the "new public sector." *Management Accounting Research*, 11(3), 281–306.

Broadbent, J., & Laughlin, R. (2009). Performance management systems: A conceptual model. *Management Accounting Research*, 20(4), 283–295.

Bureau of Finance (2014). *Tokyo Metropolitan Government*. Available at: http://www.zaimu.metro.tokyo.jp/bond/en/ir_library/ir/ir_document/ir_document2014.pdf.

Bureau of Taxation (2017). *Guide to Metropolitan Taxes 2017*. Available at:

http://www.tax.metro.tokyo.jp/.

Cabinet Office (2017). Basic Policy on Economic and Fiscal Management and Reform 2017: Increasing productivity through investment in human resources. Retrieve at: https://www5.cao.go.jp/keizai/index-e.html.

Cassete, A. & Paty, S. (2010). Fiscal decentralization and the size of government: a European country empirical analysis. *Public Choice*, 143, 173-189.

Carlin, T. (2005). Debating the impact of accrual accounting and reporting in the public sector. *Financial Accountability & Management*, 21, 309–337.

Cavalluzzo, K. S., & Ittner, C. D. (2004). Implementing performance measurement innovations: Evidence from the government. *Accounting, Organizations and Society*, 29(3–4), 243–267.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2(6), 429–444.

Chenhall, R. H. (2007). Theorizing Contingencies in Management Control Systems Research. In Chapman, Hopwood, and Shield (Eds.), *Handbook of Management Accounting Research*, Volume 1, Elsevier.

Chin, W., & Newsted, P. (1999). Structural equation modelling analysis with small samples using partial least squares. In: Hoyle, R. (Ed.), *Statistical strategies for small sample research.* Thousand Oaks, CA: Sage Publications, 307-41.

Chin, W. W. (1998). Commentary: Issues and Opinion on Structural Equation Modeling. *MIS Quarterly*, 22(1).

Chin, W. W. (2010). How to Write Up and Report PLS Analyses. In Vinzi, V.E., Chin, W.W., Henseler, J., Wang, H. (Eds.) *Handbook of Partial Least Squares: Concepts, Methods and Applications* (665-690). Springer

Cooper, D. R., & Schindler, P. S. (2014). *Business research methods*. McGraw-Hill Education

Cordero, J. M., Pedraja-Chaparro, F., Pisaflores, E. C., & Polo, C. (2017). Efficiency assessment of Portuguese municipalities using a conditional nonparametric approach. *Journal of Productivity Analysis*, 48(1), 1–24.

Council of Local Authorities for International Relations (CLAIR) (2013). *Local Government in Japan 2012*. Available at: http://www.clair.or.jp/j/forum/pub/pdf/jichi12_en.pdf.

Craig, F. (1999). State and local government performance : It's time to measure up!. *The Government Accountants Journal*, 48(1), 40–46.

Crain, W. M. (2003). *Volatile States*. Ann Arbour, MI: University of Michigan Press.

Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage. CA: Thousand Oaks.

Cronbach, L. J., & Meehl, P. E. (1955). Psychological Bulletin. *Freud's Psychology*, 52(4), 281–302. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19586159

Cuadrado-Ballesteros, B., García-Sánchez, I. M., & Prado-Lorenzo, J. M. (2013). Effect of modes of public services delivery on the efficiency of local governments: A two-stage approach. *Utilities Policy*, 26, 23–35.

Cuganesan, S., Guthrie, J., & Vranic, V. (2014). The Riskiness of Public Sector Performance Measurement: A Review and Research Agenda. *Financial Accountability & Management*, 30(3), 279–302.

Cruz, N. F. & Marques, R. C. (2014). Revisiting the determinants of local government performance. *Omega*, 44, 91-103.

De Borger, B. & Kertens, K. (1996). Cost efficiency of Belgian local governments: An comparative analysis of FDH, DEA, and econometric approaches. *Regional Science and Urban Economics*, 26, 145-170.

De Bruijn, H. (2002). Performance measurement in the public sector: Strategies to cope with the risks of performance measurement. *International Journal of Public Sector Management*, 15(7), 578-594.

De Bruijn, H. (2007). *Managing Performance in the Public Sector*. New York, NY: Routledge.

Denison, D. V., & Guo, Z. (2015). Local Government Debt Management and Budget Stabilization. In Hou, Y. (eds.), *Local Government Budget Stabilization*. Springer, Switzerland; 121-139.

DiMaggio, P. J., & Powell, W. W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, 48(2), 147–160.

Doi, T. (1996). Nihon no Toshi Zaisei niokeru Flypaper Koka. *Financial Review* 40 (September), 95-119.

Dooren, W. V., Bouckaert, G., & Halligan, J. (2015). *Performance Management in Public Sector*. Routledge.

Doumpos, M., & Cohen, S. (2014). Applying data envelopment analysis on accounting data to assess and optimize the efficiency of Greek local governments. *Omega* (United Kingdom), 46, 74–85.

Drew, J. Kortt, M. A., & Dollery, B. (2016). Did the Big Stick Work? An Empirical Assessment of Scale Economies and the Queensland Forced Amalgamation Program. *Local Government Studies*, 42 (1), 1-14.

Drew, J., & Dollery, B. (2016). Does Size Still Matter? An Empirical Analysis of the

Effectiveness of Victorian Local Authorities. *Local Government Stu*dies, 42(1), 15–28.

Drew, J., Dollery, B., & Kortt, M. A. (2016). Peas in a Pod: Are Efficient Municipalities also Financially Sustainable? *Australian Accounting Review*, 26(2), 122–131.

Drew, J., Kortt, M. A., & Dollery, B. (2017). No Aladdin's Cave in New South Wales? Local Government Amalgamation, Scale Economies, and Data Envelopment Analysis Specification. *Administration and Society*, 49(10), 1450–1470.

Drew, J., Kortt, M. A., & Dollery, B. (2015). What determines Efficiency in Local Government? A DEA Analysis of NSW Local Government. *Economic Papers*, 34(4): 243-256.

Dunleavy, P., & Hood, C. (1994). From old public administration to new public management. *Public Money and Management*, 14(3), 9–16.

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). New public management is dead - Long live digital-era governance. *Journal of Public Administration Research and Theory*, 16(3), 467–494.

Easterby-Smith, M., Thorpe, R., & Jackson, P. (2012). *Management Research*. Sage, Thousand Oaks, CA.

Eshima, Y., Katayama, T., & Ohno, T. (2001). Public Management Innovation in Japan: Its Characteristics and Challenges. *International Review of Administrative Sciences*, 67(4), 699–714.

Eyraud, L., & Lusinyan, L. (2011). Decentralizing Spending More than Revenue: Does It Hurt Fiscal Performance*?. IMF Working Paper No. 11/226*.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society*, 120(3), 253–290.

Ferreira, A., & Otley, D. (2009). The design and use of performance management systems: An extended framework for analysis. *Management Accounting Research*, 20(4), 263–282.

Fogarty, J., & Mugera, A. (2013). Local Government Efficiency : Evidence from Western Australia. *The Australian Economic Review*, 46(3), 300–311.

Folan, P., & Browne, J. (2005). A review of performance measurement: Towards performance management. *Computers in Industry*, 56(7), 663–680.

Foremy, D. (2014). Sub-national Deficits in European Countries: The Impact of Fiscal Rules and Tax Autonomy. *European Journal of Political Economy*, 34, 86-110.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39-50.

Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., … Neely, A. (2007). Towards a definition of a business performance measurement system. International

Journal of Operations and Production Management, 27(8), 784–801.

Franco-Santos, M., Lucianetti, L., & Bourne, M. (2012). Contemporary performance measurement systems: A review of their consequences and a framework for research. *Management Accounting Research*, 23(2), 79–119.

Frey, B. S., Homberg, F., & Osterloh, M. (2013). Organizational Control Systems and Pay-for-Performance in the Public Service. *Organization Studies*, 34(7), 949–972.

Fried, H. O., Lovell, C. A. K., & Schmidt, S. S. (2008). *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press.

Fryer, K., Antony, J., & Ogden, S. (2009). Performance Management in the Public Sector. *International Journal of Public Sector Management*, 22(6), 478–498.

Fukuyama, H. et al. (2017). Dynamic Network Efficiency of Japanese Network. In Tone, K. (Eds.), *Advances in DEA theory and applications: With extensions to the forecasting model*. Oxford: Wiley.

Furceri, D. & Poplawski-Ribeiro (2008). Government Spending Volatility and the Size of Nations. *Organisation for Economic Co-operation and Development (OECD) Working Paper No. 924*, Paris.

Furceri, D. (2007). Is Government Expenditure Volatility Harmful for Growth? A Cross-country Analysis. *Fiscal Studies*, 28, 103-120.

Furceri, D., Sacchi, A., & Salotti, S. (2016). Can Fiscal Decentralization Alleviate Government Consumption Volatility?. *Open Economies Review*, 27(4), 611-636.

Grant, B., & Drew, J. (2017). *Local Government in Australia: History, Theory and Public Policy*. Springer Nature, Singapore.

Gruening, G. (2001). Origin and Theoretical Basis of New Public Management. *International Public Management Journal*, 4, 1–25.

Guthrie, J. (1998). Application of accrual accounting in the Australian public sector - rhetoric or reality. *Financial Accountability & Management*, 14(1), 1–19.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Sage.

Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2009). *Multivariate Data Analysis*. Pearson.

Haneda, S., Hashimoto, A., & Tsuneyoshi, T. (2012). Evaluating administrative efficiency change in the post-merger period: A study on Ibaraki prefectures (1979-2004). *International Regional Science Review*, 35(2), 237-262.

Hansen, S. C., & Van der Stede, W. A. (2004). Multiple facets of budgeting: An exploratory

analysis. *Management Accounting Research*, 15(4), 415–439.

Hattori, T., & Miyake, H. (2015). Empirical analysis of yield determinants in Japan's municipal bond market: Does credit risk premium exist? *MPRA Paper*.

Henri, J. F. (2006). Organizational culture and performance measurement systems. *Accounting, Organizations and Society*, 31(1), 77–103.

Hernandez, D. (2002). Local government performance measurement. *Public Management,* 84, 10–11.

Hood, C. (1991). A public management for all seasons?. *Public Administration*, 69, 3–19.

Hood, C. (1995). The "new public management" in the 1980s: Variations on a theme. *Accounting, Organizations and Society*, 20(2–3), 93–109.

Hood, C. (2012). Public management by numbers as a performance-enhancing drug: Two hypotheses. *Public Administration Review*, 72(1), 85–92.

Hori, M. (2004). Japanese Public Administration and its Adaptation to New Public Management. *Ritsumeikan Law Review*. Retrieved from http://ci.nii.ac.jp/naid/110001838535/

Ittner, C. D., Larcker, D. F., & Meyer, M. W. (2003). Subjectivity and the Weighting of Performance Measures : Evidence from a Balanced Scorecard. *The Accounting Review*, 78(3), 725–758.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice. *Field Methods*, 18(1), 3–20.

Johnsen, Å. (2005). What does 25 years of experience tell us about the state of performance measurement in public policy and management?. *Public Money and Management*, 25(1), 9–17.

Kaganova, O., & Kopanyi, M. (2014). Managing Local Assets. In Farvacque-Vitkovic, C. & Kopanyi, M. (Eds.), *Municipal finances: A handbook for local government*, World Bank, Washington, D.C.

Kaganova, O., & Nayyar-Stone, R. (2000). Municipal real property asset management : An overview of world experience, trends and financial implications. *Journal of Real Estate Portfolio Management*, 6(4),307-326.

Kaganova, O. & Telgarsky, J. (2017). Management of Capital Assets by local governments: An Assessment and benchmarking surveys. *International Journal of Strategic Property Management, 22*(2), 143-156.

Kalb, A. (2010). The impact of intergovernmental grants on cost efficiency: Theory and evidence from Germany municipalities. *Economic Analysis and Policy*, 40(1), 23-48.

Kalb, A. (2014). What determines local governments' cost-efficiency? The case of road

maintenance. *Regional Studies*, 48 (9), 1483-1498.

Kalb, A., Geys, B., & Heinemann, F. (2012). Value for money? German local government efficiency in a comparative perspective. *Applied Economics*, 44, 201-218.

Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard − Measures that drives performance. *Harvard Business Review*, January-February issue.

Kobayashi, M., Yamamoto, K., & Ishikawa, K. (2016). The Usefulness of Accrual Information in Non-mandatory Environments: The Case of Japanese Local Government. *Australian Accounting Review*, 26(2), 153–161.

Kudo, H. (2003). Between the "governance" model and the policy evaluation act: New public management in Japan. *International Review of Administrative Sciences*, 69(4), 483–504.

Lampe, H. W., Hilgers, D., & Ihl, C. (2015). Does accrual accounting improve municipalities' efficiency? Evidence from Germany. *Applied Economics*, 41, 4349-4363.

Lau, C. M., & Roopnarain, K. (2014). The effects of nonfinancial and financial measures on employee motivation to participate in target setting. *British Accounting Review*, 46(3), 228–247.

Lebas, M. J. (1995). Performance measurement and performance management. International *Journal of Production Economics*, 41, 23–35.

Lee, L., Petter, S., Fayard, D., & Robinson, S. (2011). On the use of partial least squares path modeling in accounting research. *International Journal of Accounting Information Systems*, 12(4), 305–328.

Litvack, J., Ahmad, J., & Bird, R. (1999). Fiscal Decentralization in Developing Countries. *Sector Studies Series*.

Local Government Finance in Japan (2014). Naohiko Jinno, Sachio Konishi.

Martí, C. (2013). Performance budgeting and accrual budgeting: A study of the United Kingdom, Australia, and New Zealand. *Public Performance & Management Review*, 37(1), 33-58.

Martinez-Vazquez, J., & McNab, R. M. (2003). Fiscal decentralization and economic growth. *World Development*, 31(9), 1597–1616.

Martin-Rodriguez, M. & Ogawa, H. (2017). The Empirics of the Municipal Fiscal Adjustment. *Journal of Economic Surveys*, 31(3), 831-853.

Mclure, C. E., & Martinez-Vazquez, J. (2000). The Assignment of Revenues and Expenditures in Intergovernmental Fiscal Relations. *In Paper prepared for the core course on Intergovernmental Relations and Local Financial Management (40). Washington, DC: World Bank*.

Melnyk, S. A., Bititci, U., Platts, K., Tobias, J., & Andersen, B. (2014). Is performance

measurement and management fit for the future?. *Management Accounting Research*, 25(2), 173–186.

Ministry of Internal Affairs and Communication (MIC) (2017). *FY2015 Settlement*. White Paper on Local Public Finance.

Ministry of Internal Affairs and Communications (MIC) (2012). The situation in the local governments of the preparation of the financial statements for 2010. *MIC Press,1-9.*

Miyara, I. & Fukushige, M. (2001). Nihon no Shicyoson Zaise niokeru Flypaper Koka. Nihhon Keizai Kenkyu, 42(1), 44-61.

Mochida, N. (2006). Local Government Organization and Finance: Japan. In Shah, A. (Ed.), *Local Governance in Industrial Countries*. Washington, DC: World Bank.

Mochida, N. (2008). *Fiscal Decentralization and Local Public Finance in Japan*. Routledge, New York.

Mochida, N. (2008). Subnational Borrowing in Japan : from ' Implicit Guarantee ' to Market Discipline and Fiscal Rule Subnational Borrowing in Japan : from ' implicit guarantee ' to market discipline and fiscal rule. *CIRJE Discussion Papers.*

Mochida, N. (2015). Fiscal relations across levels of government and sub-national fiscal rules in Japan. In Kim, J.H., & Blochliger, H. (Eds.), *Institutions of Intergovernmental Fiscal Relations:* Challenges Ahead. OECD Fiscal Federalism Studies.

Modell, S. (2004). Performance Measurement Myths in the Public Sector: A Research Note. *Financial Accountability & Management*, 20(1), 39–55.

Modell, S. (2009). Institutional Research on Performance Measurement and Management in the Public Sector Accounting Literature: A Review and Assessment. *Financial Accountability & Management*, 25(3), 277–303.

Moynihan, D. P. (2006). Managing for results in state government: Evaluating a decade of reform. *Public Administration Review*, 66(1), 77–89.

Moynihan, D. P. (2008). *The Dynamics of Performance Management: Constructing Information and Reform*. Georgetown University Press, Washington, D.C.

Nakazawa, K. (2013). Cost inefficiency of municipalities after amalgamation. *Procedia Economics and Finance*, 5, 581-588.

Nakazawa, K. (2014). Does the method of amalgamation affect cost inefficiency of the new municipalities? *Open Journal of Applied Sciences*, 4(4), 143-154.

Narbón-Perpiñá, I., & De Witte, K. (2018). Local governments' efficiency: a systematic literature review—part I. *International Transactions in Operational Research*, 25(2), 431-468.

Neely, A., Gregory, M., Platts, K., Franco-Santos, M., Kennerley, M., Micheli, P., Platts, K.

(1995). Performance measurement system design: A literature review and research agenda. *International Journal of Operations & Production Management*, 15(4), 80–116.

Newberry, S., & Pallot, J. (2004). Freedom or coercion? NPM incentives in New Zealand central government departments. *Management Accounting Research*, 15(3), 247–266.

Nijkamp, P., & Suzuki, S. (2009). A Generalized Goals-achievement Model in Data Envelopment Analysis: an Application to Efficiency Improvement in Local Government Finance in Japan. *Spatial Economic Analysis*, 4(3), 249–274.

Nitzl, C. (2016). The use of partial least squares structural equation modelling (PLS-SEM) in management accounting research: Directions for future theory development. *Journal of Accounting Literature*, 37, 19–35.

NSW government (2013). *Integrated Planning and Reporting Manual for local government in NSW: Planning a sustainable future*. Premier & Cabinet, Division of Local Government.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory*. McGraw-Hill, New York.

Oates, W. E. (1972). *Fiscal Federalism*. NY: Harcourt Brace Jovanovich.

Oates, W. E. (1999). Essay on Fiscal Federalism. *Journal of Economic Literature*, 37(3), 1120–1149.

Oates, W. E. (2008). On The Evolution of Fiscal Federalism: Theory and Institutions. *National Tax Journal*, LXI(2).

Oates, W. E. (2011). *Financing the New Federalism*. New York: RFF Press.

Ohsugi, S. (2009). The Organization of Local Government Administration in Japan. *Papers on the Local Governance System and Its Implementation*, (11).

Ohsugi, S. (2011). The Large City System of Japan. *Papers on the Local Governance System and its Implementation in Selected Fields in Japan No.20*.

Olson, M. (1969). The Principle of "Fiscal Equivalence": The Division of Responsibilities among Different Levels of Government. *The American Economic View*, 59(2).

Otley, D. (1999). Performance management : a framework for management control systems research. *Management Accounting Research*, 10(4), 363–382.

Otsuka, A., Goto, M., & Sueyoshi, T. (2014). Cost-efficiency of Japanese local governments: effects of decentralization and regional integration. *Regional Studies, Regional Science,* 40(4), 649-669.

Owaga, H., & Tanahashi, K. (2008). Effect of New Public Management: Data Envelopment Analysis. *Government Auditing Review*, 15, 47-62.

Phelps, A. (2011). Municipal property asset management – a comparative study of UK and Russia. *International Journal of Strategic Property Management*, 15(4), 416–437.

Pidd, M. (2012). *Measuring the performance of public services*. Cambridge University Press.

Pollitt, C. (2016). Managerialism Redux?. *Financial Accountability & Management*, 32(4), 429–447.

Qian, Y., & Weingast, B. R. (1997). Federalism as a Commitment to Preserving Market Incentives. *Journal of Economic Perspectives*, 11(4), 83–92.

Radnor, Z. J., & Barnes, D. (2007). Historical analysis of performance measurement and management in operations management. *International Journal of Productivity and Performance Management*, 56(5/6), 384–396.

Reichard, C., & Helden, J. van. (2017). Why cash-based budgeting still prevails in an era of accrual-based reporting in the public sector. *Accounting, Finance & Governance Review*, 23, 1–2.

Robinson, M. (2009). Accrual budgeting and fiscal policy. *OECD Journal on Budgeting*, 1, 75–103.

Robinson, M. (2013). Performance Budgeting. In Allen, R., Hemming, R., and Potter, B. H. (Eds.), *The International Handbook of Public Financial Management,* Palgrave Handbooks.

Robinson, M. (2016). Budget reform before and after the global financial crisis. *OECD Journal on Budgeting*, 16(1), 29–63.

Sacchi, A., & Salotti, S. (2016). A Comprehensive Analysis of Expenditure Decentralization and of the Composition of Local Public Spending. *Regional Studies*, 50(1), 93-109.

Sacchi, A., & Salotti, S. (2017). The Influence of Decentralized Taxes and Intergovernmental Grants on Local Spending Volatility. *Regional Studies*, 51(4), 507-522.

Schick, A. (2007). Performance budgeting and accrual budgeting: Decision rules or analytic tools?. *OECD Journal on Budgeting*, 7(2), 109-138.

Schick, A. (2013). Reflections on Two Decades of Public Financial Management Reforms. In Cangiano, M., Curristine, T., and Lazare, M. (eds), *Public Financial Management and Its Emerging Architecture*, International Monetary Fund, Washington.

Shah, A. (2007). Introduction : Principles of Fiscal Federalism. In A. Shah (Ed.), *The Practice of Fiscal Federalism: Comparative Perspectives*. McGill-Queen's University Press.

Shirai, S. (2006). The Impact of Intergovernmental Transfers in the Japanese Local Government Fiscal System. Asian Economic Paper 4(2), 26-58.

Sickles, R. C. & Zelenyk, V. (2019). *Measurement of Productivity and Efficieny: Theory and Practice*. Cambridge University Press.

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

Simons, R. (1995). Levers of Control: How Managers Use Innovative Control Systems to Drive Strategic Renewal. *Harvard Business School Press*.

Simons, R. (2000). *Performance Measurement and Control Systems for Implementing Strategy: Text and Cases*. Prentice-Hall, Inc, New Jersey.

Speklé, R. F., & Verbeeten, F. H. M. (2014). The use of performance measurement systems in the public sector: Effects on performance. *Management Accounting Research*, 25(2), 131–146.

Staley, T. (2015). The Effect of TELs on State Revenue Volatility: Evidence From the American States. *Public Budgeting & Finance*, 35(1), 29-48.

Staley, T. (2017). The Impact of Fiscal Limits on State Revenue Volatility. *American Review of Public Administration,* 47(4), 443-454.

Steccolini, I., Jones, M., & I. Saliterer (2017). Government Financial Resilience: International Perspectives on How Local Governments Face Austerity. Emerald Publishing: Bingley.

Stede, V. Der, Wim, A., Chow, C. W., & Lin, T. W. (2006). Strategy, choice of performance measures, and performance. *Behavioral Research in Accounting,* 18(2002), 185–205.

Storto, C. (2013). Evaluating technical efficiency of Italian major municipalities: a Data Envelopment Analysis model. *Procedia - Social and Behavioral Sciences*, 81, 346-350.

Storto, C. (2016). The trade-off between cost efficiency and public service quality: A non-parametric frontier analysis of Italian major municipalities. *Cities*, 51, 52-63.

Talbot, C. (2006). Performance regimes and institutional contexts: Comparing Japan, UK and USA. *International Symposium on Policy Evaluation*, (June), 1–16.

Tanaka, H. (2009). Evaluation in Local Governments in Japan. *Papers on the Local Governance System and Its Implementation in Selected Fields in Japan, (14).*

Tanaka, Y. (2011). Local Bonds in Japan. *Papers on the Local Governance System and its Implementation in Selected Fields in Japan, (19).*

Tanzi, V. (1995). Fiscal Federalism and Decentralization: A Review of Some Efficiency and Macroeconomic Aspects. In M. Bruno & B. Pleskovic (Eds.), *Annual World Bank Conference on Development Economics*. Washington, DC.

Tanzi, V. (2003). Pitfalls on the Road to Fiscal Decentralization. *Carnegie Endowment Working Paper.*

Tanzi, V., & Prakash, T. (2000). The cost of government and the misuse of public assets. *IMF Working Paper.*

Tashakkori, A., & Teddlie, C. (2010). *SAGE Handbook of Mixed Methods in Social & Behavioral* Research. Sage.

Ter Bogt, H., Budding, T., Groot, T., & Van Helden, J. (2010). Current NPM Research: Digging

Deeper and Looking Further. *Financial Accountability & Management,* 26(3), 241–245.

Tessier, S., & Otley, D. (2012). A conceptual development of Simons' Levers of Control framework. *Management Accounting Research*, 23(3), 171–185.

Thompson, F., & Gates, B. L. (2007). Betting on the Future with a Cloudy Crystal Ball? How Financial Theory Can Improve Revenue Forecasting and Budgets in the States. *Public Administration Review,* 67(5), 825-836.

Tiebout, C. M. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy,* 64(5), 416–424.

Tokyo Metropolitan Government (TMG) (2018). City Profile and Government. Available at: http://www.metro.tokyo.jp/english/about/index.html.

Tone, K. (2001). A slack-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research, 130*(3), 498-509.

Tran, T. V., Drew, J., & Noguchi, M. (2018). The Role of Revenue Volatility in Local Expenditure Volatility: A Comparison of Tokyo Metropolitan Local Governments. *Economic Papers,* 1–13.

Van de Ven, A. H., & Ferry, D. L. (1980). *Measuring and Assessing Organizations*. John Wiley & Sons.

Van der Hoek, M. P. (2005). From cash to accrual budgeting and accounting in the public sector: The Dutch experience. *Public Budgeting and Finance*, 3(Spring), 32–45.

Van Dooren, W., Bourkaert, G., & Halligan, J. (2015). *Performance Management in the Public Sector*. Routledge Masters in Public Management. Routledge.

Van Thiel, S. (2014). *Research methods in public administration and public management: An introduction.* Routledge, London and New York

Verbeeten, F. H. M. (2005). 'New' Performance Measures : Determinants of Their Use and Their Impact on Performance REPORT SERIES. *ERIM Report Series Research in Management, (ERS-2005-054-F&A).*

Verbeeten, F. H. M. (2008). Performance management practices in public sector organizations. *Accounting, Auditing & Accountability Journal*, 21(3), 427-454.

Verbeeten, F. H. M., & Speklé, R. F. (2015). Management Control, Results-Oriented Culture and Public Sector Performance: Empirical Evidence on New Public Management. *Organization Studies,* 36(7), 953–978.

Vo, D.H. (2010). The Economics of Fiscal Decentralization. *Journal of Economic Surveys*, 24(4), 657-679.

Walker, R. M., Boyne, G. A., & Brewer, G. A. (2010). *Public Management and Performance:*

Research Directions. Cambridge University Press.

Warren, K. (2015). Time to look again at accrual budgeting. *OECD Journal on Budgeting,* 14(3), 1–17.

Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach.* Cengage Learning.

Worthington, A. C, & Dollery, B. E. (1999). Fiscal illusion and the Australian local government grants process: How sticky is the flypaper effect?. *Public Choice*, 99, 1-13.

Worthington, A. C. & Dollery, B. E. (2001). Measuring efficiency in local government: An analysis of New South Wales municipalities' domestic waste management function. *Policy Studies Journal*, 29(2), 232–249.

Worthington, A. C. & Dollery, B. E. (2001). Measuring efficiency in local government: An analysis of New South Wales municipalities' domestic waste management function. *Policy Studies Journal*, 29(2), 232-249.

Worthington, A. C. (2000). Cost efficiency in Australian local government: A comparative analysis of mathematical programming and econometric approaches. *Financial Accountability & Management*, *16* (3), 0267-4424.

Xue, M., & Harker, P. (1999), Overcoming the Inherent Dependency of DEA Efficiency Scoring: A Bootstrap Approach. *Working paper*, Wharton Financial Institutions Center, Philadelphia.

Yamamoto, H. (2003). New Public Management − Japan's Practice. *Institute for International Policy Studies,* (January), 1–26.

Yamamoto, K., & Noguchi, M. (2013), Different scenarios for accounting reform in non-Anglophone contexts: The case of Japanese local governments since the 1990s. *Accounting History*, 18(4), 529–549.

Yin, R. (2009). *Case Study Research: Design and methods*. Thousand Oaks, CA: Sage.

Zhu, M., & Peyrache, A. (2017). The quality and efficiency of public service delivery in the UK and China. *Regional Studies*, 51(2), 285–296.