

The Measurement of Operating Results in the Public Sector: a Systematic Reevaluation and Academic Interpretation of Performance Aspects

Co Nguyen Trong^a, Huy Pham Quang^{b*}, ^aAcademy of Finance, No 58 Le Van Hien Road, Duc Thang Ward, Bac Tu Liem District, Hanoi, Vietnam, ^bAccounting and Auditing Faculty University of Economics Ho Chi Minh City, Vietnam, Email: ^{b*}pqh.huy@gmail.com

With the recognition of the limitation in the literature of public sector accounting in the world as well as in Vietnam, this research aims to identify and analyse the emerging literature on the performance measurement (PM) in the public sector (PS) and shed a light for other research in the future. The systematic literature review (SLR) methodology proposed by Denyer and Tranfield (2009) was applied with a total of 80 papers published in 38 scientific journals for evaluating and establishing current trends. Apart from indicating an increasing demand on a better understanding in PM of PS despite regionally heterogeneous, vision and mission for improving performance, the results also revealed gaps stemming from choice of number as well as the type of keywords and journals through seven categories in those research trends. Taking these results into account, academics can keep on exploring this type of subject area on their own topic, while practitioners and policy makers can employ in their empirical work.

Key words: *Performance measurement, public management, public sector, systematic literature review, Accounting public, Accounting Public Sector.*

JEL Classification: *H83-M41-M48*

Introduction

All the PS entities are nowadays functioning in the dynamical environment that imposes the organisations to align with the consecutively variable environmental conditions. In order to measure the performance of public sector organisations (PSOs), the application of suitable PM methods is worth an important attention (Balabonienė & Večerskienė, 2015). For this reason, interest in PM, in recent years, has grown exponentially and has been capturing growing attention in PSOs, international audience, and substantial academic researchers (Alach, 2016).

It is obvious that there is abundant research carried out toward the PM in the PS whose subjects conventionally focused on enumerating the performance measurement system (PMS) design and implementation (e.g. Langham & Paulsen, 2018), which significantly affect PM in PS (e.g. Barbato & Turri, 2017), the improvement of PM in the PS (Buathong & Bangchokdee, 2017). Additionally, the sparse extent has been studies on specific topics about PMS, which based predominantly on enumerating the different situation of the research context. Owing to the mentioned above, a systematisation of knowledge regarding PSPM seems to be hard to put forward.

In correspondence with this requirement, our study has made the inductive examination through SLR approach to determine what divergent researchers implied when they mentioned PM in the PS. Our contribution is to delineate how previous affairs have conducted the topic under inspection through employing a multidimensional configuration to categorise the studies, synthesising the principal themes inquired until now and resolving new research questions. As a consequence, the value of our study depends on galvanising an alternative from primarily descriptive research into the advancement and integration in studies on PMPS. Hence, on a better degree, this study has set a goal for addressing these ensuing research questions:

RQ1. How is the PMPS literature developing during the period of time?

RQ2. What is focused on within the PMPS literature?

RQ3. What are the implications to shed light for future research?

In order to demystify our consideration, the rest of this paper is configured as followed. The primary theoretical inputs in this study were introduced in Section 2. This is pursued by a discussion of the methods employed in our study depicting at length how our systematic review was performed, and the systematic process was applied in the literature analysis. Our main affair is accentuated in Section 4 in which we provide results and comments on the gaps to shed light for further research on PSPM. The final section draws some concluding thoughts and implications.

Literature Review and Methodology

Literature Review

According to Kaplan (1990), PM is the prerequisite of performance improvement. On the other hand, it can be defined as “the process of quantifying the efficiency and effectiveness of action” (Neely et al., 2005). By the 1980s, owing to the increasingly competitive markets and limitation of sole financial measures, organisations had to respond more to external activities and long-term success so as to satisfy customers’ increasing demands and maintain market share (Parker, 2000). However, due to the lack of focus on long-term success and insight into how an organisation’s performance is achieved and improved (Kaplan, 1986) as well as failing to indicate a set of intangible critical factors (e.g., customer satisfaction and strategy (Parker, 2000), a plethora of general PMSs had been devised over the last 20 years.

In the mid-1930s, Ridley and Simon (1938) stressed the importance of measurement, discussed what, why, and how municipal activities should be measured and pointed out steps for improvement of the PM. However, the research on PM has not gained more than modest attention until the last 2–3 decades (Streib & Poister, 1999). Starting from the New Public Management (NPM) movement, the government was expected to show its efficiency in expending public resources as well as to prove that outcomes related to a program have been generated effectively (Carlucci et al., 2014). Although there have been attempts to develop measurement and assessment systems for PS during a long period, fundamental weaknesses persisted in the quality and use of performance information. In parallel with practical developments, the body of literature seeking to document, explain and improve public sector performance measurement (PSPM) has also grown (Carlucci et al., 2014).

Research Method

A literature review can be carried out through a narrative review (as was the case for many years in management science), or SLR and meta-analysis (as was the case for many years in medicine) (Tranfield et al., 2003). Moreover, the traditional narrative literature review has been criticised by experts as it can be heavily influenced by research bias. Using the SLR framework, the bias and systematic error can be explicitly minimised (Petticrew & Roberts, 2006). Following Denyer and Tranfield (2009), a five-step procedure of content analysis is as follows: 1) question formulation, 2) relevant study location and selection, 3) study selection and evaluation, 4) analysis and synthesis, and 5) dissemination of review findings.

Question formulation

The formation of review questions was considered specifically important in case the literature is sparse (Denyer & Tranfield, 2009). In this study, the Context-Intervention-Mechanism-Output framework reformulated from the work of Pawson (2006) by Denyer and Tranfield (2009) was employed to focus on the review question. Thus, the explicit research questions were presented as followed.

- RQ1. Which immense systems are under investigation?
- RQ2. The impacts of what system are being explored?
- RQ3. Under what conditions are these apparatuses operated or not operated?
- RQ4. What are the planned and unplanned effects?

Relevant study location and selection. Five databases were employed in this study including Emerald plus, Sage, ScienceDirect journals, Taylor and Francis Online, and Wiley Online Library due to the tremendous depositories of business research and a mass of publishing outlets of highest impact on the research community (Sandberg & Aarikka-Stenroos, 2014) and bias lessening. Besides, the timeframe of our literature review was stretched from 1990 to 2019 because PMSs have been widely adopted among PSOs since the 1990s (Jääskeläinen & Sillanpää, 2013). Following the process of Tranfield et al. (2003), a search string by pointing out all synonyms for “public sector measurement performance” has been developed and broadened. In this regard, the set of keywords shaped a search string in which all synonyms for “performance measurement” were coordinated with all the cognates for “public sector”.

Table 1: Keywords or key terms used in the literature search

| Public sector | Performance measurement |
|---|---|
| Public Administration, Public sector bank, Public Education sector, Government, Federal University, Public Health sector, Public High school, Public Hospital, Public Library, Local government, Ministry, Public Service sector, Public University | Evaluation Performance, Measurement Method, Measuring Method, Measurement Technique, Measuring Technique, Performance Indicator, Performance Assessment |

Source: Researchers’ elaboration, 2019.

For each database, the search string first had to be adjusted according to the specific search guidelines and was then applied in the search field offering the most relevant results. In total, all the related articles were identified: 58,388 from Science Direct, and 40,766 from Emerald, 5,380 articles from Wiley Online Library and 130,168 Taylor and Francis Online, the search

process produced 234,720 documents comprising journal articles and conference papers. The searches generated a large number of duplicate articles. The differences in the number of articles found were attributable to the scope of each respective repository.

Study selection and evaluation

The limitation of our literature review was between January 1990 and January 2019. The investigation was limited to peer-reviewed journals (Denyer & Tranfield, 2009) in order to ensure that only high-quality manuscripts were considered. Subsequently, the inclusion/exclusion criteria list was developed and agreed on by all the authors and was shown in Table 2.

Table 2: The Inclusion and Exclusion criteria

| Inclusion criteria | Exclusion criteria |
|---|---|
| Articles must be obtainable as full-text, recorded in English, published during the period between 1990 and 2019, relevant to the research questions; belong to selected digital databases, and focused on the PM only in PS. | Articles are not obtainable as full-text, external search timeframe, non-English manuscript, not concerned with the research questions, duplicated studies, not only concentrated on PS, and the type of literature review. |

Source: Researchers' elaboration, 2019.

Eliminating duplicate articles and analysing the abstracts according to the predefined inclusion and exclusion criteria were carried out by the authors. The inclusion/exclusion decisions were based solely on analysis of the abstracts, without any additional information and with the tendency to be inclusive rather than exclusive. Conference proceedings were not included, and, for that reason, some papers identified in our search were excluded. Consequently, a final sample of 80 articles was shown in Table 3.

Table 3: Overview of search result and study selection

| S. No. | Database | Period of search | After automated and manual search | After reading complete study and quality assessment |
|--------|-------------------------|------------------|-----------------------------------|---|
| 1 | Emeraldinsight.com | 1990-2019 | 40,766 | 30 |
| 2 | ScienceDirect | 1995-2019 | 58,388 | 12 |
| 3 | Taylor & Francis Online | 1990-2019 | 130,168 | 39 |
| 4 | Wiley Online | 1990-2019 | 5,380 | 3 |
| Total | | | 234,702 | 80 |

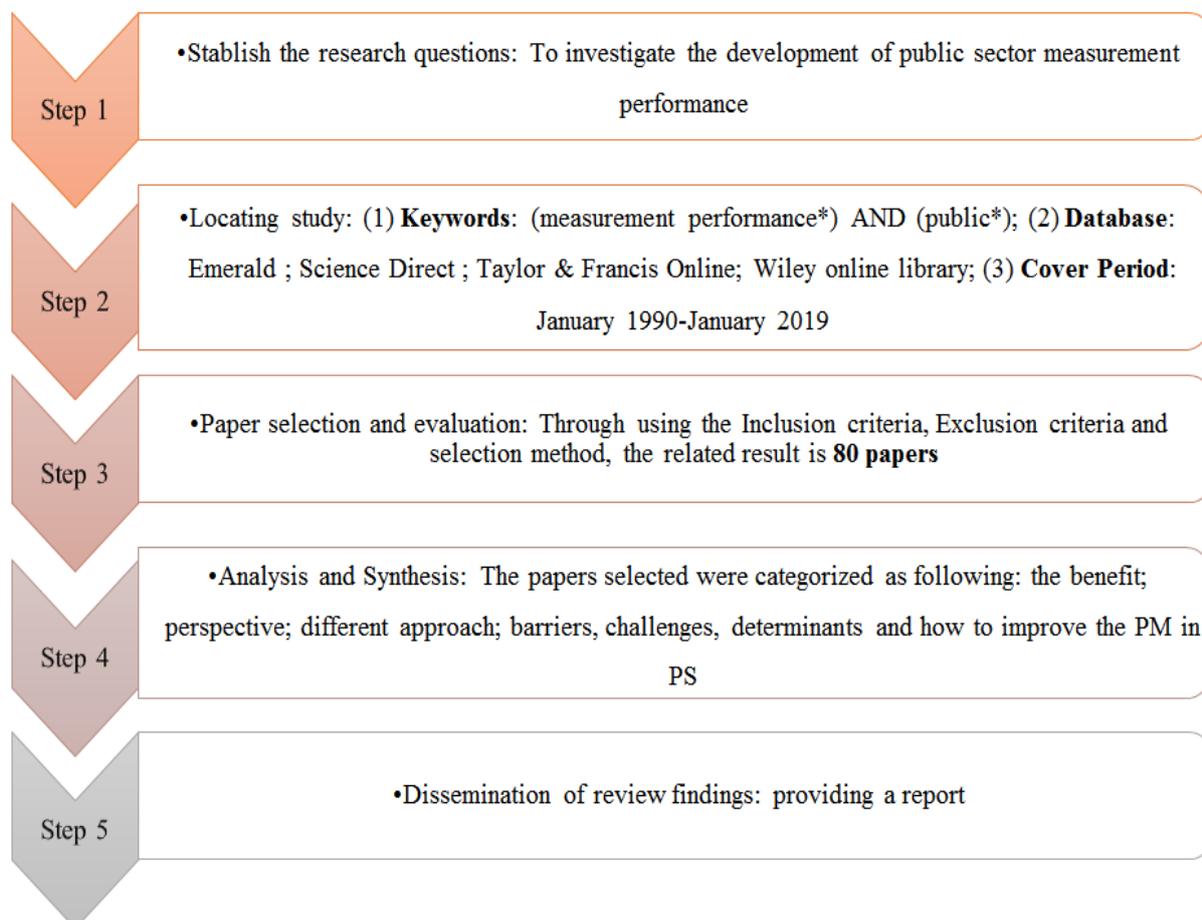
Source: Researchers' calculation, 2019.

Analysis and synthesis

Owing to the fact that analysis and synthesis themes were colligated with related data isolation, tabulation, paradigm tracking down, and common ideas aggregation (Ghadge et al., 2012), the thematic analysis was employed in this study to systematically classify the content of the papers and determine relationships (Lane et al., 2006) between PM with other relating domains of operation in the PS. Simultaneously, an interpretative synthesis method was conducted through encrypting due to heterogeneous nature and the majority of qualitative studies contented in the sample. The two researchers also verified the codes in order to eliminate inhomogeneity and reinforce the internal validity and credibility of the data (Elliott et al., 1999).

Dissemination of review findings. In the SLR paper, the final step was to provide a report which illustrated comprehensibly the review process and the overall results (Denyer & Tranfield, 2009).

Figure 1. Defining the search process for the systematic literature review



Source: Adapted by the authors from Denyer and Tranfield, 2009.

Results Review

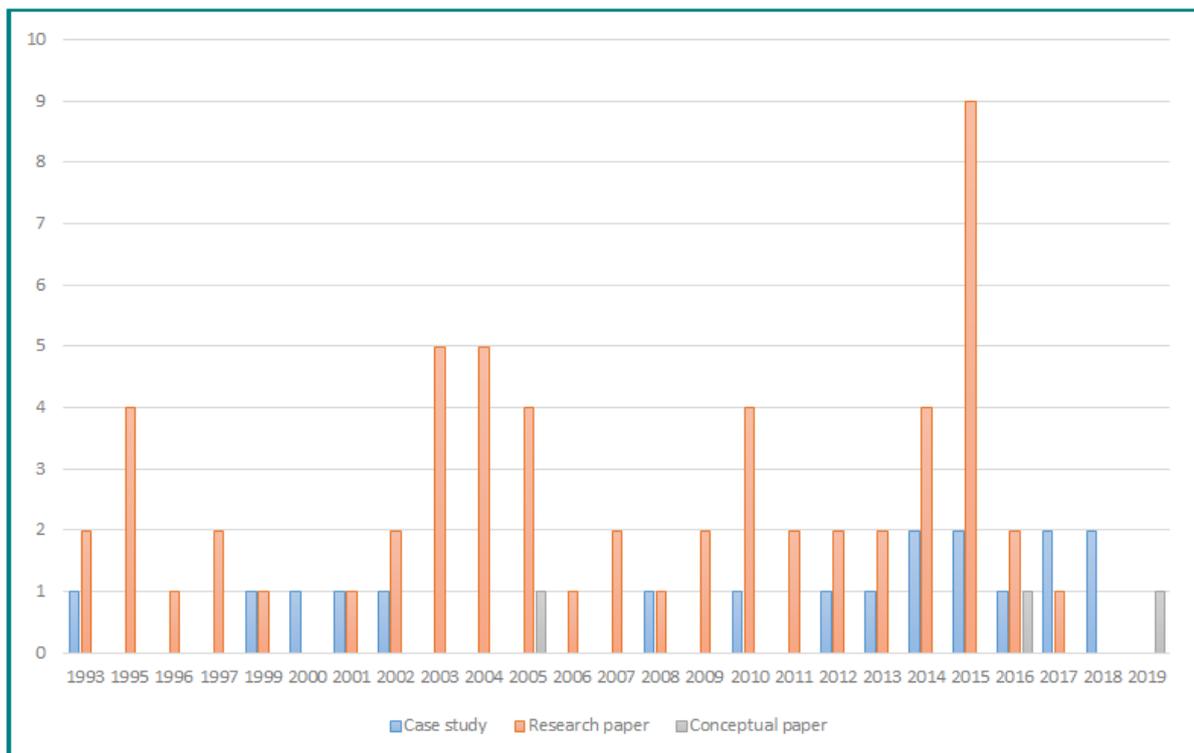
The systematic review identified 80 scholarly journal articles. We further reviewed these articles and classified them under various categories. The descriptive analyses, categories that defined the articles were selected, such as the year of publication, journal, and countries focused on in studies, the classification of papers and findings.

Descriptive Analysis of the Reviewed Articles

Annually with selected journals

The chart provides an insight into the development and research to the topic annually. Although the number of published articles on the topic at hand was still very limited throughout the considered period, it was possible to note a relatively positive trend from 2003 to 2005. The highest number of publications was recorded in 2015 and then slightly downfall in the number of papers from 2016 to 2019. There was a drastic boost in the number of research type publications in comparison with case studies and conceptual papers.

Figure 2. Annual report of related papers in selected journals

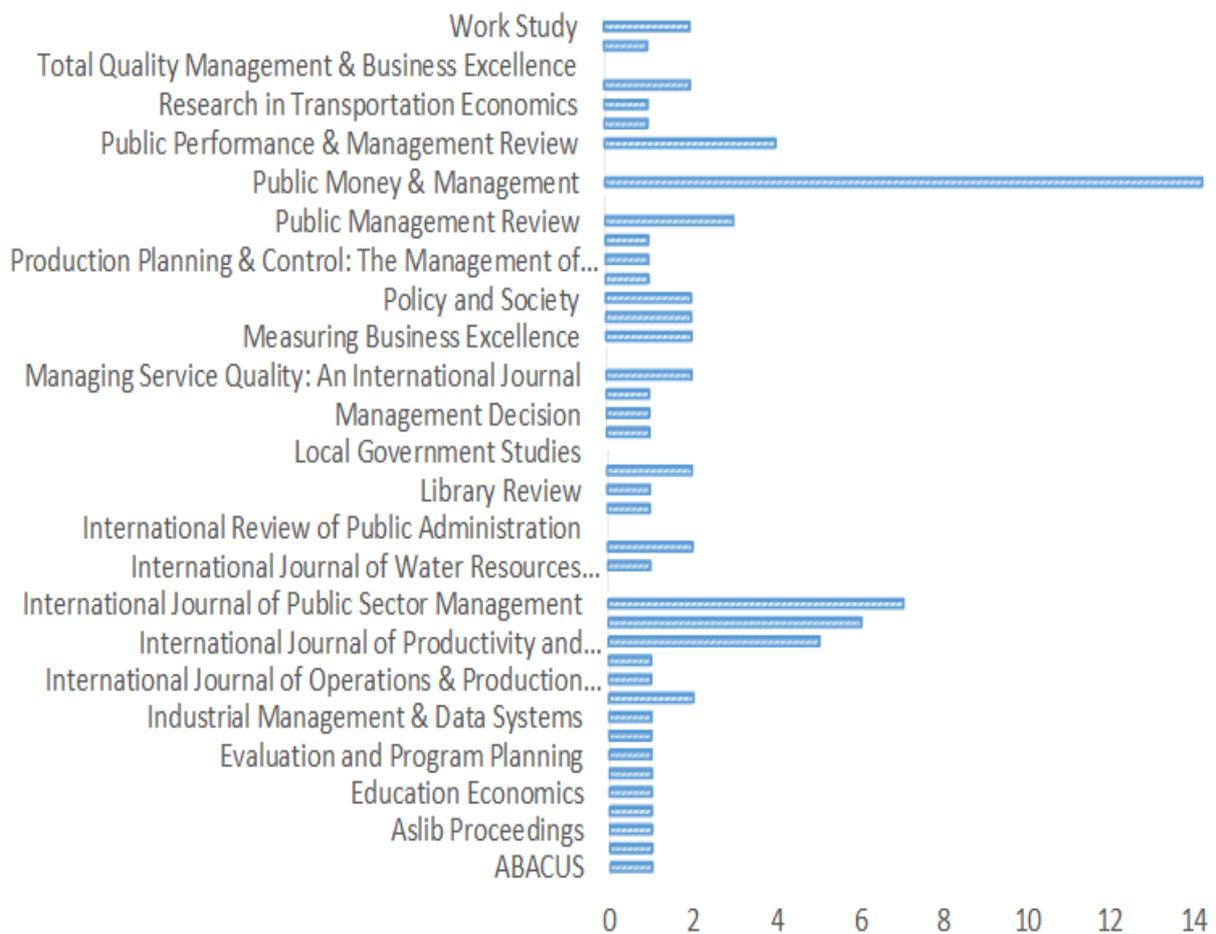


Source: Researchers' calculation, 2019.

Academic Journals. The top contributor was the Public Money & Management (14 papers), followed by the International Journal of Public Sector Management (7 papers), International

Journal of Productivity and Performance Management (5 papers) and International Journal of Public Administration (6 papers). The remaining identified journal was published one according to the results demonstrated in Figure 3. On the other hand, PSPM had a more diverse spread. Each journal article available on the topic was published in different journals, illustrating that no journal has yet made a focus on this niche area of research.

Figure 3. Number of academic journals



Source: Researchers’ calculation, 2019.

Distribution of articles by country

The descriptive analysis also highlighted that the studies in this research focused on diverse countries from almost all continents. It was interesting to observe that one article presented a “multi-country” perspective while most of the studies adopted in single-country focus. This means the relationship between PSPM and the context where it was implemented is an interesting research direction focus. The ‘no country’ category was devoted to this review, which does not show any geographical focus. Besides, results of systematic literature review

showed that most of the studies of PSPM were conducted in developed countries, especially in the UK, US and Australia. Conversely, despite the importance of application and improving PSPM in developing countries, only 17 studies explored the subject matter in developing countries. This indicates that the literature gap exists for the study of PSPM in developing countries. In this respect, one can anticipate more research on the subject matter in developing countries.

Table 4: Distribution of articles by country

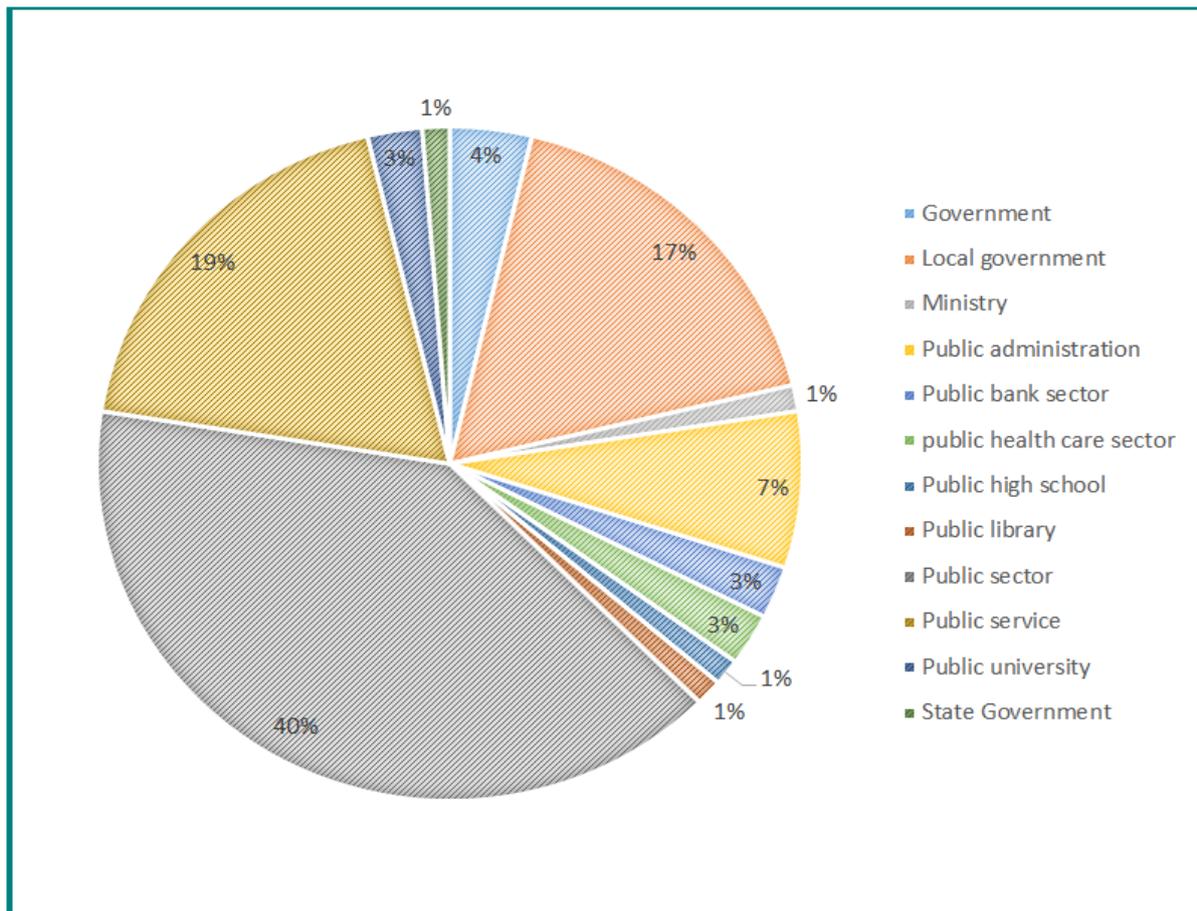
| Country | Orientation | | Count of No. |
|-----------------------|------------------|------------------|------------------|
| | Developed | Developing | |
| Australia | x | | 8 |
| Belgium | x | | 1 |
| Canada | x | | 1 |
| Denmark | x | | 4 |
| Finland | x | | 7 |
| France | x | | 2 |
| Italia | x | | 4 |
| Italia and Spain | x | | 1 |
| Netherlands | x | | 1 |
| Portugal | x | | 2 |
| Sweden | x | | 2 |
| Turkey | x | | 1 |
| United Kingdom | x | | 15 |
| United State | x | | 8 |
| United State & Canada | x | | 1 |
| Total | <u>16</u> | | <u>58</u> |
| China | | x | 6 |
| Estonia | | x | 1 |
| Greece | | x | 1 |
| India | | x | 2 |
| Indonesia | | x | 2 |
| Israel | | x | 1 |
| Lithuania | | x | 1 |
| Serbia | | x | 1 |
| Thailand | | x | 2 |
| Total | | <u>10</u> | <u>17</u> |
| Multiple countries | | | 1 |
| Non country | | | 4 |
| Total | | | <u>5</u> |

Source: Researchers' calculation, 2019.

The associated domains of selected articles

Another outcome of the SLR is demonstrated in Figure 4, where it is evident that majority of articles published on the topic of PSPM were published in the PS, whereas there were 38 articles based purely on PSPM in the PS, followed by public service (16 papers) and local government (14 papers).

Figure 4. Associated domains in selected journals



Source: Researchers' calculation, 2019.

Description of Categories Identified

The outcome of the content analysis was presented in line with the seven themes identified. Each section of the findings covered coherent information about the subject of identified themes.

Category 1: Benefits of Performance Measurement in the Public Sector. Usefulness consisted of a more comprehensive picture of PSPM, supporting decision-making processes

better and, at the same time, providing reports to the public and decision makers, supporting assessment and evaluation of program results and outcomes (Bernstein, 2001), and increasing the success of PMS implementation management commitment through good leadership (Akbar et al., 2012). PM tools, while delivering better service and greater accountability (Bolton, 2003), would support the enhancement in performance (Speklé & Verbeeten, 2014) and prove to be central to the attainment of anti-corruption (Fritzen, 2007). Additionally, PMSs design could not only be treated as a central thrust for establish strategic planning and management framework (Hoque, 2008) but also enhanced organisational learning, interactive use of SPMS as well as strategic alignment (Srimai et al., 2011).

Category 2: Perspectives of Application PM in the PS. As proposed by Moullin (2004), there were eight essentials perspectives of establishing PM including (1) Using a balanced set of measures; (2) Making sure you measure what matters to service users and other stakeholders; (3) Involving staff in determining the measures; (4) Including both perception measures and performance indicators; (5) Using a combination of outcome and process measures; (6) Taking account of the cost of measuring performance; (7) Having clear systems for translating feedback from measures into a strategy for action; (8) Focusing on continuous improvement of measurement systems, not a blame culture. Besides, based on different countries holding different cultures, politics and choice preferences, Woelert (2015) argued that effective PM application in the PS in Australia depend largely on PM application supporters who could be budget officers and agency personnel in state governments (Willoughby, 2004). Simultaneously, the POs should take notice of output specification; measuring quality and effectiveness and identifying the client in order to release better performance (Di Francesco, 1999). The PMSs must be based on the drivers of success to support the quest for Organisational excellence (Mani et al., 2003).

Category 3: Exploring Different Measurement Approaches. Most of the research in SLR about implementing PM in the PS depended on the object of the POs, including coping with environmental and organisational change (Brignall, 1993); providing service to public facilities (Parkan, 1999); addressing organisational goals and achievements (Wang, 2002); measuring the performance of utilities (Kumar & Managi, 2010); setting business excellence in the organisation (Kanji & Moura E Sá, 2007); depending on the objectives of the individual undertaking the analysis (Smith, 1995); and assessing the financial performance of local governments (Turley et al., 2015). Moreover, with the goal to develop a new method for measuring productivity in large POs, a measurement method was presented to meet the need for developing more sophisticated measurement systems (Jääskeläinen & Uusi-Rauva, 2011). Building on the importance of PM in enhancing accountability and responsibility of government action, financial data from annual financial reports (Carmeli, 2002) or performance and strategic management system-had been introduced (Lilian Chan, 2004). The advantages of scorecard went well beyond simply measuring performance: strategic,

organisational and human benefits were gained (Dreveton, 2013). The pairing of a continuum approach to PM along with specific management activities provided public managers with a new understanding of how PMSs may be designed to reflect the complexity of POs (Oh & Bush, 2015). The complex combination of PM methods was known in the business sector and applied in the PS (Balabonienė & Večerskienė, 2015) due to the implementation of the BSC experienced problems with measurement of performance metrics and aggregation of data (Carmona & Grönlund, 2003). The Public Sector Scorecard (PSS) was, then, proved to be an effective framework which helped organisations improve outcomes for service users and stakeholders without increasing overall costs, and developed measures of performance that help them improve and assure quality without motivating staff to achieve arbitrary targets at the expense of poor service to the public (Moullin, 2017).

Category 4: Common Barriers in the Implementation PM in the PS. As stated by Siverbo et al. (2019), one of the greatest initial barriers of PSPM was the dysfunctional consequences. It was, in particular, a matter of interactions between PM design and PM use, between control practices in control system and between PM and the context. According to Pidd (2005), the control aspects dominating the others were the reasons why PMSs in public services could lead to dysfunctional consequences even when people operated with the best of intentions. In addition, PM current systems in the PS could have a significant influence on improving services (Adcroft & Willis, 2005).

Category 5: Challenge in Implementing and Improving PM in the PS. Measuring the performance in the PS in multiple countries was very challenging due to the culture (Bavon, 1995). Another challenge was the limitation of the effectiveness of cascaded down to lower levels of the organisation due to various inter-related problems of leadership, training, feedback, employee empowerment and weak incentive schemes combined to cause communication difficulties (Umashev & Willett, 2008).

While Phusavat et al. (2009) argued that four important roadblocks in implementing PM in an organisation relate to staff empowerment, budgeting, external knowledge, and linkage with software usages, Jääskeläinen and Sillanpää (2013) claimed that the greatest challenge when implementing PSPM was the commitment of the operative level achievement, and the chosen measurement tool was suitable for the identified managerial requirements of the organisation. Additionally, the challenge in implementing PSPM was the respect to an NPM-based system lied in both measures and methods (Agostino & Arnaboldi, 2015) and treating audiences as consumers and citizens, the range of stakeholders (Hastings, 2004).

Category 6: Factors shaping PM in the PS. In the scope of the PS, another critical success factor for PM were implementation cognitive factors (e.g., strategic leadership and change, managing stakeholders, goal coherence, defining Pis); behavioural factors (e.g.,

entrepreneurship, knowledge management, e-governance, HR empowerment, citizen satisfaction) and ethical factors (e.g., conceptual freedom, meritocracy) increased in significance at that time (Sotirakou & Zeppou, 2006). In a study on PSPM in Finland, Rautiainen (2010) reported that contending legitimations have an effect on PM. Nevertheless, in the developing countries, the instability, poor strategic planning and policy analysis, an implementation gap and uncritical, uninformed transfer of Western performance management initiatives could impact the PM in Estonia (Nõmm & Randma-Liiv, 2012). Research undertaken in China provided some appreciation of the unified political and administrative system, the Chinese developmental state and its performance legitimacy strategy. However, decentralised intergovernmental systems, and the bureaucratic culture and informal rules had negative effects on performance measures (Jing et al., 2015). Consequently, Turkish public administration only got a high success in applying PM through taking notice of the organisational support, technical capacity, quality of performance measures (Eliuz et al., 2016). Furthermore, the finding of Carlucci et al. (2014) proved that the political, cultural and rational factors might affect the PM in Italy. At the ministry level, the characteristics of the activity could play a significant role in the success of PM (Barbato & Turri, 2017).

Category 7: The method for improvement PM in the PS. As PM was becoming increasingly important in the PS, the development of a PMS was best served by a strategic planning approach with top management support, active management participation in goal setting, a small manageable number of goals, a strong link with the budget process, and an independent evaluation process (Tuck & Zaleski, 1996). In addition, legislators should make an effort to define more precisely the intended results of the law and program they authorise and should leave managers enough latitude to select what should be done and which measures to use to demonstrate progress toward achieving the results prescribed by lawmakers (Halachmi, 2002). Besides, the empirical analysis stressed the importance in a democracy that both central and local governments provided information on the efficiency and effectiveness so that voters were in a better position to judge whether they were getting value for money (Haselbekke, 1995). Moreover, a better understanding in causal relationships, improving the informativeness of used measures and the role of the measures as communication devices, and using the measures as incentives for improvement (Pekkanen & Niemi, 2013) with significant changes in the practices of accountability for PM and risk management within the PS, the better result in the improvement of performance would take place (Rana et al., 2018).

Research Gaps and Future Research Avenues

Firstly, PMSs was stated to represent the principal element driving the management process and that the information produced by PMSs could be used for diverse purposes (Sole & Schiuma, 2010). Based on the evidence presented in the descriptive analysis, there was a slight difference in the prior literature on applying PM in the PS due to regional

heterogeneity. The literature argued that these studies were mostly conducted in the USA, followed by the Italia, Finland, UK and Australia. The SLR also proved that although the PSPM literature on less developed and developing countries only gained scholarly attention since the last decade, studies from these economies are still nascent. All in all, the PSPM were suggested to be interpreted differently, based on its geographical area.

Secondly, despite the significant volume of studies on PSPM, a scarcity of empirical research depicted in detail the systems and application processes at diverse levels of government to evaluate and manage performance. One of the most vital matters was that there was an urgent claim for a generalised model of PM for all types of the PS not only to implement but also to overcome the challenging demands in their operations. On the other hand, remarkable renovations in performance and results-based accountability have not been entirely fully accomplished, and research has been still in need of defining key determinants of successful design, implementation, and using of PMSs.

Thirdly, there were no plain PM tool usages for almost all various phases of the PS. As being established to be served for several purposes, the main activity of PSOs was possibly assumed to offer the social, educational, scientific, cultural, sport and other paid or free services stipulated in the legal acts to the residents in corresponding to the set terms and conditions. The PS's activities were arduous to differentiate from one another, had numerous principals, are produced as well as other organisations, and demonstrated over a lengthy period and not all PSOs were analogous, additionally. There has been still an increasing claim for a strong PM framework to meet the challenging demands in their activities including generalising the basic standards of the PS requirements as well as adjusting the characteristics of each type of organisations in the PS effortlessly.

Conclusion and Implications

In a nutshell, the scientific literature of the past thirty years to identify the features and obstacles of PSPM has been systematically conducted through SLR. The adoption of an SLR is not only treated as an instrument to answer our key research questions appeared to be very useful but also provides important insights that can bring out several benefits.

First of all, the finding of this current research provides a great deal of useful knowledge of PSPM which plays an important role in helping managers who deal with steering organisations toward the achievement of their performance goals to determine whether the associated benefits are worth the respective costs and investments in evaluating alternative PM for PSOs. Furthermore, taking these results into consideration, managers can gain an in-depth understanding to release a better decision on given options, of not only the associated upsides but also the related downsides and pitfalls of the PM tool and how to use it efficiently



and effectively. Next, legislators should make an effort to define more precisely the intended results of the law and the program they authorise and should leave managers enough latitude to select what should be done and which measures to use in order to demonstrate progress toward achieving the results prescribed by lawmakers (Halachmi, 2002).

Based on the analysis of prior research, this work identified detrimental PSPM antecedent effects and potential routes through which they can jeopardise performance. Thereby, the study uncovered implied causal relationships of high relevance for managers. The study fosters managers' conceptual understanding of the reasons for PSPM failure. Thus, managers should take notice of evaluating the evidenced and suggested performance effects in this paper to both broaden their perspective of possible unfavourable PS scenarios and develop their preparedness for and resilience to them.

In recognition that lack of success can be reached without sharing foundations embedded in the academic arena among practitioners, validity tests of the foundations will thus be welcomed to verify and complement them in contexts outside of the academic literature.

Notwithstanding the rigorous SLR approach used by this study, a dash of limitations is still in need of being tackled. Firstly, the study was limited to publications within a 30-year period gathered from 4 databases, the choice of number and keyword type, the resulting selection of studies, the choice of relevant outcomes, the interpretation, the generalisation and the application of results. Secondly, the synthesised articles were restricted to only articles published in peer-review journals although references were made to evidence from other sources to reduced bias, excluding items such as books and grey literature. However, the value of this study lied within its robust and exhaustive methodology of synthesising a large body of evidence from prior studies on the topic for effortless comprehension (Denyer & Tranfield, 2009).



REFERENCES

- Adcroft, A & Willis, R (2005), The (un)intended outcome of public sector performance measurement, *International Journal of Public Sector Management*, 18(5), 386–400.
- Agostino, D & Arnaboldi, M (2015), The New Public Management in hybrid settings: New challenges for performance measure, *International Review of Public Administration*, 20(4), 353–369.
- Akbar, R, Pilcher, R & Perrin, B (2012), Performance measurement in Indonesia: the case of local government, *Pacific Accounting Review*, 24(3), 262–291.
- Balabonienė, I & Večerskienė, G (2015), The Aspects of Performance Measurement in Public Sector Organisation, *Procedia - Social and Behavioral Sciences*, 213, 314–320.
- Barbato, G & Turri, M (2017), Understanding public performance measurement through theoretical pluralism, *International Journal of Public Sector Management*, 30(1), 15–30.
- Bavon, A (1995), Innovations in performance measurement systems: a comparative perspective, *International Journal of Public Administration*, 18(2-3), 491–519.
- Bernstein, D J (2001), Local government measurement use to focus on performance and results, *Evaluation and Program Planning*, 24(1), 95–101.
- Bolton, M (2003), Public sector performance measurement: delivering greater accountability, *Work Study*, 52(1), 20–24
- Brignall, S (1993), Performance measurement and change in local government: A general case and a childcare application, *Public Money & Management*, 13(4), 23–30.
- Buathong, S & Bangchokdee, S (2017), The use of the performance measures in Thai public hospitals, *Asian Review of Accounting*, 25(4), 472–485.
- Carlucci, D, Schiuma, G & Sole, F (2014), The adoption and implementation of performance measurement process in Italian public organisations: the influence of political, cultural and rational factors, *Production Planning & Control*, 1–14.
- Carmeli, A (2002), A Conceptual and Practical Framework of Measuring Performance of Local Authorities in Financial Terms: Analysing the Case of Israel, *Local Government Studies*, 28(1), 21–36.



- Carmona, S & Grönlund, A (2003), Measures vs actions: the balanced scorecard in Swedish Law Enforcement, *International Journal of Operations & Production Management*, 23(12), 1475–1496.
- Denyer, D & Tranfield, D (2009), Producing a Systematic Review. In: Buchanan; D. and Bryman, A., Eds., *The Sage Handbook of Organisational Research Methods*, Sage, London, 671-689.
- Di Francesco, M (1999), Measuring performance in policy advice output: Australian developments, *International Journal of Public Sector Management*, 12(5), 420–431.
- Dreveton, B (2013), The advantages of the balanced scorecard in the public sector: beyond performance measurement, *Public Money & Management*, 33(2), 131–136
- Eliuz, S, Kapucu, N, Ustun, Y & Demirhan, C (2016), Predictors of an Effective Performance Measurement System: Evidence from Municipal Governments in Turkey, *International Journal of Public Administration*, 40(4), 329–341
- Elliott, R, Fischer, C T & Rennie, D L (1999), Evolving guidelines for publication of qualitative research studies in psychology and related fields, *British Journal of Clinical Psychology*, 38(3), 215-229.
- Fritzen, S A (2007), Crafting performance measurement systems to reduce corruption risks in complex organisations: the case of the World Bank, *Measuring Business Excellence*, 11(4), 23–32.
- Ghadge, A, Dani, S, & Kalawsky, R (2012), Supply chain risk management: present and future scope, *The International Journal of Logistics Management*, 23(3), 313–339.
- Halachmi, A (2002), Performance measurement: a look at some possible dysfunctions, *Work Study*, 51(5), 230–239.
- Haselbekke, A G J (1995), Public policy and performance measurement in the Netherlands, *Public Money & Management*, 15(4), 31–38.
- Hastings, C (2004), Discussion of performance measures in public service broadcasting, *Aslib Proceedings*, 56(5), 301–307
- Hoque, Z (2008), Measuring and reporting public sector outputs/outcomes, *International Journal of Public Sector Management*, 21(5), 468–493



- Jääskeläinen, A & Sillanpää, V (2013), Overcoming challenges in the implementation of performance measurement, *International Journal of Public Sector Management*, 26(6), 440–454
- Jääskeläinen, A & Uusi-Rauva, E (2011), Bottom-up approach for productivity measurement in large public organisations, *International Journal of Productivity and Performance Management*, 60(3), 252–267.
- Jing, Y, Cui, Y, & Li, D (2015), The politics of performance measurement in China, *Policy and Society*, 34(1), 49–61.
- Kanji, G & Moura E Sá, P (2007), Performance Measurement and Business Excellence: The Reinforcing Link for the Public Sector, *Total Quality Management & Business Excellence*, 18(1-2), 49–56.
- Kaplan, R S (1986), Accounting lag – obsolescence of cost accounting systems, *California Management Review*, 28(2), 174-199
- Kaplan, R S (1990), *Measures for Manufacturing Excellence*, Harvard Business School Press, Boston.
- Kumar, S & Managi, S (2010), Service Quality and Performance Measurement: Evidence from the Indian Water Sector, *International Journal of Water Resources Development*, 26(2), 173–191.
- Lane, P, Koka, B, & Pathak, S (2006), The reification of absorptive capacity: a critical review and rejuvenation of the construct, *Academy of Management Review*, 31(4), 833-863.
- Langham, J & Paulsen, N (2018), Applying the Experience Effectiveness (XE) Framework in the Canadian Public Sector, *Design Management Journal*, 13(1), 53–69
- Lilian Chan, Y (2004), Performance measurement and adoption of balanced scorecards, *International Journal of Public Sector Management*, 17(3), 204–221
- Mani, T P, e Sá, P M & Kanji, G (2003), Finding the path to organisational excellence in Portugese local government: A performance measurement approach, *Total Quality Management & Business Excellence*, 14(4), 491–505.
- Moullin, M (2004), Eight essentials of performance measurement, *International Journal of Health Care Quality Assurance*, 17(3), 110–112
- Moullin, M (2017), Improving and evaluating performance with the Public Sector Scorecard, *International Journal of Productivity and Performance Management*, 66(4), 442–458.



- Neely, A, Gregory, M & Platts, K (2005), Performance measurement system design: a literature review and research agenda, *International Journal of Operations and Production Management*, 25(12), 1228-1263.
- Nõmm, K & Randma-Liiv, T (2012), Performance Measurement and Performance Information in New Democracies, *Public Management Review*, 14(7), 859–879.
- Oh, Y & Bush, C B (2015), Assessing public sector performance and untangling the complexity of public performance measurement, *International Review of Public Administration*, 20(3), 256–272.
- Parkan, C (1999), Performance measurement in government services, *Managing Service Quality: An International Journal*, 9(2), 121–135.
- Parker, C (2000), Performance measurement, *Work Study*, 49(2), 63–66.
- Pekkanen, P & Niemi, P (2013), Process performance improvement in justice organisations—Pitfalls of performance measurement, *International Journal of Production Economics*, 143(2), 605–611.
- Petticrew, M & Roberts, H (2006), *Systematic Reviews in the Social Sciences: A practical guide*, Blackwell, Oxford
- Phusavat, K, Anussornnitisarn, P, Helo, P & Dwight, R (2009), Performance measurement: roles and challenges, *Industrial Management & Data Systems*, 109(5), 646–664.
- Pidd, M (2005), Perversity in public service performance measurement, *International Journal of Productivity and Performance Management*, 54(5/6), 482–493.
- Rana, T, Hoque, Z & Jacobs, K (2018), Public sector reform implications for performance measurement and risk management practice: insights from Australia, *Public Money & Management*, 1–10.
- Rautiainen, A (2010), Contending legitimations, *Accounting, Auditing & Accountability Journal*, 23(3), 373–391.
- Ridley, C E, & Simon, H A (1938), *Measuring municipal activities. A survey of suggested criteria and reporting forms for appraising administration*, International City Managers Association, Chicago.
- Sandberg, B & Aarikka-Stenroos, L (2014), What makes it so difficult? A systematic review on barriers to radical innovation, *Industrial Marketing Management*, 43(8), 1293–1305.



- Siverbo, S, Cäker, M & Åkesson, J (2019), Conceptualising dysfunctional consequences of performance measurement in the public sector, *Public Management Review*, 1–23.
- Smith, P (1995), Performance indicators and outcome in the public sector, *Public Money & Management*, 15(4), 13–16.
- Sole, F & Schiuma, G (2010), Using performance measures in public organisations: challenges of Italian public administrations, *Measuring Business Excellence*, 14(3), 70–84.
- Sotirakou, T & Zeppou, M (2006), Utilising performance measurement to modernise the Greek public sector, *Management Decision*, 44(9), 1277–1304.
- 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.
- Srimai, S, Damsaman, N, & Bangchokdee, S (2011), Performance measurement, organisational learning and strategic alignment: an exploratory study in Thai public sector, *Measuring Business Excellence*, 15(2), 57–69.
- Streib, G & Poister, T (1999), Assessing the validity, legitimacy and functionality of performance measurement systems in municipal governments, *American Review of Public Administration*, 29(2), 107–123.
- Tranfield, D, Denyer, D & Smart, P (2003), Towards a Methodology for Developing Evidence- Informed Management Knowledge by Means of Systematic Review, *British Journal of Management*, 14(3), 207-222.
- Tuck, N & Zaleski, G (1996), Criteria for developing performance measurement systems in the public sector, *International Journal of Public Administration*, 19(11-12), 1945–1978.
- Turley, G, Robbins, G & McNena, S (2015), A Framework to Measure the Financial Performance of Local Governments, *Local Government Studies*, 41(3), 401–420.
- Umashev, C & Willett, R (2008), Challenges to Implementing Strategic Performance Measurement Systems in Multi-Objective Organisations: The Case of a Large Local Government Authority, *Abacus*, 44(4), 377–398.
- Wang, X (2002), Perception and reality in developing an outcome performance measurement system, *International Journal of Public Administration*, 25(6), 805–829.



Willoughby, K G (2004), Performance Measurement and Budget Balancing: State Government Perspective, *Public Budgeting Finance*, 24(2), 21–39.

Woelert, P (2015), The “logic of escalation” in performance measurement: An analysis of the dynamics of a research evaluation system, *Policy and Society*, 34(1), 75–85.