Factors Promoting the Collection of Performance Measurement: Evidence from US Local Governments

Rashmi Krishnamurthy
Stephen J.R. Smith School
of Business
Queen's University
rashmi.krish16@gmail.com

Kevin C. Desouza
School of Public Affairs
Arizona State University
kev.desouza@gmail.com

Gregory S. Dawson
W.P. Carey School of
Business
Arizona State University
Gregory.Dawson@asu.edu

Alfred Tat-Kei Ho
School of Public Affairs
and Administration
University of Kansas
alfredho@ku.edu

Abstract

Performance measurement has gained significant importance around the world. Many governments are adopting performance measurement as a part of reform efforts. Despite the widespread practitioner attention, academic studies are inconclusive about the impact of performance measurement in the public sector. Moreover, while studies have examined what factors influence the adoption of performance measures and its impact, they have paid relatively less attention to the use of different types of performance measures. To fill this gap, this study examines: (1) what types of performance measures are collected by US local governments and; (2) how organizational, technical, and external factors influence the collection of performance measurement. Leveraging survey data, we conduct cluster analysis and qualitative comparative analysis (QCA) to study factors that impact the collection of performance measures. The findings of QCA highlight that no single condition dominates the collection of performance measurement, rather different combinations of organizational and external factors influence the outcome. The paper concludes by discussing the implications for local policymakers and managers.

1. Introduction

Over the last few decades, performance measurement has gained considerable attention around the world. Many governments are adopting and implementing performance measurement as a part of reform efforts [34]. Scholars are also investigating the effect of performance measurement [22, 27], factors influencing the adoption and use of performance measurement [1, 19], and perception of performance measurement in the public sector [16, 17].

Further, performance measurement is favored by elected officials and citizens as a mechanism to make public agencies more accountable, transparent, and responsive [35]. It is anticipated that the collection and use of performance measurement will help public agencies increase efficiency and effectiveness in operations, uncover problems, reduce costs, cut redundancies, and encourage innovative solutions [34]. Public agencies can leverage performance measurement to set agendas, monitor programs, track initiatives, and communicate with stakeholders [23]. In other words, performance measurement is the collection and use of quantitative indicators to continuously measure how an agency is meeting its organizational goals and tasks [16]. Thus, we define *performance measurement* as a collection of data on various outcomes related to caseload, efficiency, perception, trend and benchmarking measures.

Advancement in information and communication technologies (ICTs) have opened up new platforms to obtain, analyze, and share performance information [23]. Moreover, the adoption of ICTs will become a critical asset for effectively and efficiently collecting performance measures in the public sector [10]. As a result, performance management has never been more important, and its usage will continue to rise in the foreseeable future.

However, public agencies cannot ignore performance measurement during times of economic uncertainty. For instance, many local government are under pressure to provide high-quality services with severe budget constraints and downsized staff [21]. Despite these expectations and pressure, many public agencies are struggling to use performance measurement information effectively in the decision-making processes [19, 21]. Moreover, past studies are inconclusive about the impact and value of performance measurement in the public sector [17, 22].

While past studies offer mixed results about the use and effect of performance measurement in the public sector, these studies often focus on a narrow set of indicators to capture performance measurement. Further, many studies in recent years have focused on how performance measures are adopted and its impact on organizational outcomes, they often pay relatively less attention to the different types of performance measures collected by different departments [18].

HICSS

This study addresses this gap by examining what types of different performance measures are collected by US local governments. We first asked local government about the types of performance data they collected. Particularly, we asked four departments - Police, Code Compliance, Parks and Recreation, and Public Works – about what types of performance data they collected and how frequently. We specifically chose these departments as they vary in terms of the type of service they provide (technical, non-technical) and their interactions with citizens (coercive, non-coercive). *Coercive services* refer to services that lead to non-discretionary, involuntary official-citizen interactions as a result of certain individual behaviors or actions [3]. *Technical level* refers to the level of technicality that is needed to perform the job. Thus, the four departments surveyed can be mapped into each quadrant (see Figure 1).

	Coercive Services	Non-Coercive Services
Highly Technical	Police	Streets/Public Works
Not Highly Technical	Code Compliance	Parks and Recreation

Figure 1: Model of city services

Second, we also asked our survey respondents about their organizational and external environment. We use this information to examine how different factors influence the collection of performance measurement in local governments.

To answer our research questions, we conducted two-step cluster analysis to classify the grouping of different departments based on their performance measure usage. Next, we ran qualitative comparative analysis (QCA) to understand how different combination of organizational and external conditions influence the collection of performance measurement. The results indicate that no *single* condition dominates the collection of performance measurement, but rather *several* configurations explain the collection of performance measurement.

Our study contributes to the current literature on performance measurement in several ways. First, it takes a holistic view to understand what types of performance data are collected in local governments and how collection of performance data differs across departments. Second, it identifies how different configurations of factors promote the collection of performance measurement. The findings show that high use of performance measures is associated with different combinations of factors of organizational and external conditions, i.e. the "best" performance data collection strategy may vary, contingent upon different organizational contexts. No one factor explains the

collection of performance measurement, but rather a combination of factors influences the outcome. Public agencies can use different organizational and external configurations to improve the collection of performance measurement.

The rest of the paper is organized as follows. The next section discusses how organizational and external factors influence the collection of performance measures. Then, we present our data collection strategy. The following section outlines the findings of cluster and QCA analysis. The final section discusses the implications for local policymakers and managers.

2. Performance Measurement and ICT

Advocates of performance measurement often suggest that the collection of data is critical for improving organizational operations, processes, and outcomes. Successful organizations leverage new technologies to not only collect a wide variety of data such as project status, employee response time, and customer feedback, but also combine data from different sources to arrive at evidence driven decisions on wide ranging topics [5, 9]. For example, the New York Police Department developed real-time Crime Information Warehouse to collect and share data. police officers can utilize crime data in real-time to monitor trends and respond to public safety issues [23].

While the collection and use of performance measurement can offer many potential benefits, public agencies are still struggling to institutionalize performance measurement into organizational decision making [17, 19]. Several factors such as organizational culture, staff capacity, and information technology (IT) infrastructure often influence the adoption of performance measurement [19, 33, 34]. In addition to internal organizational environment, external pressures also influence the adoption of performance measures [26]. The following sub-section discusses how organizational and external environments influence the collection of performance measures.

2.1. Organizational Factors

Organizational culture is a key factor that influence the collection of performance measurement. An organization considering to move towards evidence-driven management may need to think about internal culture [19]. The culture of an organization often influences the attitude and beliefs of employees about change and risk taking [20]. Employees working in an innovation-driven organization are more willing to take risks and have a positive attitude towards performance measures. Further, an innovation-driven organization

has processes in place that reward risk taking behavior to encourage innovation.

Organizations adopting performance measurement also need to train their employees to work in a result-oriented work environment. Employees with adequate knowledge about data can then leverage performance data to improve work tasks [33]. Hence, capacity building is important so that employees can use performance measurement to track and monitor programs, respond efficiently, and communicate with stakeholders on policy outcomes.

2.2. Technical Factors

A critical component for the collection of performance measurement is adequate information technology (IT) infrastructure. IT infrastructure includes hardware, software, and wide-area networking that support the collection of performance measurement. Having an adequate IT infrastructure that can collect wide range of organizational data is important for understanding and evaluating the operations, processes, and outcomes. The collection of quality performance measurement is highly dependent on well-established systems and processes that automatically collect data on input and output measures [13, 34].

Technical factors also include human resources that support the planning and operations of technology infrastructure [11]. It is important for an organization to map its technical infrastructure with human resource and knowledge so that its physical asset and human resources are able to set up processes and systems for collecting performance measurement. This could include having dedicated data staff to manage the collection of performance measurement. Further, having own data staff may help public agencies avoid the burden of feeling overwhelmed with the collection of performance measurement. Public agencies often face the challenge of managing performance measurement in addition to their daily tasks, which could easily create backlogs and result in employee resistance [16]. Therefore, sufficient internal data staff support is likely to help public agencies effectively collect performance data to meet organizational needs and goals.

2.3. External Pressures

Previous studies have consistently reported that the elected officials' attitudes can influence the collection of performance measurement [17, 19, 21]. Elected officials support and pressure often plays a critical role in the adoption of organizational performance measurement [19]. When elected officials show an interest in the collection of performance measurement, public agencies are more likely to collect and report performance measures [11, 20]. If public managers perceive that elected officials are not interested in hearing and using performance measures, their interest in performance measurement declines [17]. This is because if elected officials are not interested in performance measurement, they are less likely to engage in dialogues with department heads about what types of data needed to be collected and analyzed, and they are less likely to use performance data in organizational decision-making.

Based on the above understanding, this paper hypothesizes that a *combination* of organizational and external pressures influences the collection of performance measures. Figure 2 displays our theoretical framework.
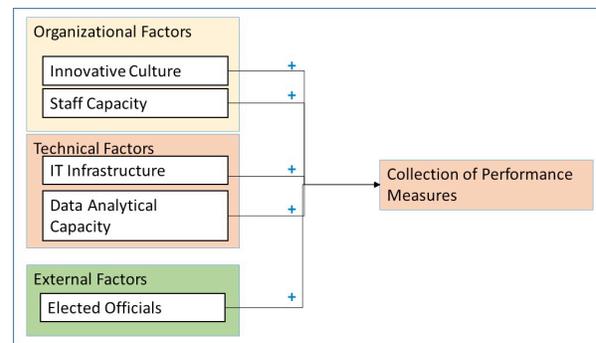
Figure 2: Theoretical framework

Using this theoretical framework, we hypothesize that:

H1: Innovative culture will be positively associated with the collection of performance measurement.
H2: Staff capacity will be positively associated with the collection of performance measurement.
H3: Having adequate IT infrastructure will be positively associated with the collection of performance measures.
H4: Availability of data analytical capacity will be positively associated with the collection of performance measures.
H5: Elected officials support and pressure will be positively associated with the collection of performance measures.

2.4. Variation in Service Delivery and Citizen Engagement

Local governments also collect performance measures to gauge the quality of service provided [24, 29]. It is anticipated that the collection of performance measures will help local governments provide efficient, responsive, and effective services that meets the

demands of the citizens [15]. However, to effectively collect and utilize performance measures, it is important to differentiate the type of service provided. For instance, citizen may opt for certain type of service (e.g. parks and recreation) and other times, they may be forced into certain services (e.g. police). Depending on the type of service, it is possible that citizen may experience different interactions with service providers, resulting in how these services are interpreted [3].

One type of service provided by local government can be classified into coercive services. In simple terms, *coercive services* are those where a local government specifically exercise authority to control citizen's behavior and actions [3, 22]. Governments, in general, engage in coercive services to prevent citizen's from harming each other (e.g. theft) and/or causing harm to society (e.g. encroaching land). It is often reported that citizens who are forced into interactions with coercive service providers often report negative interactions [3]. Past studies have consistently found that people provide negative feedback about their interactions with police officers, particularly if they encountered offensive or slow response from police officers [8, 31].

On the other hand, when citizens choose voluntarily to have interactions with service providers, they are likely to display different behaviors and attitudes. Citizens who choose to interact with service providers often provide positive feedback [3].. For example, public works department does not usually impose involuntary citizen-official interactions on street users, and parks and recreation programs cannot force citizens to sign up for certain programs or use certain parks and facilities. Due to different service nature and citizen-official relationships, departments may choose to focus on different aspects to measure performance [3, 24].

In addition to the type of service, it is important to differentiate departments based on their level of *technical capacity*, i.e. the extent of technicality expected of the departments to carry out their mission. For example, highly technical departments are more likely to have specific measurement of intermediate and final outcomes [18], while less technical departments are more likely to focus on outputs, user satisfaction, and process. .

Putting it all together, it can be anticipated that coercive and highly technical departments, such as police departments, are likely to collect most amount of performance measures, especially outcome measures, public sentiment, and comparative performance benchmarks, because they face pressures from public and elected officials for accountability and responsiveness. Coercive and non-technical units,

such as nuisance complaints and code enforcement in neighborhoods, may also face a lot of public accountability pressure but due to their non-technical nature, they may focus less on quantifiable efficiency and outcome measures or comparative benchmarks and more on public sentiment and satisfaction.

On the other hand, non-coercive departments, such as public works or parks and recreation, are likely to collect responsiveness, satisfaction and citizen complaint measures and focus on the voluntary, discretionary user experiences of the public. Relatively to parks and recreation, public works may also focus more on process, workload, and cost-efficiency due to its technical nature.

3. Data and Methodology

This study uses data from a survey of 280 counties and cities in the US. The survey asked Police, Code Compliance, Parks and Recreation, and Public Works departments within each county/city to respond to several questions related to the collection of performance measures. Before sending the full survey, a pilot survey was created and tested. The pilot survey was sent to six city managers, who distributed the survey to the heads of Police, Code Compliance, Parks and Recreation, and Public Works departments for completion. The comments received from the pilot survey were incorporated and a full survey was sent out using SurveyMonkey. The Alliance for Innovation (http://transformgov.org/) distributed that survey to counties and cities in their network. The survey was administered from February 27, 2017 to April 7, 2017.

In total, 132 responses were received. Not all departments in all communities responded, and in some cities/counties other departments such as the Mayor's offices responded to the survey. The response rate was 21 percent. To ensure comparability, responses received from other departments were removed before running the following analysis. This reduced our sample size to 100. Table 1 shows the distribution of survey responses across the four departments.

Department	N
Police	26
Code Compliance	23
Parks and Recreation	27
Public Works	24
Total	**100**

Table 1. Number of departments

The survey asked respondents about how often their department collected 11 different types of performance measures. These performance measures can be broadly classified into three main categories: caseload measures, efficiency and outcome measures, and trend and benchmarking measures. The responses include "never", "once/twice a year", "three/four times a year", "monthly", and "routinely". We ran a cluster analysis to see the pattern of performance measure usage. Cluster analysis combines variables into homogenous groups if they are strongly related and provide the same information [4].

A two-step process was used in to conduct cluster analysis. In the first step, we transformed 11 performance measures into binary variables to avoid scaling errors and biases. This transformation yielded 54 categories. Next, we used the *dendoextend* package in R to run a cluster analysis using the squared Euclidean distance matrix method with Ward coefficient matching coefficients [14, 36]. The results of cluster analysis yielded two groups.

4. Findings

The results of the first step in cluster analysis are reported in Tables 2 and 3 (Please note: due to space limitations, we are not including all tables). As Table 2 shows, about 61.5 percent of the Police departments in our sample reported that they routinely collect workload, complaint, and public sentiment measures. Moreover, 57.7 percent of Police departments reported that they collect responsiveness measures monthly.

variable	Never	Once/ Twice a year	Three/ Four a year	Mon thly	Routi nely
workload	0.0	0.0	7.7	30.8	61.5
cost efficiency	26.9	34.6	11.5	7.7	19.2
responsiveness	0.0	23.1	19.2	57.7	0.0
outcomes	0.0	15.4	15.4	26.9	42.3
satisfaction	11.5	50.0	19.2	11.5	7.7
complaint	0.0	7.7	11.5	19.2	61.5
public sentiment	7.7	7.7	11.5	11.5	61.5
trend	11.5	15.4	11.5	30.8	30.8
spatial	15.4	11.5	15.4	26.9	30.8
benchmark-peer	11.5	50.0	19.2	3.9	15.4
benchmark-national	7.7	65.4	3.9	11.5	11.5

Table 2. Collection of performance measures – Police

Parks and Recreation departments in our sample reported that they routinely collect workload (48.2 percent), complaint (63 percent), and public sentiment (44.4 percent) measures. However, unlike Police departments, they also collect cost efficiency (33.3 percent) and responsiveness (44.4 percent) measures regularly.

variable	Never	Once/ Twice a year	Three/ Four a year	Mon thly	Routi nely
workload	7.4	7.4	11.1	25.9	48.2
cost efficiency	3.7	11.1	25.9	33.3	25.9
responsiveness	7.4	22.2	25.9	44.4	0.0
outcomes	3.7	18.5	29.6	18.5	29.6
satisfaction	0.0	29.6	25.9	14.8	29.6
complaint	11.1	0.0	14.8	11.1	63.0
public sentiment	7.4	3.7	14.8	29.6	44.4
trend	11.1	22.2	40.7	7.4	18.5
spatial	29.6	33.3	7.4	14.8	14.8
benchmark-peer	7.4	66.7	18.5	0.0	7.4
benchmark-national	22.2	66.7	7.4	0.0	3.7

Table 3. Collection of performance measures – Parks and Recreation

The results of cluster analysis across the four departments are reported in Figures 3 and 4 (Again due to space constraints, we are not including all cluster analysis results). The results indicate that two groups have emerged from the 54 performance measures. For Police departments, we find that cluster one includes Goodyear, Phoenix, Mesa, and 11 other departments. Table 4 reports the characteristics of the two clusters. The results reported in Table 4 show that Police departments in cluster one rarely fewer different types of performance measures.

Cluster two includes Brisbane, Arlington, Scottsdale and 10 other departments (Figure 3). As reported in Table 4, we find that Police departments in cluster two are the frequent users of performance measures, routinely collecting workload (92.3 percent), cost efficiency (38.5 percent), and outcomes (84.6 percent) measures.

Figure 3. Police department cluster analysis

		worklo ad	cost efficien cy	respons iveness	outcom es
Cluster 2	Never	0.0	23.1	0.0	0.0
	Once/ Twice a year	0.0	30.8	0.0	0.0
	Three/ Four a year	7.7	0.0	0.0	0.0
	Monthly	0.0	7.7	100.0	15.4
	Routinely	92.3	38.5	0.0	84.6
Cluster 1	Never	0.0	30.8	0.0	0.0
	Once/ Twice a year	0.0	38.5	46.2	30.8
	Three/ Four a year	7.7	23.1	38.5	30.8
	Monthly	61.5	7.7	15.4	38.5
	Routinely	30.8	0.0	0.0	0.0

Table 4. Police performance measures between two clusters (Limited to few categories due to space constraints)

For Parks and Recreation departments, we find that cluster one, the less frequent users, includes Dallas, Mesa, Phoenix, and 8 other departments (Figure 4). More frequent users of performance measures in cluster two include Dublin, Durham, Goodyear, and 13 more departments (Figure 4). As Table 5 shows, departments in cluster two tend to collect workload (68.8 percent), cost efficiency (43.3 percent), and outcomes (43.8 percent) measures routinely.

Figure 4. Parks and Recreation department cluster analysis

		worklo ad	cost efficien cy	respons iveness	outcom es
Cluster 2	Never	6.3	6.3	12.5	0.0
	Once/ Twice a year	6.3	0.0	6.3	0.0
	Three/ Four a year	6.3	43.8	12.5	37.5
	Monthly	12.5	6.3	68.8	18.8
	Routinely	68.8	43.8	0.0	43.8
Cluster 1	Never	9.1	0.0	0.0	9.1
	Once/ Twice a year	9.1	27.3	45.5	45.5
	Three/ Four a year	18.2	0.0	45.5	18.2
	Monthly	45.5	72.7	9.1	18.2
	Routinely	18.2	0.0	0.0	9.1

Table 5. Parks and Recreation performance measures between two clusters

4.1. QCA Results

QCA is a set-based theoretical approach that combines qualitative (variable-based) and quantitative (case-based) techniques to provide greater insights. It describes relationships among variables in terms of set membership [12]. Thus, a condition variable (X) explains the outcome variable (Y). Further, QCA also identifies several condition variables (X*Z) that combine to explain the outcome variable (Y). The use of QCA approach differs from regression approaches because it can recommend how a *combination* of condition variables explains an outcome variable [28]. The use of QCA approach is more suited to organizational studies to understand how different combinations can help achieve an outcome [2].

QCA analysis is conducted in three steps [32]. First, the raw data is converted into categorical cases 0 and 1. Second, a truth table is derived, which represents all possible configurations that explain the outcome. Finally, the outcomes of the truth table are reduced to Boolean minimization. Boolean minimization provides configurations that cannot be achieved by another configuration.

As discussed in the previous section, the results of cluster analysis show that for each of the four departments, the cases can be divided into two groups: the frequent users (cluster two) and less frequent users (cluster one). We coded departments in cluster two as

1 and departments in cluster one as 0. We use crisp set QCA (csQCA) to run the analysis.[1]

We also dichotomized condition variables. In this study, we use five condition variables. We combined three factors to construct an innovation culture variable. Then, we used csQCA calibration to dichotomize innovation culture, i.e., highly innovative departments were coded as 1 and 0 otherwise. If a department provided staff training, the variable was coded as 1 and 0 otherwise. If a department had own data staff, the variable was coded as 1 and 0 otherwise. Adequate IT infrastructure was coded as 1 if a department answered strongly agree or agree that they have sufficient hardware and software to support performance data and 0 if they responded - strongly disagree, disagree or neutral. If a department consulted elected officials about performance measures review, the variable was coded as 1 and 0 otherwise.

Once the raw data was converted into categorical variables, we tested the dataset for necessity and sufficiency conditions. Testing for whether a variable is necessary and sufficient for producing the outcome variable is a critical step before constructing a truth table [32]. Based on the results of the truth table, we derived Boolean minimization.

The results of the csQCA are presented in Tables 6 through 10. To present csQCA finding, we use Ragin and Fiss notation system [30]. The black shaded circles (● or •) indicate the presence of a condition. The crossed-out circles indicate the absence of a condition (⊗ or ⊗). The large circles represent core conditions, and smaller circles represent peripheral conditions. Blank spaces denote that the causal condition may be present or absent. Raw coverage provides comparative dominance of each configuration and consistency provides how a given configuration displays the outcome [30, 32].

QCA results for Police departments are reported in Table 6 (please note we included non-essential solution sets that were present with essential solution sets). The results indicate that no condition variable dominates the outcome. For instance, innovative culture is present in three (one peripheral) out of five configurations. However, condition variables such as availability of data staff and adequate IT infrastructure are not present in two configurations but present in one. Interestingly, elected officials review of performance measures is not

[1] An excellent comparison between crisp set and fuzzy set QCA in the public sector can be found in Dawson, Denford and Desouza 2016 Journal of Strategic Information Systems [6].

present in three configurations and only present in two (one peripheral).

	1	2	3	4	5
Innovative culture	⊗	●	•		●
Provision of staff training		●		●	⊗
Availability of data staff			⊗	⊗	●
Adequate IT infrastructure	⊗	●	⊗		
Elected officials review	●	•	⊗	⊗	⊗
Consistency	1	1	1	1	1
Raw coverage	0.08	0.25	0.17	0.08	0.08
Unique coverage	0.08	0.25	0.00	0.00	0.00
Overall consistency	1				
Overall coverage	0.58				

Table 6. Results of QCA – Police

Similarly, QCA results for Code Compliance departments indicate that no condition variable dominates the outcome (See Table 7). For instance, innovative culture is present in two (one peripheral) configurations and not present in one. Providing staff training is only present in two configurations out of five. However, condition variables such as availability of data staff and elected officials review of performance measures are not present in two configurations but present in one. The variable adequate IT infrastructure is not present in three configurations, and is only present in two (one peripheral).

	1	2	3	4
Innovative culture		•	⊗	●
Provision of staff training		●	●	
Availability of data staff	⊗	●		⊗
Adequate IT infrastructure	⊗	⊗	•	●
Elected officials review	⊗		⊗	●
Consistency	1	1	1	1
Raw coverage	0.44	0.22	0.22	0.11
Unique coverage	0.33	0.22	0.11	0.00
Overall consistency	1			
Overall coverage	0.89			

Table 7. Results of QCA – Code Compliance

The results of QCA for Parks and Recreation departments indicate the availability of data staff dominates the configurations (See Tables 8 and 9. Due to formatting issues, the results of QCA are presented in two tables). Other conditions variables such as innovative culture, staff training, adequate IT infrastructure and elected officials review of

performance measures are present in some configurations (including peripheral) and not present in others. Therefore, the availability of data staff is an essential condition that promotes the collection of performance measures in Parks and Recreation departments.

	1	2	3	4	5
Innovative culture	•	•	⊗	●	⊗
Provision of staff training	●		•		⊗
Availability of data staff		•	●	●	●
Adequate IT infrastructure	⊗	●	⊗	●	•
Elected officials review	●	⊗		⊗	●
Consistency	1	1	1	1	1
Raw coverage	0.25	0.31	0.13	0.31	0.06
Unique coverage	0.19	0.00	0.00	0.00	0.00
Overall consistency	1				
Overall coverage	0.69				

Table 8. Results of QCA – Parks and Recreation

	6	7	8	9
Innovative culture	⊗		⊗	⊗
Provision of staff training	⊗	⊗	●	●
Availability of data staff				●
Adequate IT infrastructure		●	⊗	
Elected officials review	●	●		⊗
Consistency	1	1	1	1
Raw coverage	0.06	0.06	0.13	0.06
Unique coverage	0.00	0.00	0.00	0.00
Overall consistency	1			
Overall coverage	0.69			

Table 9. Results of QCA – Parks and Recreation (Contd.)

For Public Works departments, the results of QCA indicate that no condition variable dominates the outcome (See Table 10). For instance, innovative culture is present in two configurations and not present in one. Providing staff training is not present in two configurations out of six. Out of six, the variable availability of data staff is not present in three configurations. The variables adequate IT infrastructure, and elected officials review of performance measures are not present in three configurations, but present in three (including two peripherals). Thus, the results highlight that for all four departments, different configurations can facilitate the collection of performance measures.

	1	2	3	4	5	6
Innovative culture		●	●	⊗		
Provision of staff training			⊗			⊗
Availability of data staff	⊗	⊗			⊗	
Adequate IT infrastructure	●	•	•	⊗	⊗	⊗
Elected officials review	●	•	•	⊗	⊗	⊗
Consistency	1	1	1	1	1	1
Raw coverage	0.2	0.4	0.4	0.1	0.1	0.1
Unique coverage	0.2	0	0	0	0	0
Overall consistency	1					
Overall coverage	0.7					

Table 10. Results of QCA – Public Works

5. Discussion and Conclusion

In this paper, we examined (1) what types of performance measures are collected by US local governments and (2) how organizational, technical, and external factors influence the collection of performance measurement. We investigated how innovation culture, staff training, availability of data staff, adequate IT infrastructure, and pressure by elected officials influence the collection of performance measurement. To test our theoretical framework, we adopted a unique approach. First, we conducted cluster analysis to determine how departments group based on 54 performance indicator variables. The results of cluster analysis divided departments into two groups. Second, we conducted QCA to examine how different organizational and external factors influence the collection of performance measurement. The findings of our QCA analysis reveals that no single condition influence the collection of performance measurement. Public agencies can adopt different organizational and external configurations to improve the collect performance measures.

Before discussing the contributions of this study, we highlight several limitations. This study utilizes survey data collected from cities in the Alliance for Innovation network. It may be possible that certain types of cities self-select into the network. Future studies should collect information from cities beyond the Alliance for Innovation network. Moreover, this study uses QCA to provide a holistic view of the collection of performance measures in the public sector. However, QCA method is new and evolving. At

the same time, it should be noted that many studies in information systems and public administration are starting to utilize QCA [6, 7]. Despite these limitations, this study provides valuable insights about the collection of performance measures in US local governments.

The findings of our study offer several contributions to research and practice. First, our study collected data on a broad range of performance measures. Previous studies have noted that no single performance measure can completely capture the complexity in the public sector. At the same time, many local governments do not have the resources or the need to collect a wide range of performance measures. The four departments in our study are more likely to collect workload, compliant, and public sentiment measures. However, other performance measures (e.g., trends, peer benchmarks) are less regularly collected.

Second, we also find that there is a considerable difference among departments in their collection of performance measures. For instance, we find that Police departments tend to collect workload and outcomes measure more routinely than Parks and Recreation departments while Parks and Recreation departments more frequently collect responsiveness measures than Police departments. These differences can be attributed to the types of services provided by these departments, such as whether they are reliant on discretionary user support and participation. Moreover, as our results show that it is important to differentiate departments based on the type of service they provide. Differentiating departments based on their services may help public managers collect better performance measures to meet the demands of the public and provide better quality services.

Third, our results indicate that Police departments are clearly more extensive in their collection of performance measures. They frequently collect wide range of performance measures (e.g. spatial, trend) compared to Parks and Recreation departments in our sample. The findings highlight that Police departments who provide coercive and highly technical service are collecting most amount of performance measures. One reason for this could be attributed to the nature of service provided by Police departments. Police departments often face high pressures from public, elected officials, and media. Often, they became targets of media attention based on their performance. Moreover, as Police departments provide coercive services, they are more likely to be aware of negative feedback from the public. As a result, it is likely that Police departments collect data to highlight their performance on a routine basis for providing quality

services and evidence-driven response to any criticism from the public, media and elected officials.

Fourth, we also find that Parks and Recreation departments collect cost-efficiency measures more frequently compared to Police departments. This finding can again be interpreted in terms of the type of service delivered. It is possible that Parks and Recreation departments - being non-coercive and non-technical service providers - are likely to focus on performance measures to improve service delivery. Thus, these departments are more likely to collect cost-efficiency measures to understand the quality of service delivered and cost of service delivery. Further, as citizens have a choice to choose whether they use the services provided by Parks and Recreation departments, it is logical for these departments to focus on providing efficient services to meet the needs of the citizens.

Finally, our study utilizes QCA to provide a more in-depth view of performance measures. The use of QCA allows us to capture multiple configurations that impact the collection of performance measures. The results of QCA highlight that multiple configurations can help public agencies collect performance measures. Previous studies on performance measures often focus on the influence of a single variable on the collection of performance measures. However, public agencies differ significantly in their social and technical contexts. The findings of this study suggest that public agencies should examine social and technical contexts of their organization to arrive at the best performance measurement collection strategy.

6. Acknowledgments

7. References

[1] Ammons, D.N., and W.C. Rivenbark, "Factors influencing the use of performance data to improve municipal services: Evidence from the North Carolina benchmarking project", Public Administration Review 68(2), 2008, pp. 304–318.
[2] Andrews, R., M.J. Beynon, and A.M. McDermott, "Organizational capability in the public sector: A configurational approach", Journal of Public Administration Research and Theory 26(2), 2015, pp. 239–258.

[3] Brown, T., "Coercion versus choice: Citizen evaluations of public service quality across methods of consumption", Public Administration Review 67(3), 2007, pp. 559–572.

[4] Chavent, M., V. Kuentz, B. Liquet, and L. Saracco, "ClustOfVar: An R Package for the clustering of variables", arXiv:1112.0295 [stat], 2011.

[5] Davenport, T.H., "Competing on analytics", Harvard Business Review 84(1), 2006, pp. 98-106.

[6] Dawson, G.S., J.S. Denford, and K.C. Desouza, "Governing innovation in U.S. state government: An ecosystem perspective", The Journal of Strategic Information Systems 25(4), 2016, pp. 299–318.

[7] Dawson, G.S., J.S. Denford, C.K. Williams, D. Preston, and K.C. Desouza, "An examination of effective IT governance in the public sector using the legal view of agency theory", Journal of Management Information Systems 33(4), 2016, pp. 1180–1208.

[8] Dean, D., "Citizen ratings of the police: The difference contact makes", Law & policy 2(4), 1980, pp. 445–471.

[9] Desouza, K.C., Realizing the Promise of Big Data, IBM Center for the Business of Government, 2014.

[10] Desouza, K.C., Creating a Balanced Portfolio of Information Technology Metrics, IBM Center for the Business of Government, 2015.

[11] Emery, F.E., and E.L. Trist, Towards a social ecology: Contextual appreciations of the future in the present, Plenum Publishing Corporation, New York, 2012.

[12] Fiss, P.C., "A set-theoretic approach to organizational configurations", Academy of Management Review 32(4), 2007, pp. 1180–1198.

[13] Fountain, J.R., T.K. Patton, and H.I. Steinberg, "Improving performance reporting for government: New guidance and resources", The Journal of Government Financial Management 53(1), 2004, pp. 60-66.

[14] Galili, T., "dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering", Bioinformatics 31(22), 2015, pp. 3718–3720.

[15] Glaser, M.A., and J.W. Bardo, "A five-stage approach for improved use of citizen surveys in public investment decisions", State and Local Government Review 26(3), 1994, pp. 161–172.

[16] Ho, A.T.-K., "Perceptions of performance measurement and the practice of performance reporting by small cities", State and Local Government Review 35(3), 2003, pp. 161–173.

[17] Ho, A.T.-K., "Accounting for the value of performance measurement from the perspective of Midwestern mayors", Journal of Public Administration Research and Theory 16(2), 2006, pp. 217–237.

[18] Ho, A.T.-K., and A.Y. Ni, "Have cities shifted to outcome-oriented performance reporting? A Content Analysis of City Budgets", Public Budgeting & Finance 25(2), 2005, pp. 61–83.

[19] Julnes, P. de L., and M. Holzer, "Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation", Public Administration Review 61(6), 2001, pp. 693–708.

[20] Krishnamurthy, R., and K.C. Desouza, "Big data analytics: The case of the Social Security Administration", Information Polity 19(3, 4), 2014, pp. 165–178.

[21] Melkers, J., and K. Willoughby, "Models of performance-measurement use in local governments: Understanding budgeting, communication, and lasting effects", Public Administration Review 65(2), 2005, pp. 180–190.

[22] Melkers, J.E., K.G. Willoughby, B. James, J. Fountain, and W. Campbell, Performance measurement at the state and local levels: A summary of survey results, GASB, Hartford, CT, 2002.

[23] Messatfa, H., L.C. Reyes, and M. Schroeck, The power of analytics for public sector: Building analytics competency to accelerate outcomes, IBM, 2011.

[24] Miller, T.I., and M.A. Miller, "Standards of excellence: U. S. residents' evaluations of local government services", Public Administration Review 51(6), 1991, pp. 503–514.

[25] Moore, M.H., "Privatizing public management", Market-based governance: Supply side, demand side, upside, and downside, 2002, pp. 296–322.

[26] Moynihan, D.P., "Why and how do state governments adopt and implement 'Managing for Results' reforms?", Journal of Public Administration Research and Theory 15(2), 2004, pp. 219–243.

[27] Newcomer, K.E., "Using performance measurement to improve programs", New directions for evaluation 1997(75), 1997, pp. 5–14.

[28] Pavlou, P.A., and O.A. El Sawy, "From IT leveraging competence to competitive advantage in turbulent environments: The case of new product development", Information Systems Research 17(3), 2006, pp. 198–227.

[29] Poister, T.H., and G.T. Henry, "Citizen ratings of public and private service quality: A comparative perspective", Public Administration Review, 1994, pp. 155–160.

[30] Ragin, C.C., Redesigning social inquiry: Fuzzy sets and beyond, University of Chicago Press, Chicago, 2008.

[31] Skogan, W.G., "Citizen satisfaction with police encounters", Police Quarterly 8(3), 2005, pp. 298–321.

[32] Thiem, A., Qualitative comparative analysis with R a user's guide, Springer, New York, 2013.

[33] Wholey, J.S., "Performance-based management: Responding to the challenges", Public Productivity & Management Review, 1999, pp. 288–307.

[34] Yang, K., and J.Y. Hsieh, "Managerial effectiveness of government performance measurement: Testing a middle-range model", Public Administration Review 67(5), 2007, pp. 861–879.

[35] Government analytics, IBM, 2010.

[36] "Hierarchical clustering / dendrograms", In NCSS Statistical Software, 445-1-445–15.