Open Government Data: The Key to Promoting Public Participation, and Fighting Against Corruption

Thalinee Sangkachan*

Abstract

Poor governance resulting from corruption is a chronic problem for the public sector. Although there have been efforts to involve citizens to participate in dealing with this issue, they tend to avoid such participation since they do not have sufficient data and have to face data asymmetry that may debase their opinions or arguments. However, open government data (OGD) are now being widely implemented around the world and are expected to produce a variety of positive impacts on society. This study investigates the effects of OGD on empowering public participation (PP) and enhancing corruption control (CC). Based on the data analysis of 113 countries using partial least squares-structural equation modeling (PLS-SEM), this study found that OGD have a positive significant effect on PP. Further, OGD can significantly increase and make more effective the efforts to CC. Indeed, OGD have empowered citizens to increasingly participate in CC. Given the importance of OGD, countries should be aware of the development of OGD and capitalize on its benefits and values.

Keywords: Open government data, public participation, corruption control

* Graduate School of Public Administration, National Institute of Development Administration (NIDA).
E-mail: thalineeofficial@gmail.com
Received: January 19, B.E.2564. Revised: September 5, B.E.2564. Accepted: September 5, B.E.2564.

ข้อมูลเปิดของภาครัฐ: กุญแจสำคัญในการส่งเสริมการมีส่วนร่วมของประชาชนและการต่อสู้กับการทุจริตคอร์รัปชัน

ฐาลินี สังฆจันทร์*

บทคัดย่อ

การบริหารปกครองที่ย่ำแย่อันเป็นผลมาจากการทุจริตคอร์รัปชันถือเป็นปัญหาเรื้อรังของภาครัฐ แม้ว่าจะมีความพยายามให้ประชาชนมีส่วนร่วมในการจัดการกับปัญหาดังกล่าว แต่ประชาชนก็มักจะหลีกเลี่ยงการเข้าร่วมเนื่องจากพวกเขาไม่มีข้อมูลที่เพียงพอและต้องเผชิญกับความไม่สมดุลของข้อมูล ซึ่งอาจทำให้ความคิดเห็นหรือข้อโต้แย้งของพวกเขาถูกลดคุณค่าลงไป อย่างไรก็ตาม ปัจจุบัน ข้อมูลเปิดของภาครัฐกำลังได้รับการดำเนินการอย่างแพร่หลายทั่วโลกด้วยความคาดหวังว่าข้อมูลดังกล่าวนี้จะเอื้ออำนวยให้เกิดผลกระทบเชิงบวกในแง่มุมต่าง ๆ ต่อสังคม การวิจัยนี้จึงมุ่งศึกษาผลกระทบของข้อมูลเปิดของภาครัฐที่มีต่อการเสริมสร้างการมีส่วนร่วมของประชาชนและการเพิ่มประสิทธิภาพในการควบคุมการทุจริตคอร์รัปชัน จากการวิเคราะห์ข้อมูลของ 113 ประเทศ ด้วยโมเดลสมการโครงสร้างแบบกำลังสองน้อยที่สุดบางส่วน ค้นพบว่า ข้อมูลเปิดของภาครัฐมีผลกระทบเชิงบวกอย่างมีนัยสำคัญต่อการมีส่วนร่วมของประชาชน นอกจากนี้ ข้อมูลดังกล่าวยังสามารถเพิ่มพูนความพยายามในการควบคุมการทุจริตคอร์รัปชันได้อย่างมีประสิทธิภาพ โดยเฉพาะอย่างยิ่ง ข้อมูลเปิดของภาครัฐได้ส่งเสริมให้ประชาชนมีส่วนร่วมในการควบคุมการทุจริตคอร์รัปชันมากขึ้นอย่างมีนัยสำคัญ ด้วยความสำคัญนี้ ประเทศต่าง ๆ ควรได้ตระหนักถึงการพัฒนาข้อมูลเปิดของภาครัฐและใช้ประโยชน์จากข้อมูลดังกล่าวนี้ให้มากยิ่งขึ้น

คำสำคัญ: ข้อมูลเปิดของภาครัฐ การมีส่วนร่วมของประชาชน การควบคุมการทุจริตคอร์รัปชัน

---

* คณะรัฐประศาสนศาสตร์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

อีเมล: thalineeofficial@gmail.com

วันที่รับบทความ: 19 มกราคม 2564 วันที่แก้ไขบทความล่าสุด: 5 กันยายน 2564 วันที่อนุมัติการตีพิมพ์: 5 กันยายน 2564

Introduction

Poor governance is a complex and chronic problem for public administration. The authority's misuse of private advantages, also known as corruption, is the most ancient disease that has plagued bureaucracy for centuries (Awan et al., 2018; Guriev, 2004). The corruption issue has gained center stage in a country's development because it can determine whether or not a government successfully drives national development (Faisal & Jafri, 2017). Corruption has many negative consequences; it ruins democracy, exploits wealth, decreases trust, devastates economic development and growth, and poses a grave threat to human rights (Bhayani, 2013; Desta, 2006; United Nations Development Programme, 1997). The global community has made intensive efforts to solve the problem, such as through public participation (hereinafter referred to as PP), in the hope that it solve corruption (Hays & Kogl, 2007; World Bank, 2018). However, government-initiated PP is often unsuccessful because a government and citizens have unequal power, capabilities, knowledge, information, or data (Bartenberger & Grubmüller, 2014; Steele, Murnane & Willett, 2010). In other words, citizens might be skeptical of participation because they lack these resources, power, knowledge, information, or data (Ansell & Gash, 2007). Citizens, who do not have sufficient knowledge, who face information asymmetry, and who lack data, tend to not participate with others or with their government (Daniel & Habsari, 2019).

Data are an important fundamental resource required throughout all stages of public participation. Currently, many governments have begun to adopt open government data (hereinafter referred to as OGD). Once government data are open to the public, information and knowledge asymmetries can be reduced (Bertot, Jaeger & Grimes, 2010; Halonen, 2012; Janssen, 2012; Kassen, 2013; Peled, 2013; Zuiderwijk et al., 2012). This movement also increases opportunities for citizens to generate new knowledge, besides encouraging them to expand and accumulate knowledge using reliable and available data (Lee, Ham, & Choi, 2016). It also helps promote better, meaningful, and insightful interaction and communication between a government and citizens so that both can work together to better society (Hays & Kogl, 2007; Jetzek, Avital, & Bjørn-Andersen, 2013; Odongo & Rono, 2016). OGD also lead to more effective decision-making (Kassen, 2013). Additionally, the public can also monitor and evaluate the government's activities and performance via OGD. OGD shift the role of citizens from passive to active or proactive citizens that are more aware of their government's policies and programs. Moreover, OGD are posited to not only boost PP, but also, in turn, reduce opportunities for corruption, and potentially reinforcing efforts to enhance corruption control (hereinafter referred to as CC) (Darusalam et al., 2019; Florez & Tonn, 2019; Hulstijn, Darusalam, & Janssen, 2018; Janssen & Zuiderwijk, 2012)

However, although OGD can be a significant driver and enabler of anti-corruption, its potential is yet to be fully realized or leveraged in the fight against corruption (Vrushi & Hodess, 2017). OGD might not be sufficient to fight corruption on its own (Florez & Tonn, 2019). Previous studies found that OGD did not immediately or directly reduce corruption (Darusalam et al., 2019; Hulstijn et al., 2018; Iglesias, 2017; Park & Kim, 2019). Astonishingly, Vrushi and Hodess (2017) found that OGD policy and anti-corruption policy improved independently, although the two policies were highly correlated. Another perspective proposed that OGD can contribute to anti-corruption reforms by influencing institutions and PP (Davies & Perini, 2016). Not only does OGD reveal corruption cases, but it also prevents corruption behavior by empowering the civil sector to mobilize against corruption. However, some have argued that OGD still has a weak effect on anti-corruption because the public’s use of data is low and only a few stakeholders use OGD (Florez & Tonn, 2019). Iglesias (2017) and Florez and Tonn (2019) have proposed that the public (PP) should work with OGD to better deal with corruption. PP could decipher the link between OGD and those that work on anti-corruption. The relationships among these issues should be further studied (Iglesias, 2017; Florez & Tonn, 2019); little is still known about how these issues are linked or how they influence each other.

Based on previous studies, this study is, therefore, an early attempt to investigate the relationships of the major contemporary public policy issues: OGD, PP, and CC. This paper will contribute to remedying the persistent problem of poor governance as OGD are now perceived as an essential strategic catalyst for anti-corruption (Gigler, Custer & Rahemtulla, 2011; Hartog et al., 2014; OECD, 2018) and as a novel standard for good governance (Hartog et al., 2014; OECD, 2018) in the 21st century. If a country can harness and unleash the power of OGD per its potential, OGD would then indeed improve and enhance governance (Gigler et al., 2011).

Literature Review and Hypothesis Development

As mentioned, there has been some concern that OGD might not be sufficient to fight corruption on its own (Florez & Tonn, 2019), and that its potential has not yet been fully realized or leveraged in the fight against corruption (Vrushi & Hodess, 2017), or OGD does not immediately or directly reduce corruption (Darusalam et al., 2019; Hulstijn et al., 2018; Iglesias, 2017; Park & Kim, 2019). This research thus applied the information systems success (ISS) model developed by Delone and McLean (1992; 2003) to draw a conceptual framework of the relationships among OGD, PP, and CC since the ISS model aims to provide a comprehensive understanding of information systems in terms of input, process, and output. ISS is a very specific model for evaluating the application of information technology. The elements of this model included system quality, information quality, service quality, service quality, system use, user satisfaction, and net benefits. The content of this research is related to the ISS model, as shown in Table 1.

Table 1. Application of ISS Model Framework for this Research

	Input(s)	Process	Output
ISS model	system quality, information quality, service quality	system use, user satisfaction	net benefits
Research framework	OGD	PP	CC

The OGD in this research are similar to the information quality element of the ISS model since they are perceived as the input in the system. PP has implications for the use of OGD similar to system usage in the ISS model. Further, CC can be perceived as the output of the system since it is the result of data use and can be perceived in terms of net benefits as recommended by the ISS model. Based on the comprehensive framework, the hypothesis development is as follows.

As poor governance is a complex and chronic problem, one or even two approaches might not be sufficient to achieve better governance. Although governance is how authorities exercise control over a system (Ionescu, 2013), authorities can still misuse the advantages given to them, otherwise known as corruption. Corruption is the most ancient disease affecting bureaucracy and is, therefore, an essential indicator of governance failure (Awan et al., 2018; Guriev, 2004). Poor governance can result in poor countries and negative growth. Empirical research has identified corruption as an impediment that negatively affects governance (Kumar, 2004; Tshepo, 2015). With the calamities resulting from corruption, governance practices can become more difficult to achieve. Poor governance can also be exacerbated because of the failure to fight corruption (Faisal & Jafri, 2017).

Some solutions to this problem have been identified, one of which is PP, a subject that has been studied for a long time in the hopes that it can solve the problem. PP is defined as the interactions among citizens and governments (World Bank, 2018) and the interaction between citizens (Hays & Kogl, 2007) in public or government activities. It is a two-way approach in which citizens take part in and contribute their opinions and knowledge for public policy decision-making, discussing and monitoring government activities, as well as the development of a better society (Mellouli, Luna-Reyes & Zhang, 2014). Previous studies have shown that citizen participation can help solve corruption; in particular, PP can positively impact social and economic aspects through the policy decision-making process (Hays & Kogl, 2007). The above actions contribute significantly to the public policy-making process, making it more transparent and highlighting governments' fulfillment of their responsibility to the people, which, in turn, enhances CC. Thus, this study hypothesizes that:

H1: Public participation positively affects the level of corruption control.

However, government-initiated PP is often unsuccessful because of the unequal power, capabilities, knowledge, and information among the government and the people (Bartenberger & Grubmüller, 2014; Evans & Campos, 2013). In other words, citizens might be skeptical of participating publicly because they lack the resources, power, knowledge, and data, as mentioned above (Ansell & Gash, 2007). Stakeholders that do not have sufficient knowledge or data will not participate with others or with the government in order to right the wrongs in society; they might even be manipulated by stronger actors. Empirical research (Daniel & Habsari, 2019) has found that information asymmetry or a lack of data can also hamper the PP process. However, since OGD are currently developing rapidly and becoming more popular, this study aims to contribute to the existing body of knowledge by investigating the effect of OGD on PP.

In order to deal with the critical problems of PP, many governments have begun to adopt the open government concept, which requires them to open their data to the public (Bertot et al., 2010; Halonen, 2012; Janssen, 2012; Kassen, 2013; Peled, 2013; Zuiderwijk et al., 2012). OGD refer to complete, timely, machine-readable, and freely accessible data (Kalampokis, Tambouris & Tarabanis, 2011; Kassen, 2013; World Wide Web Foundation, 2016). Per the PP concept (Bishop & Davis, 2002), data are an important resource required throughout all stages of participation. The OGD concept implies that citizens have the right to access government data to gain information and knowledge about its policies and performance (Bartenberger & Grubmüller, 2014). By opening government data to the public, information and knowledge asymmetries are expected to decrease. In other words, OGD increase the opportunities for citizens to generate new knowledge and expand and accumulate existing knowledge (Lee et al., 2016). Further, OGD help promote better interaction and communication between among governments and citizens, in turn enhancing their teamwork in order to better society (Odongo & Rono, 2016). This strengthened citizen participation is achieved by empowering many stakeholders to monitor and evaluate the government’s activities and performance.

According to Davies (2012), OGD can give citizens a stronger voice politically (i.e. informing citizens as voters, supporting public debate), collaboratively (i.e. co-produced planning, public services, and information), and service-wise (i.e. improving the quality of services and competitive innovation). According to Jetzek et al. (2013), OGD can generate meaningful collaborative and participative governance, as the citizens can contribute to the development of policies and programs by voicing their opinions via a government platform that would enable them to contribute. This action can reduce the barriers to collaboration and participation and encourage citizens to become more involved. Moreover, Kassen (2013) empirically investigated an OGD project in Chicago (Web portal: data.cityofchicago.org) and found that the data portal offered citizens opportunities to participate and engage in online discussions regarding social problems and how to solve them. Citizens could also provide interactive feedback and even recommend new suitable datasets. In this case, the citizens are allowed to contribute and interact with the government and obtain sufficient information and knowledge from reliable and available data.

OGD are expected to transform citizens from being passive to active or proactive ones, who contribute to the development of policies and programs. OGD can provide valuable information that enables citizens to effectively contribute to the decision-making process (Kassen, 2013). With sufficient data, active or proactive citizens will be more critical, analytical, and more interested in participating in the government policy or program decision-making process compared to passive citizens. Additionally, the citizens will be more interested in collaborating with their governments to identify problems, question public issues, seek solutions to the problems and needs, as well as provide important feedback (Hays & Kogl, 2007). Moreover, informed-active or proactive citizens effectively contribute to the democratic process as well. For the above reasons, citizens should be given OGD accessibility and availability to empower them to meaningfully and intelligently participate in government decision-making and operations. Therefore, this study hypothesizes that:

H2: Open government data positively affect the level of public participation.

Not only do OGD boost PP, but they also widely increase the public acceptance of OGD through enhanced transparency, increased accountability, reduced opportunities for corruption, as well as potentially reinforcing efforts to increase CC (Darusalam et al., 2019; Florez & Tonn, 2019; Hulstijn et al., 2018; Janssen & Zuiderwijk, 2012). Empirical research has shown that OGD can help boost CC by providing citizens with access to government data and therefore empowering them to take the necessary actions (Hulstijn et al., 2018; Rajshree & Srivastava, 2012). Further, a systematic review by Safarov, Meijer and Grimmelikhuijsen (2017) found that anti-corruption can be achieved with the implementation of OGD. Actions that are done secretly might lead to corruption, and corruption might occur when data are not released to the public. Hence, OGD are now perceived as a powerful tool for increasing public awareness while also reducing the misuse and ineffectiveness of resources due to corruption. Additionally, other studies (Renata, 2017; Vrushi & Hodess, 2017) have found a positive correlation between CC and OGD. Kim, Kim and Lee (2009) found that as part of a long-term effort (1988-2007) to fight corruption, Korea began to disclose administrative procedures and data to the public. Therefore, it can be assumed that the higher the availability of OGD, the lower is the level of corruption. When government data are disclosed, experts and the public can access and use such data (such as procurement data) to identify any risk factors that might lead to corruption. In turn, a system to prevent corruption could be developed (Izdebski, 2015). Based on the above empirical studies, this study hypothesizes that:

H3: Open government data positively affect the level of corruption control.

However, the relationship mentioned in H3 has been subject to much debate. OGD can significantly drive and enable anti-corruption but it has yet to be realized and leveraged sufficiently to fight corruption (Vrushi & Hodess, 2017). Research on 95 case studies (Goodrich, 2015) revealed that OGD correlated with 7 percent of the corruption behavior in some cases. Still, there are challenges in identifying and measuring the impact of OGD on anti-corruption (Florez & Tonn, 2019). Other works viewed OGD as not sufficient to fight corruption on its own. For example, a recent study by Iglesias (2017) found that OGD did not immediately reduce corruption, so investigations are still ongoing. Hulstijn et al. (2018) also found the same results. A longitudinal study (Park & Kim, 2019) found that OGD as part of the functions of an open government did not directly reduce corruption. One empirical work (Darusalam et al., 2019) also found that OGD on its own was not sufficient to reduce corruption. Further, Vrushi and Hodess (2017) investigated five countries and found that the development of OGD policy and anti-corruption policy in these countries improved independently, although the two policies were highly correlated. The study indicated that a lack of coordination between the two policies reduced the opportunity to effectively resolve corruption. If both were linked, the policies could more likely achieve their intended results.

Given the critical challenges mentioned above, Davies and Perini (2016) suggested that OGD can contribute to anti-corruption reforms by influencing institutions and PP. Past empirical research (Florez & Tonn, 2019) found that OGD has a weak impact on anti-corruption because the public does not use the data; only a few stakeholders do. Opening and releasing data allows experts, stakeholders, and citizens to access and identify risk factors. OGD not only reveal corruption cases but also prevent corruption behaviors by empowering the civil sector to mobilize against corruption. Past studies (Darusalam et al., 2019; Florez & Tonn, 2019; Hulstijn et al., 2018; Vrushi & Hodess, 2017) have suggested that citizens, such as NGOs, the media, journalists, pressure groups, branch organizations, watchdogs, regulators, parliaments, experts, auditors, and society as a whole, step in and fill the gap between OGD and anti-corruption by accessing, using, and analyzing such data to fight corruption. That is, OGD can empower citizens to monitor government activity and performance. The data also help specify the audience and purpose for exposing, monitoring, and controlling corruption. Iglesias (2017) and Florez and Tonn (2019) suggested strengthening the capacity of civil society or PP to work with OGD as a strategy for dealing with corruption. This strategy could be maximized if the link among those that work with OGD and those that work on anti-corruption is clearly defined. OGD might reduce corruption by prompting the public to actively participate in and use OGD. Further, the government and civil society should be aware of the benefits of using such data to solve corruption together (Vrushi & Hodess, 2017). In this case, PP might help delineate the link between OGD policies and anti-corruption. Based on the literature discussed herein, and given the posited relationship, it is interesting to determine whether PP positively mediates the relationship between OGD and CC. Therefore, this study hypothesizes that:

H4: Public participation positively mediates the relationship between open government data and corruption control.

Research Methodology

Data Collection

Secondary data based in 113 countries were retrieved from international organizations, including the World Wide Web Foundation (2016), the United Nations (2016), the Economist Intelligence Unit (2017), Transparency International (2016). In order to maintain the causality, secondary data from these sources were obtained from 2016. Table 2 summarizes the proportion of countries in greater detail.

Table 2. The Countries Categorized by the Regions

Region	Frequency (Percentage)
East Asia and Pacific	12 (10.6)
Europe and Central Asia	40 (35.4)
Latin America and Caribbean	19 (16.8)
The Middle East & North Africa	12 (10.6)
North America	2 (1.8)
South Asia	4 (3.5)
Sub-Saharan Africa	24 (21.2)
Total	**113 (100)**

Measurement of Variables

All of the variables in this study are formative constructs. The descriptions of each construct are as follows. First, OGD refer to "complete, timely, machine-readable, and freely accessible data" (Kalampokis, Tambouris & Tarabanis, 2011; Kassen, 2013; World Wide Web Foundation, 2016). It is an independent variable measured using five dimensions, i.e. (i) completeness of data (OGDC) ("do the data exist?"); (ii) timeliness of data (OGDT) ("Is the dataset being kept regularly updated?"); (iii) machine readability (OGDM) ("Is the dataset provided in a machine-readable format?"); (iv) ease of electronic access (OGDE) ("Is it available online from the government in any form?"); and (v) open licensing (OGDL) ("Are the data openly licensed?"). These dimensions were derived from the relevant literature, namely Kalampokis, Tambouris and Tarabanis (2011), Kassen (2013), World Wide Web Foundation (2016). The data for this construct were retrieved from the World Wide Web Foundation (2016).

Secondly, PP refers to the interaction between citizens and governments (World Bank, 2018), and the interaction between citizens (Hays & Kogl, 2007) in which citizens take part in and contribute their opinions and knowledge for public policy decision-making, discussing and monitoring about government activities as well as the development to better the society (Mellouli, Luna-Reyes & Zhang, 2014). It is a mediating variable measured using two aspects: (i) e-participation (PPEP) measured using three dimensions, including "enabling participation by providing citizens with public information and access to information without or upon demand," "engaging citizens inn contributions to and deliberation on public policies and services," and "empowering citizens through co-design of policy options and co-production of service components and delivery modalities," which was retrieved from the (United Nations, 2016); and (ii) political participation (PPPP) measured using many items such as "percent of population that follows politics in the news media (print, TV or radio) every day," "percent of people who have taken part in or would consider attending lawful demonstrations," "voter participation/turn-out for national elections," "percent of population for being membership of political parties and political non-governmental organizations," and "percent of people who are very or somewhat interested in politics," which was retrieved from the Economist Intelligence Unit (2017).

Lastly, CC is measured based on the perception of corruption in the public sector (CPI), ranging from 0 (highly corrupt) to 100 (very clean). The data for this construct were retrieved from Transparency International (2016). Based on the research hypotheses and the formative constructed variables, the framework is presented in Figure 1.

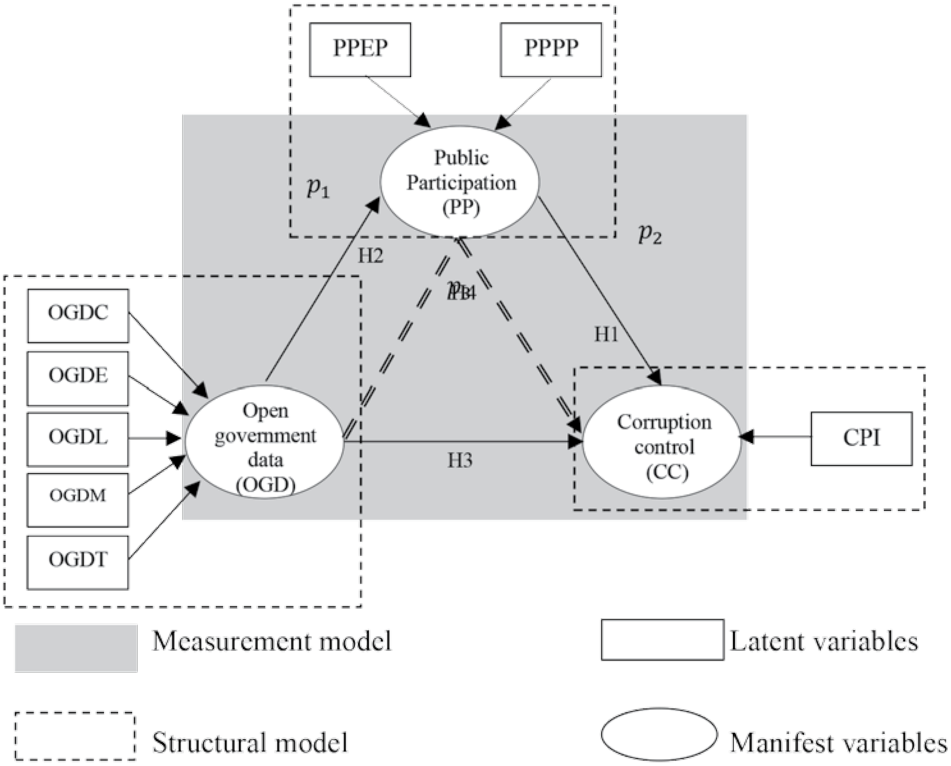

Figure 1. Research Framework and the Two Stages of PLS-SEM

Note: p_1 and p_2 are the indirect effects (path coefficients) through PP (mediator).
p_3 is the direct effect (path coefficient) between OGD and CC.

Based on the graphical form above, it can be drawn in equational form as follows:

$$PP = p_1(OGD) + e_1$$

$$CC = p_3(OGD) + p_2(PP) + e_2$$

Data Analysis

OGD are a recent and developing concept (Zuiderwijk & Janssen, 2014). Experts (Hair et al., 2019; Hair et al., 2017; Sarstedt et al., 2014) have suggested that if a study aims to conduct exploratory research or research based on a currently developed concept, the study should employ partial least squares-structural equation modelling (PLS-SEM). Additionally, PLS-SEM is suitable for formative constructs (Lee et al., 2016). It supports the formative constructs of OGD, PP, and CPI. Moreover, as this study aimed to predict the effects of OGD on PP and CC, it corresponds with the purpose of PLS-SEM, which is prediction-based. Importantly, this research used secondary data, which may have problems with normal distribution. However, PLS-SEM has no assumptions about the normality assumptions of data (data distribution) (Henseler et al., 2009; Sarstedt et al., 2014; Vinzi, Trinchera & Amato, 2010; Wong, 2013).

Thus, this study used PLS-SEM to analyze the proposed research model and used SmartPLS software, as it works well with small sample sizes (in this study, n = 113). In order to run the models, PLS algorithms and PLS bootstrapping were used with 10,000 iterations for bias correction at a 95 percent confidence level. In order to validate the hypotheses, a two-stage structural equation was applied by adopting guidelines from Sarstedt et al. (2014) and Hair et al. (2019; 2017). The measurement model as the first stage was assessed using factor analysis and construct validity was tested based on convergent validity, discriminant validity, and multicollinearity. Further, a reliability test was not required, as PLS-SEM assumes a very low correlation among the measurement items (Hair et al., 2019; Lee et al., 2016; Wong, 2013). The structural model assessment is the second stage in which the significant structural path in bootstrapping is checked to investigate the causal relationships among the independent, mediating, and dependent variables.

Table 3 lists the methods used to validate the models in this study. More importantly, after gaining the results of the analysis, the type of mediation variable was determined via the mediator analysis process in PLS-SEM developed by Lynch, Chen and Zhao (2010) and Hair et al. (2017), as presented in Figure 2.

Table 3. Model Validation

First stage: Measurement model		
Assessment	**Criteria**	**Reference**
Reliability	Not required	Hair et al. (2019), Lee et al. (2016), Wong, (2013)
Convergent validity	Significant outer weights, or values of outer loadings ≥ 0.5	Hair et al. (2019)
First stage: Measurement model		
Discriminant validity	Cross-loadings in a construct > cross-loadings with other constructs	Hair et al. (2017), Henseler, Ringle, and Sinkovics (2009)
Multicollinearity	Variance Inflation Factor (VIF) ≤ 10	Gujarati (2003), Hair et al. (2017), Henseler et al. (2009), Lee et al. (2016), Urbach and Ahlemann (2010)
Second stage: Structural model		
Nomological validity	Path coefficients between constructs are significant	Hair et al. (2019), Henseler et al. (2009)
Coefficient of determination (R^2)	$R^2 > 0.75$ (= substantial), R^2 around 0.50 (= moderate), $R^2 < 0.25$ (= weak)	Hair et al. (2019), Urbach and Ahlemann, (2010)
Blindfolding (Q^2)	Redundancy and communality > 0	Hair et al. (2019)
Second stage: Structural model		
Assessment	**Criteria**	**Reference**
Model fit	$\sqrt{\bar{R}^2\ x\ \overline{communality}}$ a value > 0.36 is high, a value between 0.25 and 36 is moderate, a value < 0.25 is low	Lee et al. (2016), Park, Lee and Chae (2017)
	Standardized root mean square residual (SRMR) for approximate model fit (estimated model); SRMR < 0.08 indicating good fit	Henseler, Hubona and Ray (2015), Henseler et al. (2014)
Checking structural model robustness in PLS-SEM	Akaike’s information criterion (AIC); smaller or more negative AIC, the better fit	Baguley (2012)
	Bayesian information criterion (BIC): smaller or more negative BIC, the better fit	Sharma et al. (2018), Sharma et al. (2019)
	Monte Carlo simulation: the higher the process capability index (Cpk), the better a model/process is.	Henseler et al. (2014)

Since a saturated model in PLS-SEM is not always straightforward, especially in a complex model, expert judgment is required to determine the saturation of a model. In order to check the robustness of the structural model in PLS-SEM, this research used AIC, BIC, and Monte Carlo simulation. AIC and BIC were developed from information theory as criteria for in-sample model selection or in-sample forecasting. Although the model performance of PLS-SEM-based fit indices can use R^2 or adjusted R^2, for example, the criteria stem from information theory balance between model fit and sophistication in order to avoid "overfitting" and to generalize the model beyond the specific sample (Myung, 2000). According to the results of Monte Carlo studied by Sharma et al. (2018; 2019), "BIC is a useful substitute for out-of-sample forecasting." Similar to the results of Monte Carlo simulated by Tolanen (2007), "BIC is most useful with small sample sizes (n < 500)." The Monte Carlo simulation shows that BIC reaches a sound trade-off between theoretical consistency (or model fit) and predictive power in the PLS model. BIC is consistent and results in a heavier penalty for model complexity than AIC (Vrieze, 2012). This research used both AIC and BIC to comprehensively consider the robust and valid results of the model. AIC and BIC can be negative values; the negative values represent a better fit of a model than positive values.

Furthermore, the Monte Carlo simulation is used to test the accuracy and robustness of PLS-SEM model performance. According to Henseler et al. (2014), this simulation helps test the model fit, which can distinguish between good-fitting models and models that do not fit so well. By taking into account the variability of the inputs, Monte Carlo simulation can be used to infer the robustness of the prediction model. In this regard, Monte Carlo simulation was performed using the mediation regression equation developed by Baron and Kenny (1986): $Y = B_0 + c' X + bM$ when B_0 is a constant value; however, because PLS is based on non-parametric or distribution-free and SmartPLS software works based on using standardized data, the constant value will thus be zero; c^' is the direct effect of OGD on CC (when considering the indirect effect through PP); X is OGD; b is the direct effect of PP on CC; and M is PP. Based on the PLS-SEM results, the equation for the Monte Carlo simulation can be stated as CC = 0.475(OGD) + 0.309(PP). The latent variable scores of OGD and PP from the PLS algorithm results were used as data inputs in Minitab workspace software. In order to define the data distribution, this research allowed the software to decide the distribution of the data; the best-fit distribution for OGD and PP was the Weibull distribution, which indicates values of OGD, including 2.033 for shape, 2.1909 for scale, and -1.9409 for threshold, as well as values of PP including 3.5086 for shape, 3.4808 for scale, and -3.1287 for threshold. The specification limits were defined based on the maximum and minimum values of CC; 2.240 (for the upper specification limit) and -1.706 (for the lower specification limit), respectively. Ten thousand iterations and 50,000 iterations as recommended by the software were performed in order to evaluate the prediction capabilities of the model. The process capability index (Cpk) was used to measure the ability of a process of a model to produce an output within the specification limits; the higher the Cpk value is, the better is the model.

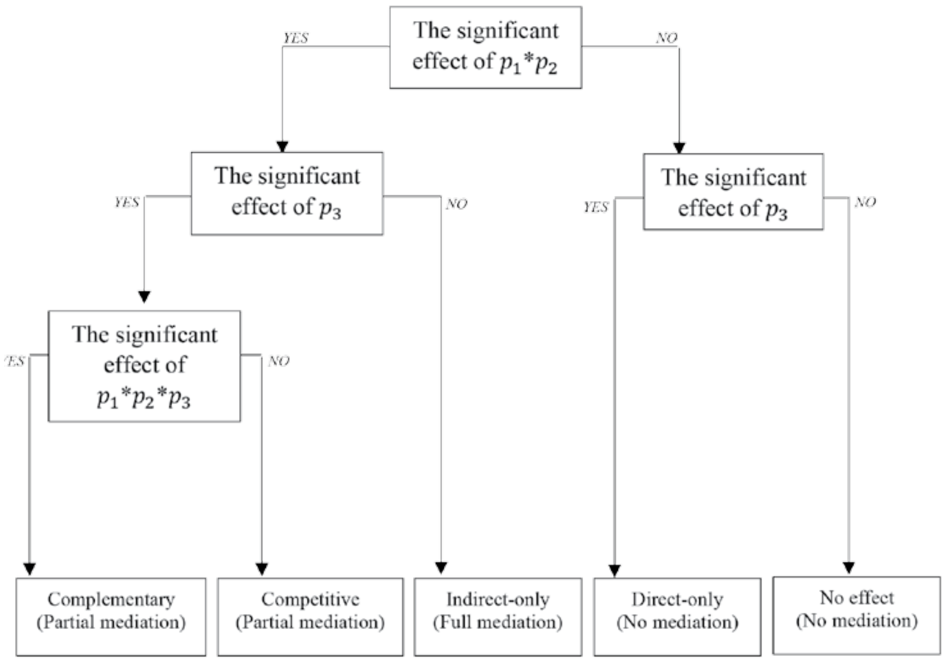

Figure 2. The Systematic Process of Mediator Analysis for PLS-SEM (Lynch, Chen & Zhao, 2010; Hair et al., 2017)

Research Results

Table 4 presents the descriptive statistics of the variables. The OGD as the independent variable included OGDC, OGDT, OGDM, OGDE, and OGDL with means of 4.87, 2.80, 6.21, 7.14, and 3.22, respectively. PP as the mediator variable obtained a mean of 0.60 for the PPEP items and a mean of 5.54 for the PPPP items. The other mediator variable, CPI, had a mean of 46.87.

Table 4. The Nature of the Dataset

Latent variable	Item	Mean	Std. Deviation	Kurtosis	Skew-ness	Source
Open Government Data (OGD)	OGDC	4.87	0.27	5.83	-2.43	World Wide Web Foundation (2016)
	OGDT	2.80	2.17	-0.21	0.57	
	OGDM	6.21	3.49	-0.74	0.24	
	OGDE	7.14	1.69	-0.14	-0.56	
	OGDL	3.22	2.61	3.28	1.95	
Public Participation (PP)	PPEP	0.60	0.23	-0.56	-0.34	United Nations (2016)
	PPPP	5.54	1.79	-0.52	0.00	Economist Intelligence Unit (2017)
Corruption (CC)	CPI	46.87	19.26	-0.51	0.73	Transparency International (2016)

The Result of the Measurement Model

The measurement model was validated via factor analysis based on three assessments. First, convergent validity was tested by considering significant outer weights. If the outer weights were not significant, the value of the outer loadings should be greater than 0.5 (Hair et al., 2019). According to Table 5, the outer weights of OGDC, OGDE, OGDL, OGDT, PPEP, and PPPP were significant. However, although the outer weight of OGDM was not significant, its outer loading was greater than 0.5 (outer loading of OGDM = 0.891). In this case, the model reached convergent validity. Secondly, discriminant validity was examined by comparing the cross-loadings within a construct, which should be greater than the cross-loading with other constructs (Hair et al., 2017; Henseler et al., 2009). Table 5 shows that all of the cross-loadings within constructs were greater than those of the other constructs. Discriminant validity was therefore achieved. Lastly, before conducting the path analysis, the correlation among items must be verified. Multicollinearity was examined using VIFs, which should be less than 10, to indicate no multicollinearity. The results showed that all of the VIF values were less than 10, indicating no multicollinearity among the items.

Table 5. Factor Analysis of the Measurement Model

Formative construct	Item	Outer weight (Outer loading)	t-value	VIF	Cross-loading		
					1	2	3
Open Government Data (OGD)	OGDC	0.175** (0.535)	1.923	1.481	0.535	0.455	0.325
	OGDE	0.207** (0.778)	2.308	2.224	0.778	0.627	0.509
	OGDL	0.315** (0.796)	2.840	1.902	0.796	0.599	0.566
	OGDM	0.079 (0.891)	0.424	4.573	0.891	0.678	0.625
	OGDT	0.470*** (0.904)	3.213	3.483	0.904	0.647	0.677
Public Participation (PP)	PPEP	0.609*** (0.884)	6.976	1.344	0.665	0.884	0.592
	PPPP	0.543*** (0.851)	6.033	1.344	0.645	0.851	0.565
Corruption (CC)	CC	1.000	-	1.000	0.708	0.667	1.000

Note *** = *p*-value < 0.001, ** = *p*-value < 0.05

The Result of the Structural Model

The structural modal was examined using four assessments. First, the significance of the path coefficients among constructs was determined, so-called nomological validity (Hair et al., 2019; Henseler et al., 2009). Table 6 shows a 0.309 path coefficient of PP on CC, which was statistically significant, as the p-value was 0.001 (p-value < .05); therefore, H1 was supported. The path coefficient of OGD on PP was 0.755, which was significant, as the p-value was 0.001 (p-value < 0.05); H2 was therefore supported. The path coefficient of OGD on CC was 0.475, which was significant, as the p-value was 0.001 (*p*-value < 0.05); H3 was therefore supported. In sum, the path coefficients of H1, H2, and H3 were significant at a p-value < 0.05.

Table 6. Significance of the Structural Model

Hypothesis	Structural relationship	Path coefficient	T statistics	P-value	Result
H1	PP → CC	0.309	3.982	0.001	Supported
H2	OGD → PP	0.755	18.337	0.001	Supported
H3	OGD → CC	0.475	5.405	0.001	Supported

In addition to the above, the mediating analysis results are presented in Table 7. The indirect effect of OGD and PP on CC was 0.233, which was significant, as the p-value was 0.001 (p-value < 0.05); hence, H4 was supported. The results reported a significant direct effect for (OGD → CC = 0.475***), a significant indirect effect for (OGD → PP → CC = 0.233***), and a positive significance of the total effect (the direct effect + the indirect effect = 0.708***). Therefore, the type of mediation was partial or complementary mediation.

Table 7. The Mediating Effect

Hypothesis	Structural relationship	Direct effect	Indirect effect	Total effect	Result
H4	OGD → PP → CC	0.475***	0.233***	0.708***	Supported
			(T-statistics = 3.780)	(T-statistics = 13.212)	

Second, the coefficient of determination (R^2) was examined in order to assess predictive fit (Hair et al., 2019; Urbach & Ahlemann, 2010). As shown in Figure 3, OGD was able to explain 57 percent of the variance in PP, indicating a moderate level of explanatory power. Together, OGD and PP explained 54.2 percent of the variance in CC, representing moderate explanatory power as well.

Third, blindfolding (Q^2) was done to assess the predictive power of the model (Hair et al., 2019), where the larger is Q^2, the higher is the predictive accuracy. As presented in Table 8, the redundancy values of PP (0.397) and CC (0.500) were larger than zero, indicating an acceptable predictive accuracy of the model path. The average redundancy of the model was 0.4485, indicating high predictive power.

Fourth, the first model fit method was examined with the steps proposed by Lee et al. (2016) and Park et al. (2017). In the first step, the average R^2-value was obtained (0.556) and the average communality was calculated (0.56433). In the following step, the average R^2-value was multiplied by the average communality to produce 0.313769333. In the final step, the square root of the result of the second step was obtained (0.560) to give the overall model fit. This model fit was considered high, as presented in Table 8. Furthermore, the second measure of model fit considered the value of SRMR. Henseler et al. (2014) introduced SRMR as a goodness of fit measure for PLS-SEM, which can be used to avoid model misspecification. The results showed that the SRMR value of the research model was 0.021 in the estimated model, which was less than 0.08. This is considered a good fit.

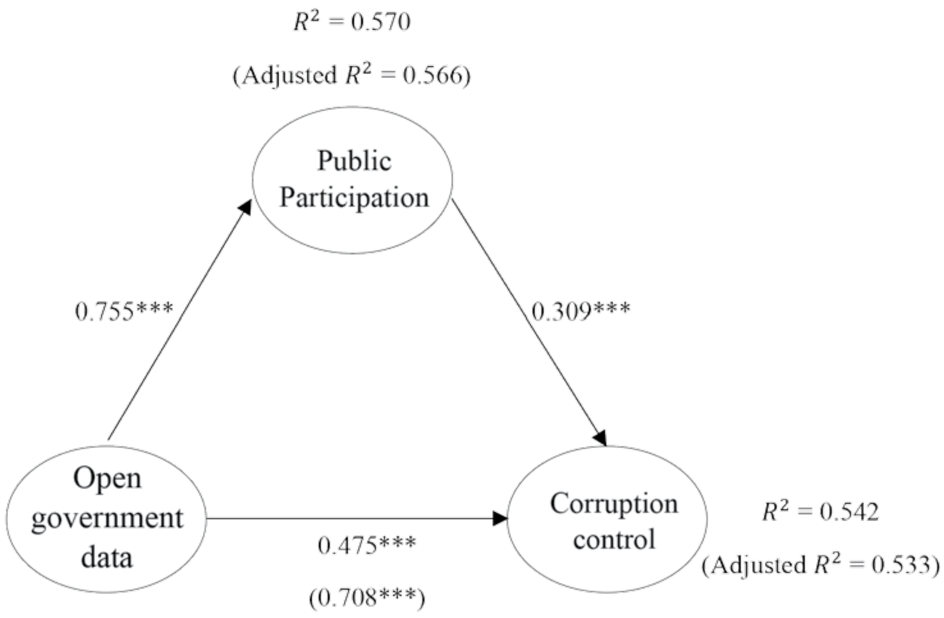

Figure 3. R^2 (Adjusted R^2) and Model Significant Path Coefficients

Note *** = p-value < 0.001, ** = p-value < 0.05

Table 8. Goodness-of-Fit

Latent variable	R^2-value	Redundancy	Communality
Open government data (OGD)			0.434
Public participation (PP)	0.570	0.397	0.259
Corruption control (CC)	0.542	0.500	1.000
Average	0.556	0.4485	0.56433
Overall model fit	0.560		

Last, the research model produced AIC and BIC in negative values for both PP and CC; the AIC value for PP was -92.416 and that for CC was -83.189; the BIC value for PP was -86.961 and that of CC was -75.007. The negative values of AIC and BIC show a good fit since the negative values indicate less data loss; minimal data loss is relative to the true model. Regarding the Monte Carlo simulation results, the Cpk value from the 10,000 iteration (n=10,000) simulation was equal to 1.17. The simulation indicates that 0 percent of the CC values fell outside the specification limits and produced the value of mean equal to 0.009 with a standard deviation equal to 0.567 (min = -1.634, median = -0.024, max = 2.080). Similar to the simulation with 50,000 iterations (n=50,000), the value of Cpk was equal to 1.17 with 0.03 percent of the CC values falling outside the specification limits. This latter simulation produced a mean value equal to 0.006, with standard deviation equal to 0.568 (min = -1.608, median = -0.029, max = 2.788). According to the values of Cpk from the 10,000 and 50, 000 iterations, a statistical explanation of the model when the curve stretches from +2.24 to -1.706 produces 99.95 percent good parts. At least 99.95 percent of the output (CC) from the model (OGD and PP) was good.

Discussion

The results can be discussed according to four perspectives as follows. First, the ISS model was quite useful in providing a comprehensive understanding of the OGD, PP, and CC systems since they can be perceived as input, process, and output, respectively.

Second, the positive effect of PP on CC was statistically significant. This result indicated that public participation can enhance CC. This result was compatible with previous claims and studies, such as those of Hays and Kogl (2007), Mellouli, Luna-Reyes and Zhang (2014), and World Bank (2018) who proposed PP as a solution to corruption.

Third, OGD significantly and positively affected PP and CC since H2 and H3 were supported. OGD positively affected the level of PP. Previous scholars (Bartenberger & Grubmüller, 2014; Daniel & Habsari, 2019; Evans & Campos, 2013) mentioned that information asymmetry or a lack of data was the essential reason why people neglect to participate with others or with the government. The results of this study proved that OGD are a fundamental resource for people to participate in government processes. The results are consistent with those of previous studies, such as those of Davies (2012), Kassen (2013), Lee et al. (2016), and Odongo and Rono (2016). Moreover, OGD were shown to positively and significantly affect CC, a result that was consistent with that of previous works such as those of Darusalam et al. (2019), Florez and Tonn (2019), Hulstijn et al. (2018), and Janssen and Zuiderwijk (2012), all of whom proposed that OGD's accessibility and availability would help increase transparency, increase accountability, reduce the likelihood of corruption behavior, and strengthen the efforts to increase CC. In short, the higher is the availability of OGD, the lower is the damage level.

Last, the mediating results supported H4, indicating that the relationship between OGD and CC was mediated by PP. This result affirms the main argument put by this study. Since the direct, indirect, and total effects were positively significant, PP had a complementary partial mediation effect on the relationship between OGD and CC. It can be interpreted that the effect of OGD on CC would be greater if PP were added to the model. In other words, PP is an important process by which OGD are used to tackle CC. This result filled the gap in the understanding of this subject in previous works, such as those of Hulstijn et al. (2018), Iglesias (2017), Park and Kim (2019), and Florez and Tonn (2019). OGD had been challenged in measuring its impact on CC. It did not have an immediate effect on CC. OGD may have a weak impact on CC since people were rarely able to access to the data or rarely used the data on that purpose. In turn, this research found that OGD still positively, significantly, and directly affected CC. Moreover, the study found a positive significant effect of OGD and PP on CC. In sum, this research indicated that the more people can access and use OGD, the more they will participate in CC, consistent with previous work, namely that of Darusalam et al. (2019), Davies and Perini (2016), Hulstijn et al. (2018), Iglesias (2017), and Vrushi and Hodess (2017). Additionally, PP was posited to link OGD policy to anti-corruption policy. This study confirmed this linkage where OGD and PP together explained 54.2 percent of the variance in CC. Providing and releasing OGD as a public service will facilitate and empower the public, i.e. citizens, non-government organizations, the media, journalists, watchdog groups, regulators, auditors, experts, and so on, to participate in monitoring, discussing, exposing, and controlling government activities and performance actively and meaningfully.

Conclusion

Corruption is a complex and chronic problem. PP is expected to support CC; however, government-initiated PP is often unsuccessful since citizens that lack data or information will not participate with others or with their government. Currently, OGD's growing popularity now provides opportunities for citizens to expand and accumulate knowledge and information about the government that, in turn, will support PP and CC. This study found that OGD positively and significantly affected PP. That is, the more freely available and accessible the government data are, the more likely the people are to be informed and to be able to use the data to participate in government activities. Indeed, OGD encourage citizens to participate in CC by empowering PP. Given the importance of OGD as discussed in this study, countries should give special consideration to the benefits and values of OGD in order to empower PP and to promote CC.

This research recommends that the policies of OGD, PP, and CC be formulated and implemented to support each other. A government should set OGD's goals to support PP and CC because anti-corruption can be driven by the OGD process when the public can access the data and leverage it to monitor and control their government. OGD should be identified in PP policy, as it is a resource of communication and participation between a government and citizens. OGD will enhance the ability of citizens to meaningfully and vigorously communicate, monitor, check, and participate in controlling corruption and also to improve the decision-making processes in government activities. Likewise, a government should create cooperation with the public as they are the data users that design the platform of OGD for CC purposes. The public will be aware of what government data should be disclosed, and how essential the data are in monitoring and checking their government performance. For example, releasing government spending data will empower the public to prevent the wasted spending of their government. Having the right and valuable data can help citizens effectively participate in anti-corruption. This complementary linkage of policies will enhance opportunities to effectively fight against corruption according to the intended results.

However, since the notion of OGD is currently being developed and this research only analyzed secondary data, future research should be conducted with primary data in order to confirm the findings of this study and to explain the subject in greater detail. Researching a local area or narrow area may produce concrete results. Further, since this research focused on the relationships among the three variables without taking into account other context variables, future research may explore OGD, PP, or CC factors that may impact the relationship among the three variables, such as laws on the protection of citizen's rights in accessing government data, national culture, socioeconomic and political factors, and so on. A study of such factors would help policymakers develop a more suitable environment for the public to use OGD for anti-corruption. Importantly, as this research did not indicate what government data should be made available, future research may employ in-depth study on the types of government data that should be disclosed, especially data that are relevant and useful for controlling corruption, for example, government spending, government procurement data, political parties’ and politicians’ financing, and so on. Finally, future research could explore how citizens or anti-corruption institutions use OGD to fight corruption.

References

Ansell, C., & Gash, A. (2007). Collaborative governance in theory and practice. *Journal of Public Administration Research and Theory, 18*(4), 543-571.

Awan, R. U., Akhtar, T., Rahim, S., Sher, F., & Cheema, A. R. (2018). Governance, corruption and economic growth: A panel data analysis of selected SAARC countries. *Pakistan Economic and Social Review, 56*(1), 1-20.

Baguley, T. (2012). *Serious Stats: A Guide to Advanced Statistics for the Behavioral Sciences.* Basingstoke: Palgrave Macmillan.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Bartenberger, M., & Grubmüller, V. (2014). The enabling effects of open government data on collaborative governance in smart city contexts. *JeDEM - EJournal of EDemocracy and Open Government, 6*(1), 36-48.

Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly, 27*(3), 264-271.

Bhayani, M. (2013). Challenges of corruption and good governance: A human rights perspective. *International Proceedings of Economics Development and Research, 71*(6), 21-27.

Bishop, P., & Davis, G. (2002). Mapping public participation in policy choices. *Australian Journal of Public Administration, 61*(1), 14-29.

Daniel, D., & Habsari, P. (2019). Informational asymmetry in public participation on environmental monitoring in Indonesian regulation: A preliminary discourse. *Journal of Infrastructure Development, 2*(1), 4-19.

Darusalam, D., Said, J., Omar, N., Janssen, M., & Sohag, K. (2019). The diffusion of ICT for corruption detection in open government data. *Knowledge Engineering and Data Science (KEDS), 2*(1), 10-18.

Davies, T. (2012). *Ten Building Blocks of an Open Data Initiative*. Retrieved March 1, 2021 from http://www.opendataimpacts.net/2012/08/ten-building-blocks-of-an-open-data-initiative/.

Davies, T., & Perini, F. (2016). Researching the emerging impacts of open data: Revisiting the ODDC conceptual framework. *Journal of Community Informatics, 12*(2), 148-178.

DeLone, W., & McLean, E. (2003). The DeLone and McLean model of information systems success: A ten-year update. *JMIS, 19*(4), 9-31.

DeLone, W. H., & Mclean, E. R. (1992). Information system success: The quest for the dependent variable. *Information Systems Research, 3*(1), 60-95.

Desta, Y. (2006). Designing anti-corruption strategies for developing countries: A country study of Eritrea. *Journal of Developing Societies, 22*(4), 421-449.

Economist Intelligence Unit. (2017). *Democracy Index 2016: Revenge of the Deplorables.* Retrieved March 1, 2021 from https://eiuperspectives.economist.com/sites/default%0A/files/TheEIU%27s2016DemocracyIndex_0.pdf%0A.

Evans, A. M., & Campos, A. (2013). Open government initiatives: Challenges of citizen participation. *Journal of Policy Analysis and Management, 32*(1), 172-185. Faisal,

F., & Jafri, A. R. (2017). Corruption as a source of failure of good governance and management in Pakistan: Proposed remedial measures. *Jpuhs, 30*(1), 57-75.

Florez, J., & Tonn, J. (2019). Accountability and anti-corruption. In Davies, T., Walker, S., Rubnistein, M., & Perini, F. (Eds.). T*he State of Open Data: Histories and Horizons.* Cape Town and Ottawa: African Minds and International Development Research Center.

Gigler, B., Custer, S., & Rahemtulla, H. (2011). *Realizing the Vision of Open Government Data: Opportunities, Challenges, and Pitfalls.* Open Development Technology Alliance and World Bank, (December), 13-35. Retrieved March 1, 2021 from https://www.academia.edu/35025372/Realizing_the_Vision_of_Open_Government_Data_Opportunities_Challenges_and_Pitfalls.

Goodrich, S. (2015). *How Open Data can Help Tackle Corruption: Policy Paper.* Retrieved March 1, 2021 from https://www.transparency.org.uk/sites/default/files/pdf/publications/Policy_Paper_-_How_open_data_can_help_tackle_corruption.pdf.

Gujarati, D. N. (2003). *Basic Econometrics.* New York: McGraw-Hill.

Guriev, S. (2004). Red tape and corruption. *Journal of Development Economics, 73*(2), 489-504.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis.* United States: CENGAGE.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM).* Thousand Oaks: SAGE Publications Ltd.

Halonen, A. (2012). *Being Open about Data: Analysis of the UK Open Data Policies and Applicability of Open Data.* Retrieved March 1, 2021 from http://http://fininst.uk/wp-content/uploads/2017/09/being-open-about-data.pdf.

Hartog, M., Mulder, B., Spée, B., Visser, E., & Gribnau, A. (2014). Open data within governmental organizations: Effects, benefits, and challenges of the implementation process. *JeDEM - EJournal of EDemocracy and Open Government, 6(*1), 49-61.

Hays, R. A., & Kogl, A. M. (2007). Neighborhood attachment, social capital building, and political participation: A case study of low-and moderate-income residents of Waterloo, Iowa. *Journal of Urban Affairs, 29*(2), 181-205.

Henseler, J. et al. (2014). Common beliefs and reality about PLS: Comments on Rönkkö & Evermann (2013). *Organizational Research Methods, 17(*2), 182-209.

Henseler, J. Hubona, G., & Ray, P. A. (2015). Using PLS path modeling in new technology research: Updated guidelines. Industrial *Management & Data Systems, 116*(1), 2-20.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing, 20*(January), 277-319.

Hulstijn, J., Darusalam, D., & Janssen, M. (2018). Open data for accountability in the fight against corruption. In Baldoni, M., Baroglio, C., & Micalizio, R. (Eds.). *Computational Accountability and Responsibility in Multiagent Systems (CARe-MAS 2017).*

Iglesias, D. (2017). *Open Data and the Fight Against Corruption in Brazil (Vol. 1).* Berlin: Transparency International, International Secretariat.

Ionescu, L. (2013). Perceptions of corruption in emerging economies. *Economics, Management and Financial Markets, 8*(1), 365-395.

Izdebski, K. (2015). Transparency and open data principles: Why they are important and how they increase public participation and tackle corruption. In *U. S. Department of State.* Retrieved March 1, 2021 from https://transparencee.org/wp-content/uploads/2015/12/open-data-principles-by-krzysztof-izdebski.pdf.

Janssen, K. (2012). Open government data and the right to information: Opportunities and obstacles. *The Journal of Community Informatics, 8*(2).

Janssen, M., & Zuiderwijk, A. (2012). *Open Data and Transformational Government.* Conference Transforming Government Workshop. United Kingdom: Brunel University.

Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2013). Generating value from open government data. *International Conference on Information Systems (ICIS 2013): Reshaping Society Through Information Systems Design, 2*, 1737-1756.

Kalampokis, E., Tambouris, E., & Tarabanis, K. (2011). A classification scheme for open government data: Towards linking decentralized data. *International Journal of Web Engineering and Technology, 6*(3), 266-285.

Kassen, M. (2013). A promising phenomenon of open data: A case study of the Chicago open data project. *Government Information Quarterly, 30*(4), 508-513.

Kim, S., Kim, H. J., & Lee, H. (2009). An institutional analysis of an e-government system for anti-corruption: The case of OPEN. *Government Information Quarterly, 26*(1), 42-50.

Kumar, R. (2004). Corruption and human rights: Promoting transparency in governance and the fundamental right to corruption-free service in India. *Columbia Journal of Asian Law, 17*(1), 31-72.

Lee, J. N., Ham, J., & Choi, B. (2016). Effect of government data openness on a knowledge-based economy. *Procedia Computer Science, 91*(Itqm), 158-167.

Lynch, J. G., Jr., Chen, Q., & Zhao, X. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*(2), 197-206.

Mellouli, S., Luna-Reyes, L. F., & Zhang, J. (2014). Smart government, citizen participation and open data. *Information Polity, 19*(1–2), 1-4.

Myung, I. J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology, 44*(1), 190-204.

Odongo, A. O., & Rono, G. C. (2016). *Open Government Data as a Right for Effective Citizen Participation.* ACM International Conference Proceeding Series, 01-03-Marc, 365-366.

OECD. (2018). *Open Government Data Report: Enhancing Policy Maturity for Sustainable Impact.* Retrieved March 1, 2021 from https://read.oecd-ilibrary.org/governance/open-government-data-report_9789264305847-en#page1.

Park, C. H., & Kim, K. (2019). E-government as an anti-corruption tool: Panel data analysis across countries. *International Review of Administrative Sciences,* 1-36.

Park, S., Lee, H., & Chae, S. W. (2017). Rethinking balanced scorecard (BSC) measures: Formative versus reflective measurement models. I*nternational Journal of Productivity and Performance Management, 66*(1), 92-110.

Peled, A. (2013). Re-designing open data 2.0. *JeDEM - EJournal of EDemocracy and Open Government, 5*(2), 187-199.

Rajshree, N., & Srivastava, B. (2012). *Open Government Data for Tackling Corruption – A Perspective Case Study: NREGA in India.* 2012 AAAI Workshop - Semantic Cities, 21-24.

Renata, M. (2017). *Measuring the Effects of Open Data on the Level of Corruption*. In Proceedings of the 21st International Conference Current Trends in Public Sector Research.

Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity, 22*(1), 1-24.

Sarstedt, M., Ringle, C. M., Smith, D., Reams, R., & Hair, J. F. (2014). Partial least squares structural equation modeling (PLS-SEM): A useful tool for family business researchers. *Journal of Family Business Strategy, 5*(1), 105-115.

Sharma, P. N., Sarstedt, M., Shmueli, G., Kim, K. H, & Thiele, K. O. (2019). PLS-based model selection: The role of alternative explanations in information systems research. *Journal of the Association for Information Systems, 20*(4), 346-397.

Sharma, P. N., Shmueli, G., Sarstedt, M., Danks, N., & Ray, S. (2018). *Prediction-Oriented Model Selection in Partial Least Squares Path Modeling,* Decision Sciences.

Steele, J. L., Murnane, R. J., & Willett, J. B. (2010). Low-performing schools attract and keep academically talented teachers? Evidence. *Journal of Policy Analysis and Management, 29*(3), 451-478.

Tolanen, A. (2007). *Latent Growth Mixture Modeling: A Simulation Study*. Department of Mathematics and Statistics, University of Jyväskylä.

Transparency International. (2016). *Corruption Perception Index 2016*. Retrieved March 1, 2021 from https://www.transparency.org/news/feature/corruption_perceptions_index_2016.

Tshepo, M. (2015). *The Nexus between Corruption and Governance since the Dispensation of Democracy in South Africa: A Time Series Analysis.* Retrieved from March 1, 2021 http://2015. essa.org.za/fullpaper/essa_2854.pdf.

United Nations. (2016). *E-government Survey 2016*. Retrieved March 1, 2021 from http://workspace.unpan.org/sites/Internet/Documents/UNPAN96407.pdf.

United Nations Development Programme. (1997). *Corruption and Good Governance.* New York: United Nations.

Urbach, N., & Ahlemann, F. (2010). Structural equation modeling in information systems research using partial least squares. *Journal of Information Technology Theory and Application, 11*(2), 5-40.

Vinzi, V. E., Trinchera, L., & Amato, S. (2010). PLS path modeling: From foundations to recent developments and open issues for model assessment and improvement. In *Handbook of Partial Least Squares.* Springer.

Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods, 17*(2), 228-243.

Vrushi, J., & Hodess, R. (2017). *Connecting the Dots: Building the Case for Open Data to Fight Corruption.*

Wong, K. K. K. (2013). Partial least squares structural equation modeling (PLS-SEM) techniques using SmartPLS. *Marketing Bulletin, 24*(1), 1-32.

World Bank. (2018). *Engaging Citizens for Better Development Results: An Independent Evaluation.* Retrieved March 1, 2021 from https://ieg.worldbankgroup.org/sites/default/files/Data/Evaluation/files/Engaging_Citizens_for_Better_Development_Results_FullReport.pdf.

World Wide Web Foundation. (2016). *Open Data Barometer Global Report.* Retrieved from http://www.opendatabarometer.org.

Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly, 31*(1), 17-29.

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012). Socio-technical impediments of open data. *Electronic Journal of E-Government, 10*(2), 156-172.