

On Robust Predictive Analytics for Advertisement Management Framework in Local Governments

Mokhtar Abdullah^{a*}, Tun Mohd Izlizam Bahardin^b, Zuraini Alias^c, ^aTime Brain Sdn Bhd, Petaling Jaya, Selangor, ^{b,c}Universiti Selangor, Selangor Business School, Petaling Jaya, Selangor, Email: ^{a*}profmokh@gmail.com

This paper presents a study on advertisement management in local governments, which are entrusted with establishment of advertisement standards and regulations that would ensure industry's compliance with the advertisement regulations. However, the lack of clear and detailed standards by the local governments has resulted in illegal and unlawful advertisement activities. Bahardin et al. (2019) developed a predictive model and performed a predictive analytics using Partial Least Squares (PLS) method for the relationship between advertisement standardization and industry's compliance with advertisement regulations using a data set from Petaling Jaya City Council, a local government in Malaysia. In this paper, a robust version of PLS, i.e., (PLSc), that is resilient to undesired effects of outliers was proposed. From the empirical analysis on the predictive model, it was shown that the robust method was able to reduce the influence of the outliers and produced more reliable results compared to its traditional counterpart, PLS.

Key words: *Advertisement Management, Regulatory Compliance, Outliers, Predictive Analytics, Robust Partial Least Squares.*

Background

Local government is an organization that deals with issues affecting people living in a specific area. It reflects the local microscopic interests contributing to a wider understanding of the well-being and happiness of its residents. As part of its third government status, it is assigned to the local government to draft legislation dealing with the different principles and practices of advertisement management in the law enforcement framework. In Malaysia, the MBPJ is a local government in Selangor that vigorously pursues the system for advertisement standardization for the local advertising industry.

Advertisement standardization agenda is fundamental for a better understanding of the whole advertisement management system. For example, the local governments have to oversee and monitor all sorts of business advertisement licensing issues related to advertisement media through cooperation with other regulatory government bodies. At present, the federal government is responsible for regulating all types of advertisement media in the advertising industry. In this regard, the government must understand the advertisement framework of advertising industry within its power and jurisdiction.

Advertisement Standardisation Management in Local Government

Local governments are bound to enforce the regulations within the advertisement framework in order to improve the advertisement management practice. This is done by providing the guiding principles for the local authorities and other regulatory government bodies involved in implementing the new advertisement regulations for advertising industry.

However, the local governments do not have official or common standards of advertisement management practices that need to be complied by advertising industry via different advertisement media. This is because the existing advertisement management does not match with all advertisement media requirements that include specification, definitions, and classification of advertisement media. In addition, the local governments have to deal with unlawful advertisement media by industry players without a proper rule and guideline. The consumers are also facing a problem with misleading information provided by advertisers in various platforms. In relation to this, the consumers feel uncertain about authenticity of products or services offered by these advertisers. On the part of the clients of the local authority, they are facing difficulty in getting the right information and guideline on how to get approval for advertisement purposes. This might be due to lack of communication in disseminating the right information between the local government staff and the clients. Inefficient and incapability of technology channel and infrastructure that provides relevant information for the clients is another issue faced by most local governments. If the standardization of advertisement management and lack of compliance by advertising industry are not addressed effectively by the local authorities and advertisement media industry, respectively, there will be undesired consequences incurred on the local government.

Lack of uniformity in the standardization of advertisement management practices across the local governments is a major issue as far as revenue generation is concerned. The different approaches to implementation of advertising standards among the local governments may influence diversity in advertisement practices by advertising industry. As a result, the opportunities to generate substantial income through revenue collection for the local governments will be affected. This situation may cause limited financial resources for the local governments.

The continuing unlawful practices and failure to meet the advertising regulations by the advertising industry will lead to more uncontrolled and widespread malpractices which would affect consumers' welfare. This undesired situation may also contribute to the loss of substantial revenue collection for the local government. In addition, operation costs involving the deployment of enforcement personnel on unlawful advertising practices will be increased. Furthermore, unnecessary budget allocation will have to be made available for the cleaning process of advertisement materials.

Therefore, radical changes are needed in the advertisement management practices and advertisement regulations by the local government. For the local governments the advertisement management practices and specific policies have to take into consideration the environmental factors before any standardization of advertisement can be pursued. For instance, a particular type of advertisement standard that is to be established in certain areas needs to consider the geographical and physical nature of the areas.

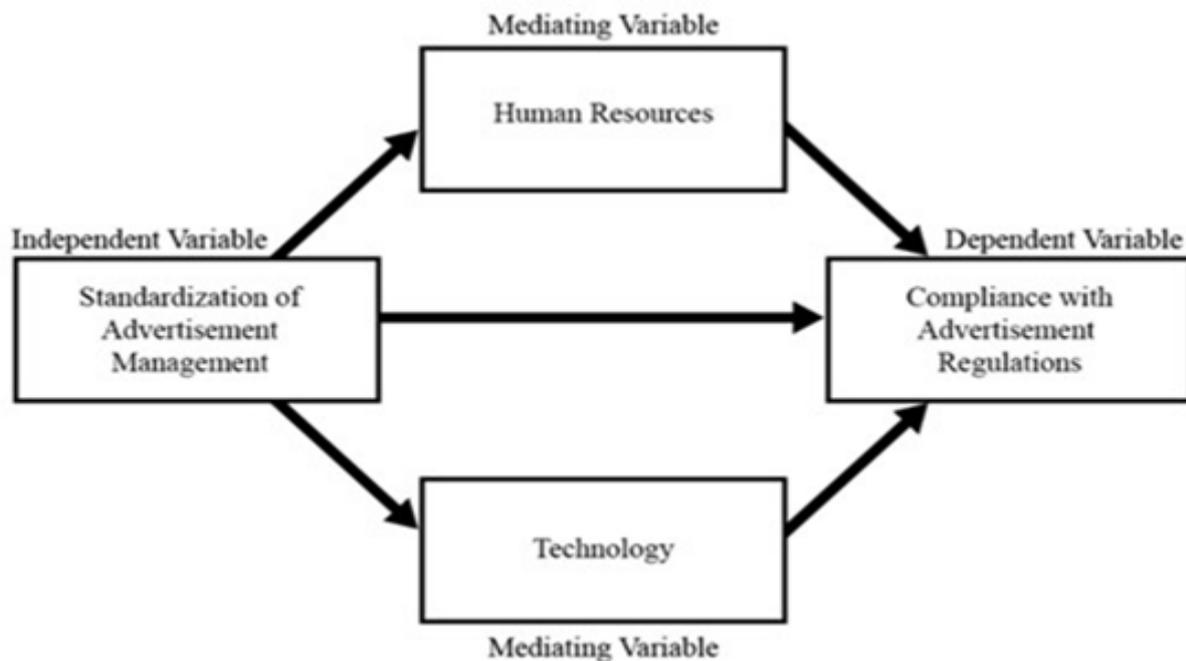
In order to establish a successful and sustainable ecosystem of the advertisement management issue, the authority such as the local government to understand the requirements for regulation. Regulation can be defined as a rule or order issued by an executive authority or regulatory agency of a government and having the force of law. Rules made by a government or other authority in order to control the way something is done or the way people behave (Michael et al., 2015). Regulations also can be defined as guidance or references to comply for correct or right action (Münstermann & Weitzel, 2008).

A Study from Beimborn et al. (2009) found that standardization have successfully improved the efficiency, quality and control of their advisory process as well as compliance rate. This means that by having standardization, the goal can be achieved and marked by the high compliance rate. Moreover, findings by Spence et al. (2015) proved that regulation compliance shows a positive relationship with standardization.

The Framework for Advertisement Standardisation Management

Having the advertisement standards is not sufficient without measuring the advertising industry's compliance level. Bahardin et al. (2019) developed a conceptual framework that displays the relationship between advertisement standardization and industry's compliance with the standards developed by the local government. Two mediating factors, namely, the human resources and technology were included in the framework. Figure 1 illustrates the inter-relationships between the standardization of advertisement management as the independent variable, the compliance with advertisement regulations as the dependent variable, and human resources and technology as two mediating variables in the framework.

Figure 1. The conceptual framework (Bahardin et al., 2019, 2020)



The key question in this framework is ‘Do enhanced levels of standardization of advertisement result in enhanced compliance level of industry players with the advertisement regulations? In addition to this, the next questions are ‘ Do human resources and technology help to strengthen the relationship between the standardization of advertisement management and the industry compliance with advertisement regulations?’

To this end, Bahardin et al. (2019) carried out an empirical study to verify the statistical significance of each of the relationships in the framework. The Structural Equation Modeling using Partial Least Squares (PLS-SEM) (Wold, 1975) was adopted to test the statistical significances of the relationships in the model based on a survey sample of 240 respondents consisting of MBPJ employees.

Predictive Analytics via Consistent PLS-SEM (PLSc)

Predictive analytics for measuring and tracking the performance of local government is at a crossroads. The emergence of big data and advances in technological and analytical tools provided opportunities for measuring and tracking a wider range of performance indicators on a timelier basis. The idea that local governments should focus more on their performance is not new. What is different today is the ability to do so regularly — and successfully — through an exponential increase in the ability to analyze massive historical data sets.

In the social and marketing research, structural equation modeling (SEM) has become a popular tool. This is because of its ability to provide a flexible framework to quantify and test interrelationships between observed and latent variables (Kaplan 2009). CSA (Jöreskog 1973) and Partial Least Squares (PLS) path modeling (Wold 1975) are two theoretically distinct solutions to SEMs (Fornell and Bookstein, 1982; Reinartz et al., 2009).

Traditional PLS (Wold, 1975) measures specific factors for observed weighted composites. The PLS constructive scores are ultimately contaminated by errors in measurement, since these composites represent reflective constructs. The relationship between any two constructs is mitigated by errors in measurement, which results in biased and inconsistent structural relations (e.g., Bollen, 1989; Cassel et al., 2000). It would be worth correcting attenuation of measurement errors, since a structural equation model usually contains one or more common factors. Dijkstra and Henseler (2015b) proposed “consistent PLS” (PLSc) as an algorithm intended to produce consistent and asymptotically normal estimates of path loadings and of correlations among latent variables for reflectively-modeled constructs. The rationale behind the development of PLSc is that the traditional PLS underestimates the R-squared value of endogenous latent variables (Dijkstra, 2010). The PLSc is also designed for fully connected common factor models in which all constructs are reflectively measured.

The Effect of Outliers on Predictive Analytics

Outliers are data points that are substantially different from most of the data (Grubbs 1969). Outliers may occur entirely unsystematically, and therefore no structure is being followed. Outliers, e.g. from a population other than the other observations, may also occur systemically. In certain cases the outliers may generate high value if they are found, value in terms of cost savings, improved efficiency, compute time savings, fraud reduction and failure prevention. Two methods are widely used to deal with outliers in empirical research. The first involves robust estimators which are not or are just distorted to a smaller degree by outliers. Additionally, before the final calculations, the outliers would be identified and removed manually. This latter approach is usually considered as the inferior one. The removal of outliers often entails the possible loss of valuable information (Gideon and Hollister, 1987), and reduced statistical power for small sets of data. When data set get bigger, there will be more likely that the number of outliers increases, and as the variety of features increases, the more difficult it becomes to detect and handle the outliers. A big data refers to a large, dynamic collections of data. These vast and complex data appear problematic for handling outliers through traditional methods. However, with a proper and efficient tool, big data may provide a considerable opportunity to uncover new insights and hidden relationships. In many instances, using a visualization tool to get your eyeballs on a whole bunch of data is the best way to identify hard-to-see outliers. Box and whisker plots, available from most statistical software such as SPSS, can help identify the presence of outliers.

In this paper a robust approach was proposed to deal with outliers in a structural equation model that involves causal relationships (such as the relationships depicted in Figure 1) and collection of big data through manifest variables in the model.

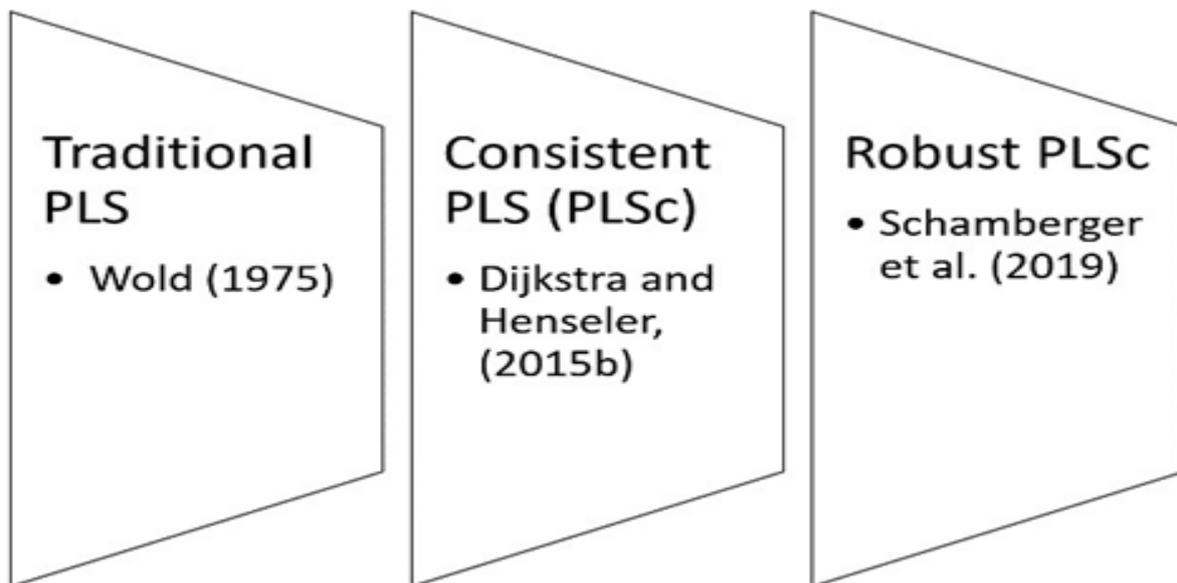
Robust Predictive Analytics Using PLSC

Outliers can have bad effects on the validity of results of empirical analyses. Many empirical studies have shown that outliers can easily affect estimators if no proper treatment is given to reduce their influence. Even detecting and eliminating outliers manually, if practiced at all, often create unnecessary problems such as loss of information and reduced statistical power (Schamberger et al., 2019). Using robust methods that are able to downweight the influence of outliers is a smart and preferred option.

Many empirical studies in the past involving structural equation models appear to have ignored the possibility of outliers in data sets. This has been pointed out by Schamberger et al. (2019), who suggests that the behavior, where potential outliers occur, of traditional PLS approach has not yet been studied. In their simulation study it was clearly shown that the traditional PLS methods are susceptible to outliers. These methods tend to create inaccurate estimates when outliers are present, and the results of the traditional PLS approach can also be very distorted for a small number of outliers.

As pointed out by Schamberger et al. (2019), there is a need for the adoption of a more effective and reliable approach to the traditional PLS in overcoming the undesired effects of outliers in the data. In this paper, the robust version of the traditional PLS similar to those proposed by Schamberger et al. (2019) was adopted as an alternative approach. The advantage of using a robust method is that outliers can be handled (i.e., detected and downweighted) automatically without manually removing them from the data. The robust PLSc estimators proposed in this study use a robust correlation estimator (Abdullah, 1990), i.e., Spearman correlation, as an alternative to the classical and non-robust Pearson correlation, in the PLS algorithm. This new mechanism of incorporating robust correlation allows the PLS algorithm and the correction for attenuation applied in PLSc to remain stable while operationalizing the process of identifying potential outliers and producing more accurate parameter estimates of the structural equation model. The formulation of the Robust PLSc as part of the earlier developments in the PLS-based methods is illustrated in Figure 2 below.

Figure 2. The development of Robust PLSc

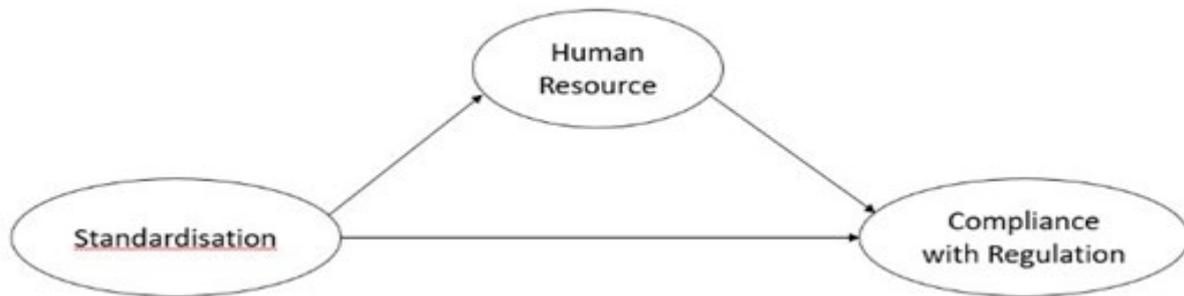


Empirical Analysis

Bahardin et al. (2019) conducted an empirical analysis on the framework in Figure 1. This was part of predictive analytics on the standardization-compliance relationship using the traditional PLS as an empirical approach to verifying the significance of the relationships. The constructs Standardization, Human Resources, Technology and Compliance are modeled as composites.. In total, 20 manifest variables or indicators used for the constructs' operationalization. Each indicator is measured on a 10-point scale ranging from 1 to 10. The data set comprises 240 observations per indicator without any missing values.

The findings from Bahardin et al. (2019) concluded that all the relationships (in Figure 1) are significant, except the relationship between 'Technology' and 'Compliance with Regulations'. This 'insignificance' of the technology-compliance relationship was due to the lack of awareness of industry players and accessibility to the relevant information provided on the digital platform (such as website, social media, etc.) by the local authority, MBPJ. As a result, the research framework in Figure 1 is reduced to a new framework that describes the inter-relationships between the three constructs as displayed in Figure 3 below.

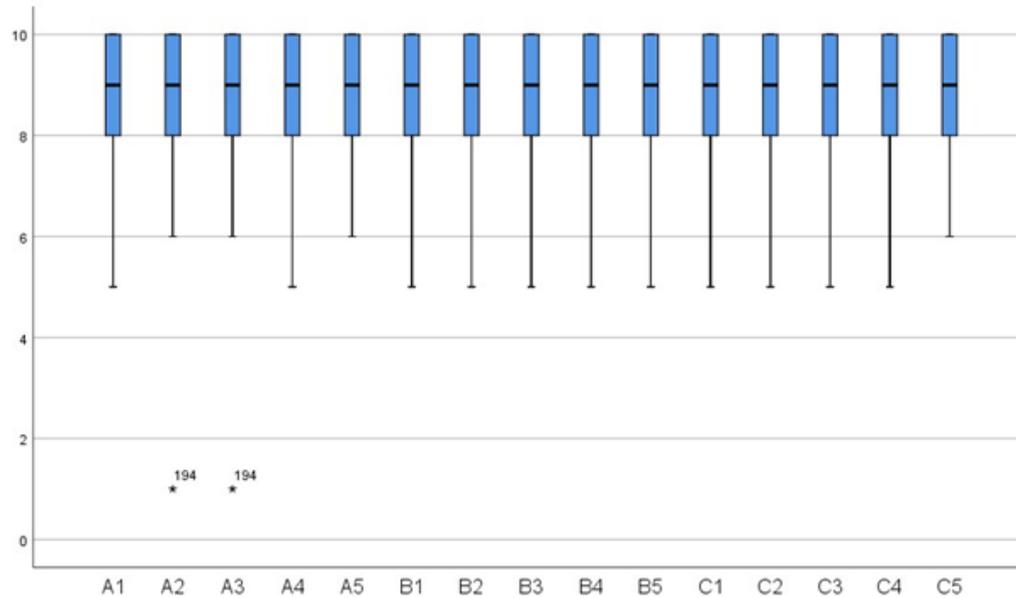
Figure 3. The Research Framework



The purpose of this paper is to present the extension of the empirical study conducted by Bahardin et al. (2019) by excluding the construct ‘Technology’ in the model. Therefore, in this study only three constructs were considered (Figure 3), i.e., standardization of advertisement management (standard), Compliance with Regulations (compliance), and Human Resources (HR), to further the analysis carried out by Bahardin et al. (2019). Some important aspects of predictive analytics which had not been considered and overlooked in their analysis were included in the new analysis. This includes the inspection and the handling of possible outliers in the data set and the adoption of a consistent PLS (Dijkstra and Schermelleh-Engel 2014; Dijkstra and Henseler 2015b) and a robust PLSc (Schamberger et al., 2019), for a more effective and reliable results. In this study, the cSEM package (Schamberger et al., 2019) and R-based ‘Lavaan’ syntax (Rosseel, 2020) were used to compute the robust PLSc.

As a preliminary analysis on the data set used by Bahardin et al. (2019), the detection of possible outliers in the data was carried out using boxplots. Figure 4 displays the boxplot for identifying potential outliers in the indicators associated with each of the constructs in the model.

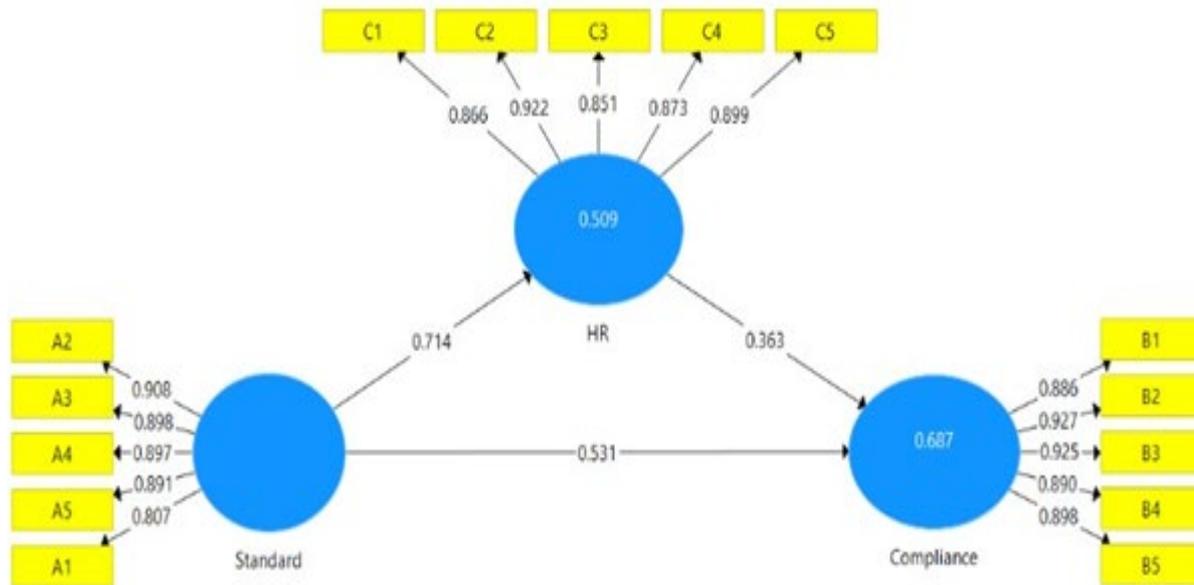
Figure 4. Box Plots for Detection of Outliers



It can be observed in Figure 4 that two potential outliers were present in the construct ‘standardization’ through indicators or manifest variables A2 (observation 194) and A3 (observation 194), respectively. Other indicators or manifest variable associated with other constructs did not indicate any potential outliers. In this paper, the effects of the outliers in A1 and A2 was examined on the goodness of fits and quality criteria of the estimated model based on the traditional PLS, PLSc, and the robust PLSc.

As an illustration the results of the traditional PLS analysis with the three constructs is shown in Figure 5 below. The results were obtained by using the software SmartPLS 3.0 (Ringle et al., 2015).

Figure 5. The estimated traditional PLS model



(Note: A1,..., A5 – indicators/items for Standard; B1,...,B5 – indicators/items for Compliance; and C1,...,C5 – indicators/items Human Resource (HR); D1,...,D5 – indicators/items for Technology)

Some key Goodness of Fits (GoF) and quality criteria in any PLS-SEM analysis are used in this study to evaluate and compare the performances of the traditional PLS, PLSc, and the robust PLSc methods, in the presence of the outliers in indicators A2 and A3, respectively. These criteria are the statistical significance of the path coefficients, coefficients of determination (R^2), and f^2 (the predictive power of independent variables on the dependent variables).

Table 1: The Goodness of Fits of Traditional PLS-SEM, PLSc and Robust PLSc

With Outliers in A2 & A3 Relationship	Traditional PLS			PLSc			Robust PLSc		
	Coeff.	R ²	f ²	Coeff.	R ²	f ²	Coeff.	R ²	f ²
Standard → HR	0.714**	0.509	1.038	0.760**	0.578	1.368	0.769**	0.591	1.447
HR → Compliance	0.363**		0.206	0.374**		0.234	0.342**		0.220
Standard → Compliance	0.531**		0.442	0.546**		0.500	0.595**		0.665
Standard + HR → Compliance		0.687			0.748			0.783	

** Significant at 1% level, * Significant at 5% level

Table 1 above illustrates the performance of the PLS-based methods in estimating the research model in Figure 3. It can be observed that all the methods produced highly significant path coefficients (at 1% level). An obvious improvement in the path coefficients

by the robust PLSc occurred in the “standard – HR” and “standard – compliance” relationships, where the values changed from 0.714 (Traditional PLS) to 0.769 (robust PLSc) and from 0.531 (Traditional PLS) to 0.595 (robust PLSc), respectively. Another notable improvement can also be observed in R^2 , where the robust PLSc raised the coefficient of determinations from 0.509 (Traditional PLS) to 0.591 (robust PLSc) for the endogenous variable, ‘HR’, and from 0.687 (Traditional PLS) to 0.783 (robust PLSc) for the endogenous variable, ‘Compliance’. More specifically, using the robust PLSc, 78.3% of the variations in ‘Compliance’ was due to the variations in the factors ‘standardization’ and ‘HR’. Most notable is that there was an improvement of almost 14% in the R^2 when the robust PLSc was used as compared to the traditional PLS in measuring the R^2 for the endogenous variable ‘Compliance with Advertisement Regulations’.

In addition to R^2 , the criterion f^2 strengthens the argument on the predictability role of the exogenous variables ,i.e., ‘standard’ and ‘HR’, on the endogenous variable, ‘Compliance’. As can be seen in Table 1, the robust PLSc has increased the f^2 values for all the direct relationships. The predictability role of the constructs, ‘standard’ on ‘compliance’, looks so notable when the f^2 values increased substantially from 0.442 (Traditional PLS) to 0.665 (robust PLSc), i.e., an improvement of 50.4% from using the robust PLSc.

Based on the empirical evidence in Table 1, the outliers in A2 and A3 associated with the construct ‘standard’ did have stronger effects on the traditional PLS and the PLSc, while their influence was lesser on the robust PLSc. In other words, the robust PLSc has been able to reduce or downweight the bad influence of the outliers as compared to the non-robust traditional PLS and the PLSc. The results in this study seem to be in line with the findings by Schamberger et al. (2019). However, the slight difference between this study and that of by Schamberger et al. (2019) is in the use of robust correlation coefficient. In this study the robust ‘Spearman Correlation’ (Abdullah, 1990) was adopted, while the robust Minimum Covariance Determinant (MCD) estimator was used in the study by Schamberger et al. (2019).

Conclusions

Predictive analytics is still a new subject in most local governments. This initiative requires certain competencies in data analytics but also the use of efficient computing technology for the speed of processing big data acquired by the local governments. However, the advancement of computing technology has accelerated the availability of tools for large-scale computing, analysis, and visualization. Also, the advent of big data and advances in technological and analytical tools have provided opportunities for addressing various issues faced by the local governments. In this paper, we presented an example of how predictive analytics can benefit the local governments in measuring the effectiveness of advertisement



management framework. The robust predictive modeling proposed in this paper will make it possible for a local government to measure, predict, and track the performance of advertisement management on a more effective and timely manner. The compliance of industry players with the standardization of advertising practices provides the basis for the local governments to develop relevant performance indicators in assessing the effectiveness of the advertisement management framework.

A special focus in this paper was the use of the advertisement management framework that takes into account possible problems in the data such as the presence of outliers or anomalies. We have shown that the use of a proper predictive analytics tool is of utmost important to ensure the credibility of the results of analysis. The empirical evidence presented in this paper clearly showed that the use of robust approach would be the preferred approach in minimizing or even eliminating the bad influence of outliers on the results of predictive analytics.



REFERENCES

- Abdullah, M.B. (1990). On a robust correlation coefficient. *Stat* 39(4):455–460.
- Bahardin et al. (2019). Standardization and Compliance of Advertisement Management Practice by Local Government in Selangor: A PLS-SEM Analysis, in *Proceeding – 20th Kuala Lumpur International Business, Economics and Law Conference*, 261-276.
- Beimborn, D., Gleisner, F., Joachim, N., & Hackethal, A. (2009). The role of process standardization in achieving IT business value. 2009 42nd Hawaii International Conference on System Sciences, 1–10. IEEE.
- Bollen, K.A. (1989). *Structural equations with latent variables*. Wiley, New York
- Cassel, C., Hackl, P., and Westlund, A.H. (2000). On measurement of intangible assets: A study of robustness of partial least squares, *Total Quality Management*, 11(7):897-907.
- Dijkstra, T.K. (2013). A note on how to make partial least squares consistent. <https://doi.org/10.13140/RG.2.1.4547.5688>
- Dijkstra, T.K., Henseler, J. (2015b). Consistent partial least squares path modeling. *MIS Q* 39(2):29–316
- Dijkstra, T.K., Schermelleh-Engel, K. (2014). Consistent partial least squares for nonlinear structural equation models. *Psychometrika* 79(4):585–604
- Fornell, C., and Bookstein, F. L. (1982). “Two Structural Equation Models: LISREL and PLS Applied to ExitVoice Theory,” *Journal of Marketing Research* (19:4), pp. 440-452
- Gideon, R.A., Hollister, R.A. (1987). A rank correlation coefficient resistant to outliers. *J Am Stat Assoc* 82(398):656–666
- Grubbs, F.E. (1969). Procedures for detecting outlying observations in samples. *Technometrics* 11(1):1–21
- Jöreskog, K.G. (1970). A general method for analysis of covariance structures. *Biometrika* 57(2):239–251
- Michael, B., Falzon, J., & Shamdasani, A. (2015). *A Theory of Financial Services Competition, Compliance and Regulation*, University of Hong Kong Faculty of Law Research Paper, 2015/001



- Münstermann, B., & Weitzel, T. (2008). What is process standardization? CONF-IRM 2008 Proceedings, 64.
- Reinartz, W. J., Haenlein, M., and Henseler, J. (2009). "An Empirical Comparison of the Efficacy of Covariance-Based and Variance-Based SEM," *International Journal of Research in Marketing* (26:4), pp. 332-344
- Ringle, C.M., Wende S, Becker JM (2015). SmartPLS 3. <http://www.smartpls.com>, Bönningstedt
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *J Stat Softw* 48(2):1–36, <http://www.jstatsoft.org/v48/i02/>
- Schamberger, T., Schuberth F., Henseler J., Dijkstra T.K. (2020). Robust partial least squares path modeling. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00088-2>
- Spence, T.B., Fanjoy, R.O., and Schreckengast, S.W. (2015). International standardization compliance in aviation, *Journal of Air Transport Management*, Vol. 49, 1-8
- Wold, H. (1975). Path models with latent variables: the NIPALS approach. In: Blalock HM (ed) *Quantitative Sociology*. Academic Press, New York, pp 307–357