Relationship of Big Data Analytics Capability and Product Innovation Performance using SmartPLS 3.2.6: Hierarchical Component Modelling in PLS-SEM

Tan Kar Hooi[#1], Noor Hidayah Binti Abu[#2], Mohd Kamarul Irwan Abdul Rahim[#3]

[#] *School of Technology Management and Logistics, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, MALAYSIA.*

[1]tan_kar@oyagsb.uum.edu.my
[2]dayah@uum.edu.my
[3]mk.irwan@uum.edu.my

***Abstract*— Partial Least Squares Structural Equation Modeling (PLS-SEM) is well-known as the second generation of multivariate statistical analysis to correlate the relationship between multiple variables namely the latent construct. Lately, the popularity using PLS-SEM is growing within the Variance-Based (VB) SEM community. There is still a great number of researcher finding VB-SEM results reporting a daunting task. Ultimately, an advanced PLS-SEM analysis utilizing product innovation performance example with SmartPLS 3.2.6 tool. Higher order construct or hierarchical component modelling is seen as an advanced tool towards the parsimony of the research variables conceptualization.**

***Keywords*—** *Partial Least Squares, Structural Equation Modelling, PLS-SEM, SmartPLS 3.2.6, Big Data Analytics Capability, Product Innovation Performance, Higher-order Constructs*

1. Introduction

Big Data is a big word recently being utilized prevalently which being described as huge volumes of high velocity, complicated and variable data which needs state-of-the-art methodologies and know-hows to allow the apprehension, packing, circulation, managing, and data study.

Since the field of Big Data Analytics is quite new, research examining its use and effects is still quite limited. However, more and more organizations began to adopt Big Data Analytics to better understand their customers and to optimize customer engagement [1]. Big Data Analytics distinguishes itself from Traditional Marketing Analytics in the four Vs of data: volume, velocity, variety, and veracity [1], and has the potential to improve business decision-making for better NPP. Here, volume indicates that the scale of the data typically ranges from gigabytes to terabytes; variety pinpoints data from social media, social sensors, transaction, and other sources; velocity implies that different sources generate continuous streams of data; veracity deals with quality or uncertainty of the data [1]

Extraordinary step of destructive innovation is resulted from the blending of digitalized technologies and subliminal technologies, global economic giants reshuffle and international crowdsourcing uprise [2]. The disruption has caused the unprecedented adversities in the entire business environment. It has coupled with the marketing and new product development being the front liner of the cultural shock. The initiatives to transform Big Data into business indicators and business performance predictions has made its difficulties unveiled [3]. Jim Gray charts the evolution of science as a process of four broad paradigms: experimental science, theoretical

International Journal of Supply Chain Management
IJSCM, ISSN: 2050-7399 (Online), 2051-3771 (Print)

science, computational science, and exploratory science [4]. The last paradigm features a new mode of ongoing proliferation of data and economic thinking decision making framework, with less emphasis on theory-laden hypotheses [5]. Detailed explanations regarding this are explained later in the taxonomy development section.

Organizations could perform real-time monitoring based on the business performance dashboard data resulted from the huge chunks of data in the beginning entry of Big Data analysis [6]. Big data has now becomes a trend in this data disruption era with increasing number of applications in a varied business function. Data analysis speed and accuracy already become indicators for business competitive advantage [7]. The ever-increasing flood of data range from Megabyte, Gigabyte, Terabyte, to Zettabyte provides a clear message to the C-suite: organize or die [8]. On the one hand, it has been reported that firms utilize less than 12% of their data. On the other hand, a lack of analytics tools and repressive data silos leads companies to ignore 88% of their customer data [1]. With more people keeping much of their lives and contributing their word of mouth online, Big Data Analytics is becoming a key agenda for firms when they consider launching new products.

Confirmatory nature analytical tools (Variance-based Structural Equation Modelling) such as AMOS and LISREL are not suitable to this research due to multivariate normality issue[9]. Partial Least Squares Structural Equation Modeling (PLS-SEM) can be used in both exploratory and confirmatory research setting with both relective and formative consturcts and the data normality is not an issue [10].

PLS-SEM has a long history starting 1960s but it is not getting interest from researchers and academicians [11]. Prof Wyne Chin has popularize the use of PLS by creating a research software for his doctoral studies namely PLS-Graph which the user interface has only limited function and not so user-friendly. WarpPLS is a powerful tool to handle non-linear structural equation models developed by Prof Ned Kock. SmartPLS has now becomes the most user-friendly software for all the researches and academicians developed by Christian Ringle, Sven Wende, Jan-Michael Becker. Many new comers in the PLS-SEM community found analysis and reporting parts alienated to them especially the hierarchical component modelling. Hence, this research exemplified the PLS-SEM analysis through a survey done in the Malaysian Electrical and Electronics Industry. SmartPLS 3.2.6 software is used to perform measurement and structural model determination [12]. Microsoft Excel and IBM SPSS will be used to perform data preparation and data cleaning.

2. Conceptual Framework and Research Hypotheses

In this research, the emphasis was put on the product innovation performance with its enabler namely big data analytics capability in the Malaysian electrical and electronics companies.

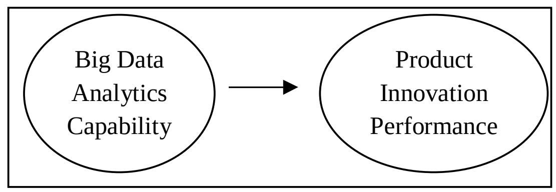


Figure 1: Conceptual Model

Operationalization of variables is about measurement of research variables [13] Through operationalization, research variables are reduced to constructs which can be observed and measured [13]. In this section, the variables examined in this study are operationalized. Thereafter, summaries of operationalization of the variables examined in this study are presented in a tabular form.

Product innovation performance. Measures for subjective product innovation performance were mostly adapted from [14] supplemented by one item suggested from the pretest. The scale is to measure management's perception of market performance of new products. Informants were asked to evaluate the contribution of new products less than three years old to sales volume, profitability, and customer satisfaction relative to their competitors and their original objectives. 5-point scales were used, with answers ranging from 1 (strongly disagree) to 5 (strongly agree).

Big data analytics capability scale was adapted from [15] The version used in this study was divided into 4 components which consists of 4 items each of the items were then rated using a 5-point Likert scale, ranging from (1) for strongly disagree to (5) for strongly agree.

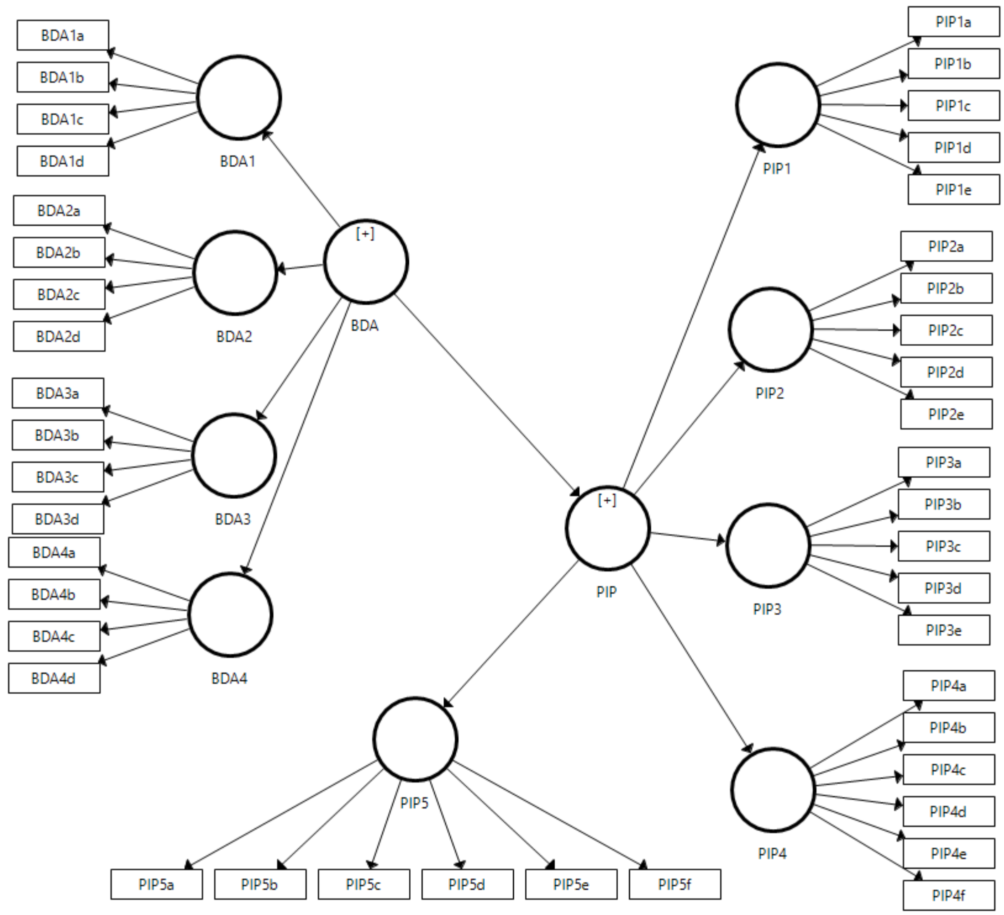


Figure 2: Statistical Model

Table 1: Product Innovation Performance Higher Order Construct

Variable	Operational Definition	Dimension	Items	Authors
Product Innovation Performance	Hard financial aspects and soft non-financial aspects of the organization's product innovation attempts.	Financial performance (Refers to Return on Sales (ROS), Return on Investments (ROI) and Return on Assets (ROA) are favored for performance evaluation)	Return on sales (profit/total sales) attributable to new products are higher than those provided by the remaining products Return on assets (profit/total assets) attributable to new products are higher than those provided by the remaining products New products have achieved the objectives set in terms of return on investment (ROI) Profits attributable to new products are higher than those provided by the remaining products New products have achieved the objectives set in terms of profit	[16], [17]
		Market performance (Refers to the end results of	New product sales are greater than those provided by the rest of the products New products have achieved the	[16]

	these policies—the relationship of selling price to costs, the size of output, the efficiency of production, progressiveness in techniques)	objectives set in terms of sales Compared with other products of your company, new products have achieved superior results in terms of market share New products have achieved the objectives in terms of market share New products have allowed the penetration of new markets	
	Technical performance (Refers to the production process outcome aspects such as manufacturing technology, quality as well as the functioning of the product)	The quality of new products is better than the rest of the products Decreasing manufacturing cost in components and materials of current products New products are launched in the deadlines and within budget Development Goals New products have reduced environmental damage, improved health and safety Developing new products with technical specifications and functionalities totally differing from the current ones	[16], [17]
	Customer performance (It is a measurable monetary or non-monetary result of a customer relationship in a defined period)	Customers are satisfied with the performance of new products Compared with other products of your company, customer complaints regarding new products are fewer New products have improved customer loyalty We actively and regularly seek customer input to identify their needs and expectations We involve customers in our product design processes The number of new products which are developed by knowledge from customers is higher in last three years	[16], [17]
	Strategic performance (Refers to business owners and managers develop activities or tasks to gauge the overall effectiveness and efficiency of their company)	New products provide the company a competitive advantage New products have reached all the goals set New products have improved the reputation of the company Our firm tries to sort out the business operation into statistics for analysis Our firm always does detailed analysis before making crucial business decisions Our firm invests the long-term project (basic research, etc.) to obtain competitive advantage in the future	[16], [17]

Table 2: Big Data Analytics Capability Higher Order Construct

Variable	Operational Definition	Dimension	Items	Authors
Big Data Analytics (BDA) Capability	Competence to provide business insights using data management, infrastructure (technology) and talent (personnel) capability to transform business into a competitive force.	Big data operations (firm's capability of identifying sources where large volumes of various kinds of data flow out in high speed, collecting, storing, and analysing such Big Data for accomplishing the firm's strategic as well as operational goals)	We can identify sources of big data that meet our needs. We can collect big data that meet our needs. We can store large volumes of data. We can process big data with a fast speed.	[15]
		Updating IT infrastructure (Applications, hardware, data, and networks to enable the BDA staff to quickly develop, deploy, and support necessary system components for a firm.)	We adopt state of the art technologies to process big data. We constantly update our computing equipment to process big data. We constantly update our IT architecture to process big data. We constantly update our IT infrastructure to process big data.	[15]
		Advanced Analytics (BDA staff's professional ability, e.g., skills or knowledge to undertake assigned tasks)	We are good at data analytics which is mainly data mining and statistical analysis. We are good at text analytics that deals with unstructured textual format data. We are good at web analytics that deals with web sites. We are good at mobile analytics that deals with mobile computing.	[15]
		Strategic Uses of Big Data (ability to handle routines in a structured (rather than ad hoc) manner to manage IT resources in accordance with business needs and priorities)	We rely on Big Data to identify new business opportunities. We rely on Big Data to develop new products. We rely on Big Data to enhance our innovativeness. We rely on Big Data to formulate our business strategy.	[15]

A total of 281 questionnaires are received from electrical and electronics companies across Malaysia; missing data had been labelled with markers prior to measurement model PLS path analysis.

Once the conceptual framework is finalized, the next step is hypothesis development. The hypothesis is developed to explore the relationship between big data analytics capability and product innovation performance.

H1: Big data analytics capability is positively related to product innovation performance.

3. Data Analysis

We utilized second generation multivariate statistical tools namely PLS-SEM to test the model developed in this study [18] which could test the entire complex model in one go by correlating the latent variable and measurement items together [18]. SmartPLS 3.2.6 [19] is an ideal tool to analyze the data collected. [18] suggested to perform bootstrapping method (500 resamples) to ascertain the levels of significance for loadings and path coefficients (beta values) and also reducing the standard error of the t-values.

Common method variance refers to the variance caused by the measurement method rather than the latent construct or measurement items itself [20]. This happens usually when the questionnaire is being measured and answered by the same respondent, for instance, the factors attributed to job satisfaction towards performance being answered by the same person. Statistical and procedural remedies are being found in the literature to counter common method variance. Harman's single factor test is being utilized to detect common method variance. Using SPSS, principal component analysis (PCA) is executed by selecting all the measurement items results in each variable [21]. When the general factor attributed to the majority of the covariance, the variables all load on one factor or one factor explains the majority of the variance, common method variance may be a problem [20]

In the principal component analysis, the results showed a seven-factor key. A total variance explained of 66.247 per cent and the first factor only explained 38.781 per cent. It denotes that common method bias is not a serious issue in this study.

4. Hierarchical Component Modelling

By referring to [22], PLS-SEM has an advance topic known as the hierarchical component modelling (HCM) which consisting of the observable lower order component (LOC) and non-observable higher order component (HOC). The advantageous of higer order construct is to reduce bias due to multicollinearity and potential discriminant validity issues eradication [23]

Explicit depictions of multidimensional constructs in a higher representation level and are related to other constructs at a similar representation level completely arbitrating the effect from or to their underpinning dimensions are known as Hierarchical latent variable models, hierarchical component models, or higher-order constructs [22]. [24] define "[...] a construct as multidimensional when it consists of a number of interrelated attributes or dimensions and exists in multidimensional domains. In contrast to a set of interrelated unidimensional constructs, the dimensions of a multidimensional construct can be conceptualized under an overall abstraction, and it is theoretically meaningful and parsimonious to use this overall abstraction as a representation of the dimensions." Uni-dimesion and multi-dimension are distinct due to the difference in their underlying construct [25]

A conceptual variable is located within the theoretical layer whereby its definition lies. Reflective and formative conceptualization of the conceptual layers is based on the operationalized definition. The ways we operationalize the construct determines the measurement model indicators characteristics namely effect, causal or composite indicators [26].

Generally, hierarchical latent constructs are often determined by the number of levels and usually up to second-order construct [27]and formative or reflective conceptualization amongst the construct within the research model [12]

In this research, reflective-reflective hierarchical component modelling (rr-HCM) is being conceptualized in this research. PLS-SEM requires the latent variable scores (LVS) computation for each latent variable in the path modelling.

There are two mainstream methods being introduced for higher order modelling namely (1) the repeated indicator approach [22] and (2) the sequential latent variable score method or two-stage approach [12]

A higher-order latent construct is created by requiring a latent construct that epitomizes all the observable variables of the underpinning lower-order latent construct for the repeated indicator approach [22]. For instance, if second order construct having five dimension. Each dimension having five manifest variables, the second-order latent variables, it can be determined using all (twenty-five) manifest variables of the underpinning first-order latent construct.

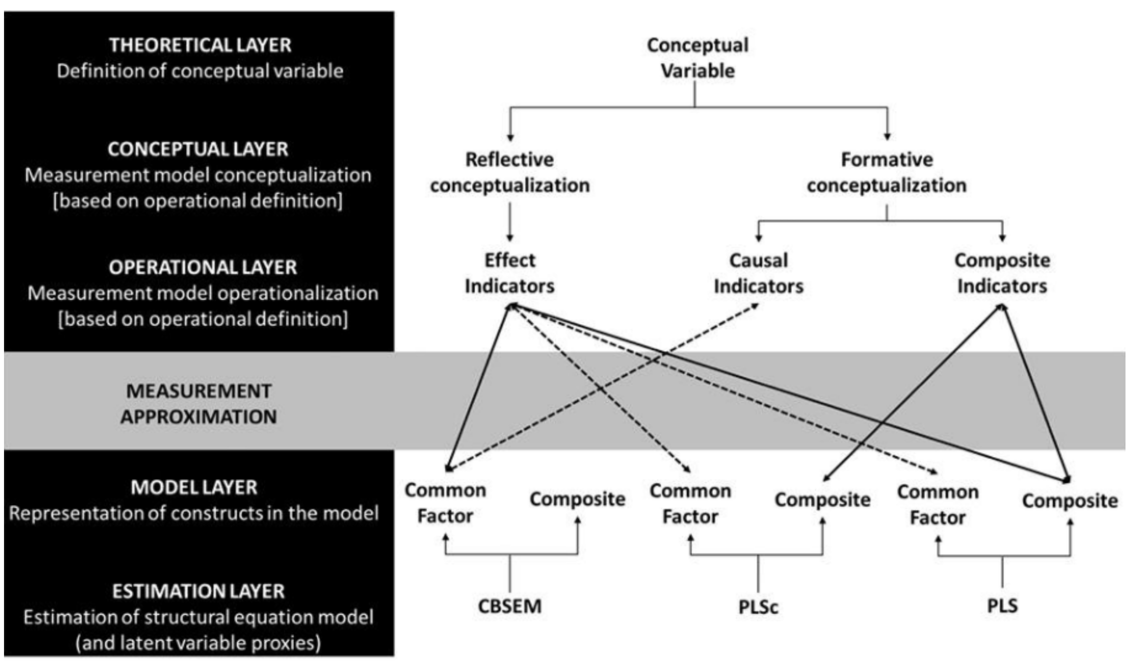
Ultimately, the manifest variables are used two times. Firstly, these variables are utilized in the first-order latent variable ("primary" loadings/weights) and secondly for the second-order latent variable ("secondary" loadings/weights).

The structural model which caters for the higher order constructs of the model, as the path coefficients between the first-order and second-order constructs represent the loadings/ weights of the second-order latent variable.

Apparently, this method can be swiftly be prolonged to higher-order hierarchical component modelling [28]

The capability to approximate all constructs concurrently rather than approximating the lower-order and higher-order dimensions discretely are the benefits of the repeated indicator approach. Therefore, it requires the entire nomological system, including the higher order and lower order of constructs. It is crucial to prevent conceptualization perplexing. Researchers have to self-determined the measurement modes for the higher-order construct and the inner weighting when utilizing the repeated indicator approach.

In any PLS-SEM models, the measurement modes for the higher-order repeated indicators needs to be pre-determined (i.e., Mode A – reflective construct or Mode B – formative construct). The typical tactic for repeated indicators on a hierarchical latent variable is by utilizing Mode A [11] which typically suitable to be used in reflective-reflective type models.



Notes: Dashed lines indicate acceptable types of measurement approximation; solid lines represent recommended types of measurement approximation. The PLSc results when estimating composite model data and composite indicators parallel those from PLS as no correction for attenuation occurs.

Figure 3: Measurement and model estimation framework *(Adopted from Sarstedt et al. 2016; Estimation Issues with PLS and CBSEM: Where the Bias Lies!)*

5. Results and Discussion

This study utilized the structural equation modelling technique to analyze the linkages between big data analytics capability and product innovation performance. Table 3 shows the internal consistencies or reliability of the measurement items.

In structural equation modelling, we should first assess the measurement model then only followed by the structural model.

Convergent validity is the degree to which multiple items to measure the same concept are in agreement. As suggested by [23] we used the factor loadings, composite reliability (CR) and average variance extracted (AVE) to assess convergent validity. The recommended values for loadings are set at >0.5, the AVE should be >0.5 and the CR should be > 0.7.

From Figure 1, it can be seen that we have conceptualized BDAC (Big Data Analytics Capability) and PIP (Product Innovation Performance) as second-order constructs. Thus, we followed the method suggested in the literature in PLS which is the repeated indicator approach to model the second-order factors in the PLS analysis. Table 4 shows that the results of the measurement model exceeded the recommended values, thus indicating sufficient convergence validity (Table 5).

After confirming the convergent validity, we proceeded to assess the discriminant validity using the [29] method. Discriminant validity is the degree to which items differentiate among constructs or measure distinct concepts.

The criterion used to assess this is by comparing the AVE with the squared correlations or the square root of the AVE with correlations. As shown in Table 4, we have used the second method which is to compare the square root of the AVE with the correlations. The criteria is that if the square root of the AVE, shown in the diagonals, is greater than the values in the row and columns on that particular construct, then we can conclude that the measures are discriminant.

From Table 5, it can be seen that the values in the diagonals are greater than the values in their respective row and column, thus indicating that the measures used in this study are distinct, demonstrating adequate discriminant validity.

After all the measurement model criteria, had passed, then we only can proceed to structural model in Partial Least Squares (PLS). To evaluate the structural models' predictive power, we calculated the R2. R2 indicates the amount of variance explained by the exogenous variables [30].

All the variables together explained 40.5 per cent of the variance. Using a bootstrapping technique with a re-sampling of 500, the path estimates and t-statistics were calculated for the hypothesized relationships.

Table 6 shows the structural model analysis. From the analysis, it was found that BDAC ($\beta = 0.636$, $p = 0.01$) was positively related to PIP.

Table 3: Internal Consistencies/ Reliability of the measurement items

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
BDA	0.915	0.918	0.927	0.497
BDA1	0.735	0.742	0.851	0.656
BDA2	0.809	0.812	0.875	0.636
BDA3	0.828	0.831	0.886	0.661
BDA4	0.773	0.775	0.898	0.815
PIP	0.958	0.960	0.962	0.526
PIP1	0.890	0.893	0.924	0.752
PIP2	0.931	0.931	0.948	0.784
PIP3	0.891	0.892	0.948	0.901
PIP4	0.885	0.888	0.913	0.636
PIP5	0.882	0.886	0.911	0.630

Table 4: Results of structural Relationship of the proposed model

First Order Construct	Second Order Construct	Item	Loadings	AVE	CR
Big data operations		BDA1a	0.839	0.656	0.851
		BDA1b	0.858		
		BDA1c	0.727		
Updating IT infrastructure		BDA2a	0.782	0.636	0.875
		BDA2b	0.797		
		BDA2c	0.829		
		BDA2d	0.781		
Advanced Analytics		BDA3a	0.751	0.661	0.886
		BDA3b	0.864		
		BDA3c	0.849		
		BDA3d	0.784		
Strategic Uses of Big Data		BDA4a	0.897	0.815	0.898
		BDA4b	0.908		
	Big Data Analytics Capability	Big data operations	0.798	0.728	0.914
		Updating IT infrastructure	0.862		
		Advanced Analytics	0.904		
		Strategic Uses of Big Data	0.845		
Financial		PIP1a	0.829	0.752	0.924
		PIP1b	0.881		
		PIP1c	0.899		

		PIP1e	0.857		
Market		PIP2a	0.853	0.784	0.948
		PIP2b	0.902		
		PIP2c	0.895		
		PIP2d	0.923		
		PIP2e	0.851		
Technical		PIP3a	0.947	0.901	0.948
		PIP3b	0.952		
Customers		PIP4a	0.758	0.636	0.913
		PIP4b	0.821		
		PIP4c	0.719		
		PIP4d	0.820		
		PIP4e	0.828		
		PIP4f	0.832		
Strategic		PIP5a	0.708	0.63	0.911
		PIP5b	0.765		
		PIP5c	0.854		
		PIP5d	0.830		
		PIP5e	0.799		
		PIP5f	0.797		
	Product Innovation Performance	Financial	0.884	0.722	0.928
		Market	0.881		
		Technical	0.736		
		Customers	0.858		
		Strategic	0.880		

Table 5: Convergent Validity

	1	**2**
1. BDA	0.705	
2. PIP	0.636	0.725

Table 6: Structural Model

		Standardized Beta	**Standard Error**	**t value**	**Decision**	**LL**	**UL**	**f2**	**VIF**
H1	BDA ➔ PIP	0.636	0.041	15.57**	Supported	0.29	0.497	0.68	1.000

*Notes: **p < 0.01; *p < 0.05*

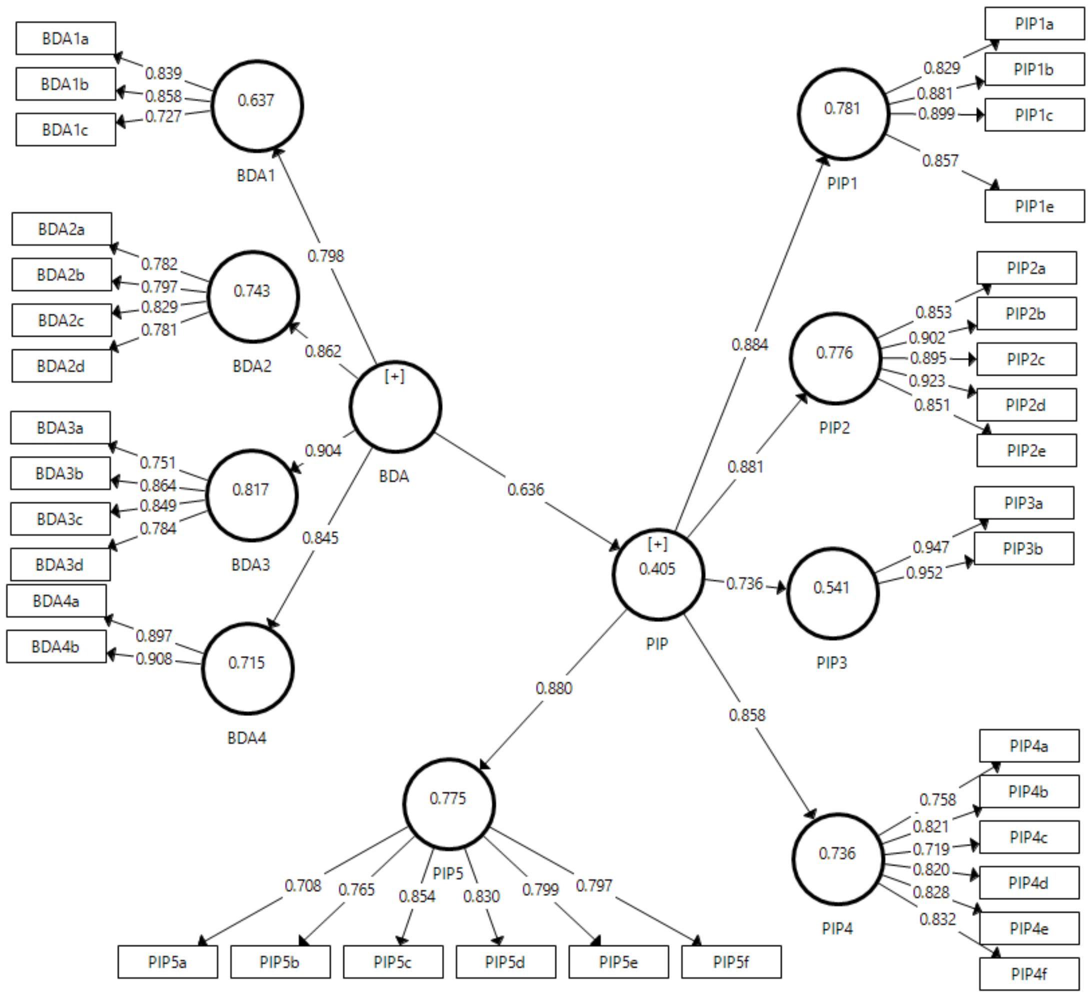


Figure 4: Measurement Model of the Study

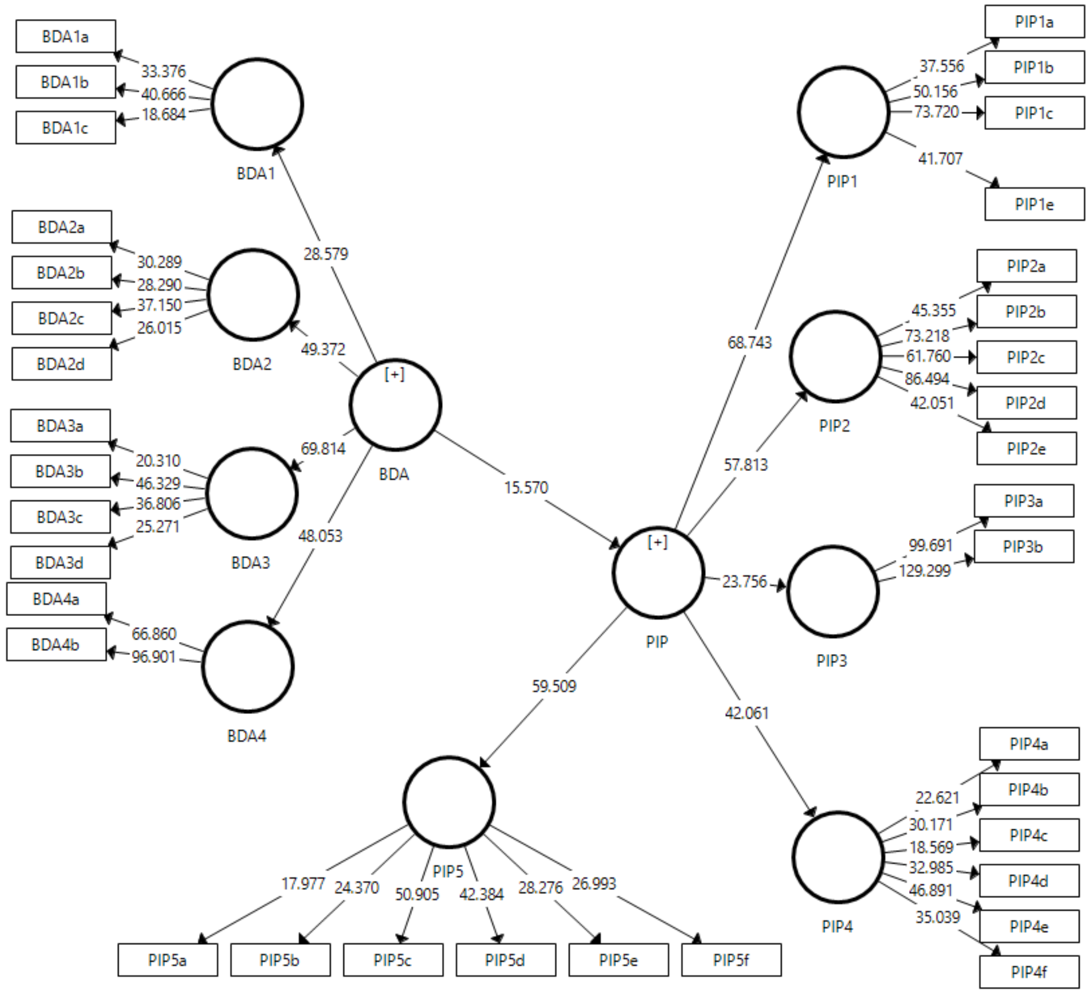


Figure 5: Structural Model of the Study

6. Conclusion and Implication

Although several researchers have provided empirical evidences on the linkage between big data analytics capability and performance, some might have overlooked that performance indicator varies. BDA is now considered as a game changer enabling improved business efficiency and effectiveness because of its high operational and strategic potential. The emerging literature on BDA has identified a positive relationship between the deployment of customer analytics and firm performance [31] For example, BDA allows firms to analyze and manage strategy through a data lens [32]. Indeed, BDA is increasingly becoming a crucial component of decision-making processes in businesses [33]. BDA is now considered as "a major differentiator between high performing and low-performing organizations," as it allows firms become proactive and forward-looking, decreases customer acquisition costs by about 47% and enhances firm revenue by about 8% [34]. In manufacturing and operations management [35], BDA is considered to be an enabler of asset and business process monitoring [36], supply chain visibility, enhanced manufacturing and industrial automation [37] and improved business transformation [38].

Acknowledgments

I would like to extend my deepest gratitude to School of Technology Management and Logistics, UUM especially to my supervisor and co-supervisors for insights, suggestions, and thoughtful review of this paper.

References

[1] Aut Forrester (2014). Expand your digital horizon with big data. Forrester, May 27 *http:// www.asterdata.com/newsletter-images/*

[2] Annunziata, M. (2015). Innovation barometer: Businesses disruption-ready, in right conditions. Ideas Lab, January 2014 *,http://www.ideaslaboratory.com/post/108552544328/innovation-barometer-businesses-disruption-ready-in-right*, last accessed Jan 28th, 2015 Monden, Y. 1981. Smoothed production lets Toyota adapt to demand changes and reduce inventory. *Industrial Engineering*. 13 (8): 42– 51

[3] Xu, Z., Frankwick, G. L., & Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research,* 69(5), 1562-1566.

[4] Xu, Z., et al (2015)., Effects of big data analytics and traditional marketing analytics on newproduct success: A knowledge fusion perspective, *Journal of Business Research*, *http://dx.doi.org/10.1016/j.jbusres.2015.10.017*

[5] Kitchin, R. (2015). Making sense of smart cities: addressing present shortcomings. *Cambridge Journal of Regions, Economy and Society*, 8(1), 131-136.

[6] Gustke, C. (2013). Big data takes turn as market darling. CNBC, Apr 15 *http://www.cnbc.com/id/100638376 Access: Apr 21st, 2014*

[7] Lamb, J. (2014). Need for speed in data analysis. Raconteur, Sep 9 http://raconteur.net/ technology/need-for-speed-in-data-analysis last accessed Dec 21st, 2014

[8] Cook, T. D. (2014). "Big data" in research on social policy. *Journal of Policy Analysis and Management*, 33(2), 544-547.

[9] Wong, K. K. K. (2016). Mediation analysis, categorical moderation analysis, and higher-order constructs modeling in Partial Least Squares Structural Equation Modeling (PLS-SEM): A B2B Example using SmartPLS. *Marketing Bulletin, 26.*

[10] Reinartz, W. J., Haenlein, M., & Henseler, J. (2009). An Empirical Comparison of the Efficacy of Covariance- Based and Variance-Based SEM. *International Journal of Market Research*, 26(4), 332–344.

[11] Wold, H. (1985). Partial Least Squares. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of Statistical Sciences* ,Vol. 6, pp. 581–591, New York: John Wiley & Sons

[12] Ringle, C. M., Wende, S., & Becker, J. M. (2015). *SmartPLS 3*. Boenningstedt: SmartPLS GmbH, http://www. smartpls. com.

[13] Sekaran, U., & Bougie, R. (2013). *Research methods for business: A skill-building approach.*

[14] Im, S., Bayus, B. L., & Mason, C. H. (2003). An empirical study of innate consumer innovativeness, personal characteristics, and new-product adoption behavior. *Journal of the Academy of Marketing Science*, 31(1), 61-73.

[15] Lin, C., Song, Z., Song, H., Zhou, Y., Wang, Y., & Wu, G. (2016). Differential privacy preserving in big data analytics for connected health. *Journal of Medical Systems*, 40(4), 97.

[16] Hannachi, Y. (2015). Development and

Validation of a Measure for Product Innovation Performance: The PIP Scale. *Journal of Business Studies Quarterly*, *6*(3), 23.

[17] Gunday, G., Ulusoy, G., Kilic, K., & Alpkan, L. (2011). Effects of innovation types on firm performance. *International Journal of production economics*, *133*(2), 662-676.

[18] Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the academy of marketing science*, *40*(3), 414-433.

[19] Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115-135.

[20] Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology*, *88*(5), 879.

[21] Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of management*, *12*(4), 531-544.

[22] Lohm€oller, J.-B., (1989). *Latent Variable Path Modeling with Partial Least Squares.* Physica-Verlag, Heidelberg, Germany.

[23] Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). *Partial least squares structural equation modeling: Rigorous applications, better results and higher acceptance.*

[24] Chin, W. W. (1998). The partial least squares approach to structural equation modeling. Modern methods for business research, 295(2), 295-336.

[25] Maxham III, J. G., & Netemeyer, R. G. (2003). Firms reap what they sow: the effects of shared values and perceived organizational justice on customers' evaluations of complaint handling. *Journal of Marketing*, *67*(1), 46-62.

[26] Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies!. *Journal of Business Research*, *69*(10), 3998-4010.

[27] Rindskopf, D., & Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research*, *23*(1), 51-67.

[28] Wetzels, M., Odekerken-Schröder, G., & Van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS quarterly*, 177-195.

[29] Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of marketing research*, 382-388.

[30] Barclay, D., Higgins, C., & Thompson, R. (1995). *The Partial Least Squares (pls) Approach to Casual Modeling: Personal Computer Adoption Ans Use as an Illustration*.

[31] Germann, F., Lilien, G. L., Fiedler, L., & Kraus, M. (2014). Do retailers benefit from deploying customer analytics?. *Journal of Retailing*, *90*(4), 587-593.

[32] Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356-365.

[33] Hafen, E., Kossmann, D., & Brand, A. (2014). Health data cooperatives—citizen empowerment. *Methods Inf Med*, *53*(2), 82-86.

[34] Hämäläinen, T. J. (2015). Governance solutions for wicked problems: Metropolitan innovation ecosystems as frontrunners to sustainable well-being. *Technology Innovation Management Review*, *5*(10).

[35] Liu, V., Lei, X., Prescott, H. C., Kipnis, P., Iwashyna, T. J., & Escobar, G. J. (2014). Hospital readmission and healthcare utilization following sepsis in community settings. *Journal of hospital medicine*, *9*(8), 502-507.

[36] Davenport, T. H., & Patil, D. J. (2012). Data scientist. *Harvard business review*, *90*(5), 70-76.

[37] Chiolero, A. (2013). Big data in epidemiology: too big to fail?. *Epidemiology*, *24*(6), 938-939.

[38] Gardner, E. (2013). The HIT approach to big data. *Health data management*, *21*(3), 34-36.