GLOBAL BUSINESS MANAGEMENT REVIEW
http://e-journal.uum.edu.my/index.php/gbmr

How to cite this article:
Pulka, B. M. (2022). Preliminary Analysis of the Determinants of SMEs Performance. *Global Business Management Review, 14*(1),1-19. https://doi.org/10.32890/gbmr2022.14.1.1

PRELIMINARY ANALYSIS OF THE DETERMINANTS OF SMES PERFORMANCE

Buba Musa Pulka

Department of Business Administration, Faculty of Management Sciences,
University of Maiduguri, Nigeria.
Corresponding author: bubapulka@unimaid.edu.ng

Received: 15/01/2022 Revised: 5/7/2022 Accepted: 10/7/2022 Published: 15/07/2022

ABSTRACT

Preliminary analysis is the inspection, scrutiny and analysis conducted on data before main analysis, to detect, manage and correct/treat errors. Preliminary analysis helps in data screening, cleaning, initial examination and correcting/treating incorrect entries, identify outliers, missing values and to identify other unusual entries in dataset. Although, preliminary analysis is vital in academic research, majority of researchers partially report, or entirely do not perform and report this stage in their studies. Evading this critical stage of research could lead to poor quality and incorrectness of the research results and misinterpretation due to under or over inflation of statistical results. Unexpectedly, the review of the extant literature discovered that there is scarcity of studies that have been conducted and reported on preliminary analysis. This has limited scholars, educators and other stakeholders in understanding the importance of conducting and reporting preliminary analysis in enhancing reliability and accuracy of research results, findings, interpretation and implications. Consequently, this research filled in these gaps by developing a framework and providing empirical evidence on these issues. This would be of great benefits to the academic world. Thus, the results indicated that the data had met all the criteria and assumptions for multivariate analysis after statistical analysis and treatment. Thus, it is recommended that studies in entrepreneurship, social sciences, management and other related disciplines should apply and follow this procedure to help in meeting the criteria for multivariate data analysis.

Keywords: data screening, cleaning, preliminary analysis, SMEs performance.

INTRODUCTION

Preliminary analysis is the inspection, scrutiny and analysis conducted on data of research before the main analysis to detect, manage and correct/treat errors therein. Preliminary data screening, cleaning and analysis, is required to be performed to fulfil the prerequisite to the main analysis using multivariate data and analysis (Hair Jr, William, Babin, & Anderson, 2014; Ibrahim & Shariff, 2014; Abduwahab, Dahalin, & Galadima, 2011). The main purpose of the preliminary data analysis is to check the data and to prepare it for further analysis. It is also used for describing the characteristics of the data, and to summarise the results (Blischke, Karim, & Murthy, 2011). Preliminary analysis is vital in any given research. This because it gives the opportunity to the researchers to examine, screen and prepare the data for analysis. This process has been emphasised by many researchers (Hair, Hult, Ringle, & Sarstedt, 2017; Hair, Ringle, & Sarstedt, 2013). According to Blischke et al., (2011) data screening, cleaning and preliminary analysis help the researchers to; (1) verify the source of the data. (2) Verify that the data include the variables specified. (3) Verify the units of measurement. (4) Clean the data by deleting or, if possible, correcting/treating incorrect entries. (5) Identify outliers or otherwise unusual results. (6) Check for missing data and to identify any other unusual data features. Though preliminary analysis vital in academic research, the majority of researchers partially report or entirely do not conduct and report data screening, cleaning and preliminary analysis. Avoiding this critical stage of research would lead to poor quality and inaccuracy of the research results and findings.

Unexpectedly, the assessment of the existing literature revealed that there is a scarcity of understanding and of research that has been carried out and reported on data screening, cleaning and preliminary analysis particularly in the context of Nigeria (Ibrahim & Shariff, 2014). This has immensely put some restrictions to the scholars, educators, practitioners, policy makers and other stakeholders in comprehending the significance of conducting and reporting data screening, cleaning and preliminary analysis in enhancing the reliability and accuracy of the research results, findings, interpretation and implications.

Thus, this research work therefore provided; firstly, developing a framework that would serve as a guide for data screening, cleaning and preliminary analysis. Secondly, by providing empirical evidence on data screening, cleaning and preliminary analysis that would be of benefits to scholars, educators, practitioners, policy makers and other stakeholders. These would go a long way in enhancing the understanding and importance of conducting and reporting data screening, cleaning and preliminary analysis in any given research work to enhance the reliability and accuracy of the research results, findings, interpretation and implications.

Thus, the main objective of this study is to present a procedure and results of data screening, cleaning and preliminary analysis. While the specific objectives are to:

1. Examine the response rate of the respondents.
2. Detect and manage the missing data in the data set.
3. Analyse and manage outliers in the data set.
4. Analyse and determine data normality.
5. Analyse and manage multicollinearity
6. Analyse non-response bias
7. Analyse and manage common method variance.

LITERATURE REVIEW (FIRM RESOURCES)

Since the study, the focus is on the data screening, cleaning and preliminary analysis; it is essential to give a brief background of the variables used in the study. The study utilised entrepreneurial competencies (EC), entrepreneurial orientation (EO), information and communication technology (ICT), entrepreneurial network (EN) and government business support (GBS) as independent variables (firm resources). Similarly, the study has SMEs performance (SP) as the dependent variable, while the external environment (EE) is a moderator.

Firstly, entrepreneurial competencies (EC) is defined as the sum of individuals' knowledge, experience, skills, and attitudes that are acquired over a period that enhances the effectiveness of the firm performance (Kaur & Bains, 2013). Secondly, entrepreneurial orientation (EO) is the process of development of specific behaviour that is targeted towards increasing SMEs propensity to take and absorb risk, to be creative, initiative, proactive and innovative toward achieving success (Covin & Slevin, 1991). Thirdly, ICT is defined in this study as the ICT resources use by the firms that include that improve the production, communication, process and design of SMEs for generating high productivity, efficiency, innovation, strengthening the customer-supplier relationship, increase the profitability of the firms and improvement in the decision-making process and overall firm performance (Bayo-Moriones, Billon, & Lera-Lopez, 2013).

Fourthly, entrepreneurial network (EN) is defined as the relationship/interactions of the SMEs with other firms, customers, suppliers, social organizations, trade organizations, research institutions, financial institutions, support institutions, friends and close relatives that provide resources and or access to the resources needed by SMEs in the external environment to improve its competitive advantage and for to achieve high SMEs performance (Pulka, 2019). Fifthly, government business support (GBS) defined as the financial and strategic support provided by institutions, organisations and agencies set up by the government to support, promote and regulate SMEs activities to enable them to achieve growth, development and high performance capable of providing employment opportunities (Pulka, 2019).

Sixthly, the external environment is defined as the degree to which the activities in the external environment of firms provide opportunities and resources, and be able to determine SMEs achievement, success and overall performance (Aminu, 2015b). Finally, SMEs performance is defined as the abilities of the SMEs to harness, integrate and utilise various internal and external resources with timely and right reconfiguration to achieve targeted set of objectives and performance capable of providing employment opportunities, the growth of GDP, export and to uplift the standard of living of the society (Pulka, 2019).

FRAMEWORK FOR DATA SCREENING, CLEANING AND PRELIMINARY ANALYSIS

Though some researchers have reported data screening, cleaning and preliminary analysis (Kura, 2014; Shamsudeen, Keat, & Hassan, 2016; Zakaria, 2016), nevertheless, there is dearth of research that has been published that has stated clearly the procedure for data screening, cleaning and preliminary analysis as in this study. Therefore, this study composed the framework for data screening, cleaning and preliminary analysis based on the previous literature. Thus, the conceptual framework is made up of various components that make the procedure for data screening, cleaning and preliminary analysis. These include the introduction, analysis of response rate, data coding, analysis of missing data/values, analysis and management of outliers, analysis of common method variance and analysis of nonresponse bias. Others are, data normality analysis, analysis of multicollinearity and analysis of descriptive latent variables of the study. Therefore, figure 1 present the conceptual framework on the data screening, cleaning and preliminary analysis developed for the study.

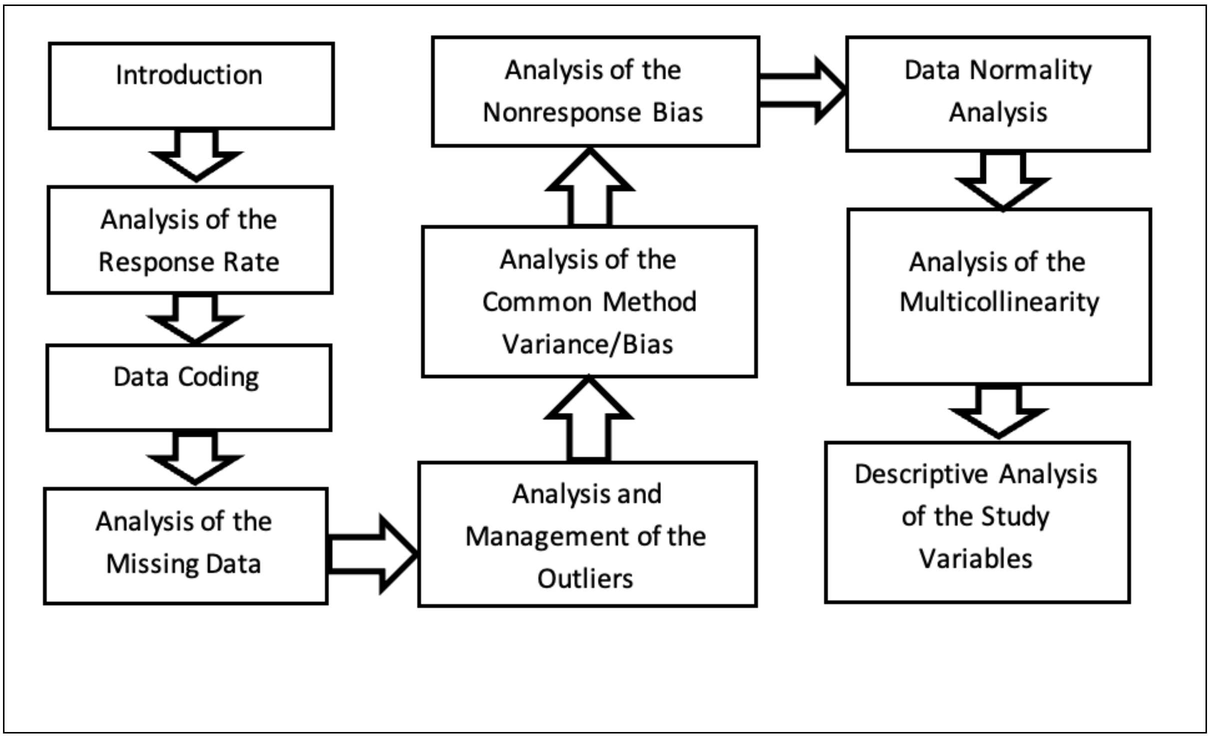

Figure 1. *Conceptual framework for data screening, cleaning and preliminary analysis.*

RESEARCH METHODS

The study utilised data for PhD study that was collected in 2019 (Pulka, 2019). The population of the study is made up of all 1,726 SMEs operating in northeastern Nigeria (SMEDAN, 2012; SMEDAN & NBS, 2013). A survey method was used to collect data from the owners/managers of the SMEs utilising the self-administered and structured questionnaire based on 5 point Likert scale. Krejcie and Morgan, (1970) sample size determination table was used to determine the sample of the study. According the table, the sample size for this study is 313. However, to succeed in managing non-response bias, 50% (157) has been added to the sample size, making 470. This is in line with the sample size adjustment by Bartlett, Kotrlik, and Higgins, (2001) and Salkind (1997). Multistage sampling technique was used in selecting the SMEs from a population of the study. First of all, the study area was clustered according to states (Adamawa, Bauchi, Borno, Gombe, Taraba and Yobe states). Secondly, proportional to size sample sampling was used to determine the number of subsample from each state. Thirdly, simple random sampling was employed to select SMEs that participated in the survey. Lastly, SPSS version 24 was used in the data analysis.

Measures

The instrument for measuring the variables in the study were adapted from previous research. SMEs performance was measured using instruments adapted from the work of Santos and Brito, (2012). The measurement of entrepreneurial competencies was adapted from the work of Man (2001). The measurement of entrepreneurial orientation was adapted from the work of Covin and Slevin (1989). The measurement of ICT was adapted from the work of Bayo-Moriones et al., (2013) and Yusuf, (2013). The measurement of entrepreneurial network was adapted from the work Naala (2016). The measurement of government business support was adapted from the work of Shamsuddin (2014). While measurement of external environment was adapted from the work of Chi (2006, 2009).

RESULTS AND DISCUSSION

Response Rate

Response rate refers to the total number of questionnaires successfully retrieved by a researcher during a research process that enable him to use the data for analysis to draw inferences. Therefore, 470 questionnaires were administered to the respondents (SMEs owners/managers) from the study area (Adamawa, Bauchi, Borno, Gombe, Taraba and Yobe states) with the help of six research assistants, one designated at each state. The process is enhanced with follow up visits at intervals, phone calls and text messages (Aminu, 2015; Sekaran & Bougie, 2010; Shamsudeen, Yeng, & Hassan, 2016; Shehu, 2014).

Consequently, 321 questionnaires were successfully retrieved from the respondents out of 470 that was initially distributed to them. Thereby achieving the response rate of 68%. However, out of the successfully retrieved questionnaires, 13 (2.7%) were invalid because the respondents wrongly filled in or left substantial parts of the questionnaires blank. Hence, 308 (65.5%) of the questionnaires were valid and used for the analysis. Thus, the response rate for the study is adequate for analysis at 65.5% since Babbie (2007), Baruch and Holtom (2008), Hair, Anderson, Babin, and Black (2010), Rubin and Babbie (2015) and Sekaran and Bougie (2013) explained that response rate of 36%, 30%, 50%, 50% and 30% respectively is sufficient and considered adequate.

Therefore, the response rate is moderate in this field. For instance, the studies of Aminu (2015) achieved the response rate of 89.46%, Gorondutse (2014) 64%, Shamsudeen (2016) 66%, Otache (2015) 32.2%, Shehu (2014) 70% and Jabeen (2014) 77%. Thus, table 1 shows the responses rate for the study.

Table 1

Response Rate of the Questionnaires

Responses	Adamawa	Bauchi	Borno	Gombe	Taraba	Yobe	Total
Number of distributed questionnaires	67	177	46	69	67	44	470
Number of retrieved questionnaires	45	119	33	48	45	31	321
Number of retrieved and valid questionnaires	43	112	33	47	42	31	308
Number of returned and invalid questionnaires	02	07	00	01	03	00	13
Number of Questionnaires not retrieved	22	58	13	21	22	13	149
Gross response rate in %	67	67	72	70	67	70	68
Valid response rate in %	64	63	72	68	63	70	65.5

Source: Field Survey

Data Coding

The responses from the valid questionnaires were coded and entered into excel worksheet. The coding of the responses was based on the 5 points Likert scale, ranging from strongly disagree = 1, disagree = 2, undecided = 3, agree = 4 and strongly agree = 5 (Vagias, 2006). This is applied to all variables in the study. After coding, the variables were assigned initials; SP = SMEs performance, EC = entrepreneurial competencies, EO = entrepreneurial orientation, ICT = information and communication technology, EN = entrepreneurial network, GBS = government business support and EE = external environment. After coding, data cleaning and preliminary analysis were performed.

Data Cleaning

Data cleaning is an important aspect of data analysis. It is carried out to detect and remove possible errors and inconsistencies from the research data with the aim of enhancing its quality (Rahm & Do, 2000). Therefore, to provide accurate and consistent research data, cleaning and preliminary analysis are essential. Similarly, Hair, Wolfinbarger, Ortinau and Bush (2010) clarified that data cleaning enables the detection of any possible violation of the key assumptions related to the use of the multivariate approach in data analysis.

Consequently, to carry out the data screening, cleaning and preliminary analysis, all the 308 usable questionnaires were coded and entered into excel worksheet and later were transferred into SPSS version 24 and performed the screening, cleaning and preliminary analysis. The screening, cleaning and analysis includes analysis of missing values, analysis of outliers, normality and multicollinearity tests (Hair, Anderson, et al., 2010; Rahm & Do, 2000).

ANALYSIS OF THE MISSING DATA

Many studies in behavioural sciences experience the cases of missing data (Acock, 2005; Pigott, 2001; Schlomer, Bauman, & Card, 2010; Streiner, 2002). Missing data usually happen when there is a failure from the respondents to answer one or more questions intentionally or unintentionally (Hair, Hult, Ringle, & Sarstedt, 2017).

Therefore, after ascertaining the missing values in the data set, it was found that out of 33, 572 data points, 28 were randomly missing. It accounts for 0.083% of the total data points. Specifically, SMEs performance has six missing values, entrepreneurial competencies five missing values, entrepreneurial orientation three missing values, ICT one missing value, entrepreneurial network three missing values, government business support three missing values and external environment seven missing values.

Furthermore, when a research data has missing values, it is suggested that it should be statistically treated. According to Schlomer et al., (2010) to treat missing values in a study, there is the need to examine the extent and nature of the missing values and the procedures for treating and managing the missing data. Thus, there is no consensus among researchers to what percentage or level of missing values are considered a problem in research (Schafer, 1999; Schlomer et al., 2010; Tabachnick & Fidell, 2013). Several researchers have stressed different levels, for example, Schafer (1999) and Hair, Ringle and Sarstedt (2013) suggested 5%, Bennett (2001) suggested 10%, and Peng, Harwell, Liou and Ehman (2006) 20% as the limit in any given research. Although the above suggestions were made, Hair et al., (2017) also recommended that any questionnaire that contains more than 15% of missing values should be removed from the data file.

Similarly, they noted that an entire observation might be removed from the data file if there are high missing values in a particular construct. As a result, Little and Rubin (1987), Raymond (1986) and Tabachnick and Fidell (2007) suggested that when the missing values in a research data are less than 5%, mean value replacement should be used to treat the missing data. Therefore, mean value replacement was used in replacing the missing values found in the data of the study (Hair, Hult, Ringle, & Sarstedt, 2017; Hair et al., 2013; Tabachnick, & Fidell, 2007). Likewise, to calculate the missing value in the data, the following formula was applied which is adapted from previous studies (Aminu, 2015; Badara, 2015).

$$\frac{\text{Number of missing values}}{\text{Total number of observations}} \times 100$$

Hence, in this study, the missing values analysis showed none of the variables is having up to 5% of the missing values. The missing values found 0.083%. Thus, this study employed mean value replacement in treating the missing values. The mean value replacement is a process been used in statistical analysis to replace missing cases present in a data set by the mean of non-missing cases of that research variable. In other words, the missing values of an indicators of a variable are replaced with the mean of valid values of that indicator in the data set (Hair et al., 2017). This is done in conformity with previous studies (Economics, Han, Wang, & Naim, 2017; Aminu, 2015; Shamsudeen et al., 2016).

Analysis and Management of Outliers

According to Hair et al., (2017) outliers are extreme responses to a particular question (s) or extreme responses to the entire items in a questionnaire. Zikmund, Babin and Griffin (2010) view outliers as observations in a research data or responses that possess unique attribute from others observations or responses. While Caroni, Karioti, Economou, Pierrakou, & Sciences, (2005), Hair, Wolfinbarger, et al., (2010), and Hodge & Austin (2004) view outliers as observations or its components in research that are inconsistent with other observations or its components. Likewise, Fidell and Tabachnick (2003) view outliers as irregular cases with unjustified effect on the results of a study that increase or sometimes decrease the mean causing concealment of the actual significance or create false significance in research. Outliers usually inflate dispersion, leading to distortions in the correlation results. Outliers in a data may affect the results of the analysis, leading to unpredictable results that cannot be generalised to the population of the study (Fidell & Tabachnick, 2003). Similarly, the existence of the outliers in a research data could interfere with the estimations and values of the coefficients and might lead to defective statistical results (Verardi, Croux, Verardi, & Croux, 2009).

Accordingly, Tabachnick and Fidell (2007) pointed out that some factors usually cause outliers in research data. These factors include inappropriate data entry into excel or SPSS worksheet, failure to identify and specify missing values in the data, outliers are not members of the study population and the variables in the population that have more extreme values in the data that are not normally distributed. Similarly, Grubbs (1969) and Rousseeuw and Hubert (2011) contended that outliers could occur as a result of gross deviation of direction from other observations in a study. Hair et al., (2017) gave the example of wrong coding (instead of 3, 33 is inserted).

Therefore, to detect outliers in a study univariate, bivariate or multivariate methods can be used (Ben-gal, 2005; Hair et al., 2017; Pallant, 2010). In detecting observations in the data set that are outside the actual value limit, the following three steps were followed. Firstly, the study detected outliers by using SPSS 24 with all the observations which might appear outside the normal value label as a result of wrong data entry. From the descriptive statistics using the frequency tables of all the research variables, minimum and maximum statistics were checked.

Secondly, the data were subjected to univariate outlier analysis by employing Z scores standardised values with a cut off of ± 3.29 ($p < .001$) as suggested by (Tabachnick & Fidell, 2007). The univariate outlier analysis takes care of the outliers among the items of a construct. The univariate outliers can be detected using standardised variable values (Z score) and or by employing frequency distribution (i.e. histograms, box plots and normal probability plots). Considering these two options, the study employed standardised variable values (Z-scores) with the threshold of ± 3.29 to detect the univariate outliers (Tabachnick & Fidell, 2007). Consequently, the results of the univariate analysis show that 34 cases (1, 3, 4, 8, 9, 11, 13, 15, 16, 17, 18, 21, 22, 23, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 43, 44, 45, 46, 48, 49, 50, 52, 54, 55,) have exceeded the thresholds of ± 3.29 and as a result, were considered as univariate outliers in the data. Hence, the 34 cases were deleted from the data, and 274 cases were considered for further analysis.

Thirdly, multivariate outlier analysis was conducted. The multivariate outlier analysis takes care of the outliers in the constructs of the study. In line with Aminu (2015), Fidell and Tabachnick (2003), Shamsudeen et al., (2016) and Tabachnick and Fidell (2007) the study used Mahalanobis D^2 measurement to detect and treat multivariate outliers. The Mahalanobis distance is "the distance of a case from the centroid of the remaining cases where the centroid is the point created at the intersection of the means of all the variables" (Tabachnick & Fidell, 2007). Linear regression analysis in SPSS software and Chi-square were employed in detecting the multivariate outliers in the study.

Therefore, the Mahalanobis analysis was conducted using SPSS, and the calculated values were compared with the critical values in the Chi-square table. In the study, 99 items were used to measure the variables. Therefore, the degree of freedom is 98 (n-1). In the Chi-square distribution table, the threshold of Chi-square is 122.108 ($p = 0.05$) (Tabachnick & Fidell, 2013). It means that any case in the data with Mahalanobis values beyond 122.108 is regarded as a multivariate outlier and ought to be removed from the data set. Consequently, all the cases have Mahalanobis values ranging from 0.45112 to 28.1868. Therefore, there is no any case that violates the assumption in the data and hence, all 274 cases were considered for further analysis in the study.

Common Method Variance (Bias) Analysis

According to Podsakoff, MacKenzie, Lee and Podsakoff, (2003) Common method bias (CMV) refers to the differences that are more likely produced by method a researcher used in measuring research variables in a study. Equally, CMV is the extent to which false correlation occurs between research variables that are generated employing the same source of the survey to measure variables in a study (Craighead, Ketchen, Dunn, & Hult, 2011). In a survey type of research that uses a self-administered questionnaire, same source of data and the data were collected from the respondents during the same period, CMV may distort the data collected from the respondents (Lindell & Whitney, 2001; Richardson, Simmering, & Sturman, 2009; Samson, 2015; Shamsuddin et al., 2016). This could lead to misinterpretation of the results or drawing invalid inferences from a particular study (Conway & Lance, 2010). Usually, the CMV manifest by deflating or inflating results of studies (Chang, van Witteloostuijn, & Eden, 2010; Podsakoff et al., 2003; Sharma, Yetto, & Crawford, 2009; Siemsen, Roth, & Oliveira, 2010).

Thus, to ascertain that no CMV in the observed scores of the study and to ascertain that the correlations are not deflated or inflated, the study used some techniques explained by some researchers (Podsakoff et al., 2003; Podsakoff, MacKenzie, & Podsakoff, 2012). These techniques are; improving the items through rewording, using simple, concise and precise words and statements, confidentiality of the respondents is assured, the respondents were enlightened that there is no wrong or right answers provided in the questionnaire and that the questionnaires are anonymous.

Furthermore, the entire research variables for the study were subjected to Harman's single factor test (Podsakoff & Organ, 1986). CMV exist when the analysis of the factor provide only a single factor or factor manifest as a greater part of the covariance among the measurement (Podsakoff et al., 2003). It is expounded that if single factor accounts for more than 50% of the variance in the predictors and criterion variables, it indicates that CMV exist, but when it is less than 50%, CMV does not exist (Jakobsen & Jensen, 2015; Lowry & Gaskin, 2014; Podsakoff et al., 2003).

Consequently, un-rotated factor analysis with 99 items of all the variables was used in the study were analysed. The results have shown that there is no single factor accounted for up to 50% of the variance. The results generated 20 factors explaining a cumulative of 69.56% of the variance. The first factor accounted for 27.39% of the entire variance, and this is far below the threshold of 50% as suggested by some researchers (Lowry & Gaskin, 2014; Podsakoff et al., 2003). This show the non-appearance of common method variance in the study as advocated by (Jakobsen & Jensen, 2015; Lowry & Gaskin, 2014; Podsakoff et al., 2003; Samson, 2015). Therefore, the data were subjected to further statistical analysis. Therefore, table 2 present the results of the CMV.

Table 2

Common Method Variance Test

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	27.115	27.389	27.389	27.115	27.389	27.389
2	5.987	6.048	33.437	5.987	6.048	33.437
3	5.537	5.593	39.030	5.537	5.593	39.030
4	4.833	4.882	43.912	4.833	4.882	43.912
5	2.803	2.832	46.744	2.803	2.832	46.744
6	2.430	2.455	49.199	2.430	2.455	49.199
7	2.169	2.191	51.390	2.169	2.191	51.390
8	1.930	1.949	53.339	1.930	1.949	53.339
9	1.743	1.760	55.099	1.743	1.760	55.099
10	1.726	1.744	56.843	1.726	1.744	56.843
11	1.505	1.520	58.363	1.505	1.520	58.363
12	1.479	1.494	59.857	1.479	1.494	59.857
13	1.432	1.447	61.304	1.432	1.447	61.304
14	1.327	1.341	62.645	1.327	1.341	62.645
15	1.288	1.301	63.946	1.288	1.301	63.946
16	1.225	1.238	65.184	1.225	1.238	65.184
17	1.178	1.190	66.373	1.178	1.190	66.373
18	1.110	1.121	67.495	1.110	1.121	67.495
19	1.042	1.052	68.547	1.042	1.052	68.547
20	1.002	1.012	69.559	1.002	1.012	69.559

Extraction Method: Principal Component Analysis.

Non-response Bias Analysis

According to Groves (2006) non-response bias refers to the difference in the responses of the respondents and non-respondents mean in research. McInnis (2006) view non-response bias as the prejudice that occurs in the results of survey research at a point when the respondents differ from those that did not respond to certain variables. The problem of non-response bias takes place in a study where answers provided by the respondents significantly be at variance from those that did not respond (Aminu, 2015; Baruch, 1999; Baruch & Holtom, 2008). Therefore, it is essential to manage non-response bias, as a certain level may affect the integrity of research results, which may affect the generalizability of the results on the population (Groves, 2006).

Therefore, non-response bias can be tested using the significant difference between the early and lately returned questionnaires by using t-tests (Armstrong & Overton, 1977; Chiang, Kocabasoglu-Hillmer, & Suresh, 2012). Therefore, this study determined non-response bias by comparing the early and late responses of the respondents (Chiang et al., 2012; Rodrigues & Carlos Pinho, 2012). Similarly, the early response was considered based on the time frame the questionnaires are returned to the researcher (Vink & Boomsma, 2008). The early responses are the questionnaire that was received on 1st June to 5th July 2018. While the late responses are the questionnaires received after 5th July 2018. Consequently, 206 owners/managers responded within the early response period. While 68 owners/managers responded within the late response period. Thus, the independent samples t-test and Levene's test for equality of variance at 0.05 significance level was employed in testing the non-response bias (Coakes, 2013; Field, 2009; Pallant, 2010a). Table 3 depicts the t-test results of non-response bias for this study.

Table 3

Test of Non-response Bias: Independent-Samples T-Test (274)

Variables	Group	N	Mean	Std. Dev.	Levene's Test for Equality of Variances	
					F	Sig.
SMEs Performance	Early Response	206	3.76	0.57	0.013	0.909
	Late Response	68	3.65	0.59		
Entrepreneurial Competencies	Early Response	206	3.72	0.66	0.503	0.479
	Late Response	68	3.88	0.6		
Entrepreneurial Orientation	Early Response	206	4.08	0.69	2.628	0.106
	Late Response	68	4.04	0.51		
Information and Communication Technology	Early Response	206	4.07	0.66	0.010	0.922
	Late Response	68	4.16	0.60		
Entrepreneurial Network	Early Response	206	4.15	0.69	2.283	0.132
	Late Response	68	4.26	0.56		
Government Business Support	Early Response	206	3.99	0.79	0.468	0.495
	Late Response	68	3.82	0.77		
External Environment	Early Response	206	3.95	0.75	3.025	0.083
	Late Response	68	4.01	0.46		

Source: Field Survey

The results from table 3 indicated that the values of the equal variance significance for each of the seven variables of the study are higher than 0.05 significance level of Levene's test for equality of variances (Coakes, 2013; Field, 2009; Pallant, 2010a). Hence the results indicated that the equal variance no statistical differences between the early and late respondents. Additionally, the response rate has exceeded the threshold of 50% as suggested by some researchers (Lindner & Wingenbach, 2002). Therefore, there is no problem of non-response bias in the study. This signifies that the sample used in this study adequately represent the entire population and the results can be generalised.

Data Normality Analysis

For the main analysis of the PhD work, PLS-SEM 3.0 was used to analyse the data collected from the respondents in the study area. PLS-SEM does not require research data to be normally distributed as a criterion for data analysis (Hair et al., 2017). However, it is essential to check and verify research data for extreme non-normal that could pose serious problems in the evaluation of parametric data (Hair et al., 2017). Extremely non-normal data inflate the standard errors obtained from the bootstrapping process and could distort the significant level of relationships among variables (Hair et al., 2017, 2013; Vinzi, Chin, Henseler, & Wang, 2010). Hence, to check for the data normality in the study, multivariate normality is used to check the data distribution by employing kurtosis and skewness measure of the distribution as suggested by some researchers (Curran, West, & Finch, 1996; Fidell & Tabachnick, 2003; Hair, Anderson, et al., 2010; Hair et al., 2017; West, Finch, & Curran, 1995).

Similarly, according to Hair et al., (2017), if the distribution of the responses stretches to the left or right tail direction, the distribution is characterised by skewness. Moreover, Kurtosis is measured to determine whether the distribution of the responses is characterised by very thin distribution with almost all the responses are at the centre or not. Therefore, Curran et al., (1996) and West et al., (1995) argue that the values of the skewness ought to be less than 2, while the values of the Kurtosis ought to be less than 7. Moreover, Kline (2015) explained that if the values of the Skewness is above 3, and the Kurtosis is above 10, then it is an indication that the problem of non-normal distribution exists in the data. Similarly, they further explained that if the values of the Kurtosis is higher than 20, it is an indication that a severe problem exists in the data. Consequently, table 4 shows that the Skewness and Kurtosis of the metric variables of the study are within the accepted limits of less than 2 and 7 respectively. Therefore, it indicated that the data is normally distributed.

Similarly, Field (2009) the study employed histogram in examining the normality assumptions of the data. This is done following the suggestion made by Field (2009) who suggested the use of normality graphs to complement the statistical method. Centred on the histogram, Figure 2 illustrates that the data collected from the respondents. The distribution of the data of the study follows the normal pattern since all the bars on the histogram are close to the normal curve (Tabachnick & Fidell, 2007).

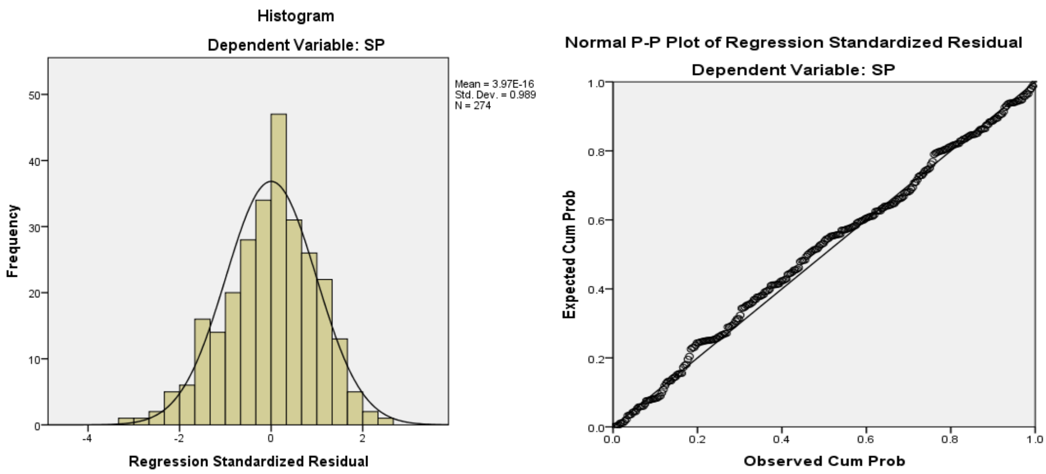

Figure 2. *Histogram Probability Plots*

Multicollinearity Test

Multicollinearity means the linear relationships that exist between two or more research variables (Alin, 2010; Belsley, 1991; Chatterjee & Hadi, 2006). Whereas to Hair, Anderson, et al., (2010) contend that multicollinearity is the relationships between two, three or more exogenous research variables, that the independent variables in research demonstrate significant correlation with the other independent variables in the same study. Multicollinearity exists when two or more variables are highly correlated in a study. Multicollinearity turns out to be a problem in a situation where two or more independent variables in a study are highly correlated to each other (Fidell & Tabachnick, 2003; Hair, Anderson, et al., 2010; Pallant, 2010b). Therefore, when two or more research variables are highly correlated in a study, it indicates that the variables have unwanted information and therefore, must be treated accordingly by dropping the section or item (s) that are highly correlated (Aminu, 2015). Hence, according to Hair, Anderson, et al., (2010) and Pallant (2010b) multicollinearity can be detected with the use of tolerance and VIF (variance inflation factor). The value of the VIF has to be between 0.1 and less than 10.

Given the above, multicollinearity was examined by applying correlation matrix, tolerance and level of VIF for the independent variables in the study. The correlation matrix was used on the independent variables to examine the level of the correlations between the research variables. According to Hair, Anderson, et al., (2010) and Pallant (2010a) when the value of the correlation matrix is 0.9 and above, it indicate the presence of multicollinearity.

Consequently, the results of the correlation matrix indicated that no any variables in the study are extremely correlated with other variables. The results from table 4 show that the values are less than 0.9. As a result, it is established that this study has no multicollinearity problem.

Table 4

Correlations Matrix of the Research Variables

Variables	EC	EO	ICT	EN	GBS	EE
Entrepreneurial Competencies	1					
Entrepreneurial Orientation	0.635	1				
Info. & Comm. Tech.	0.547	0.777	1			
Entrepreneurial Network	0.476	0.737	0.754	1		
Govt. Business Support	0.395	0.622	0.538	0.730	1	
External Environment	0.270	0.588	0.492	0.433	0.369	1

Normality Test: Skewness and Kurtosis Statistics (n=274)

Variables	Mean Stat.	Std. Dev.	Skewness Stat.	Skewness Std. Error	Kurtosis Stat.	Kurtosis Std. Error
SP	3.73	0.58	-0.64	0.15	-0.07	0.29
EC	3.76	0.65	-2.28	0.15	6.39	0.29
EO	4.07	0.65	-2.21	0.15	5.80	0.29
ICT	4.09	0.65	-1.96	0.15	5.09	0.29
EN	4.18	0.66	-2.18	0.15	5.41	0.29
GBS	3.95	0.79	-1.66	0.15	3.06	0.29
EE	3.97	0.68	-1.88	0.15	4.79	0.29

Multicollinearity Test based on Tolerance and VIF Values

Construct	**Tolerance**	**VIF**
Entrepreneurial Competencies	0.570	1.754
Entrepreneurial Orientation	0.238	4.203
Information and Communication technology	0.311	3.214
Entrepreneurial Network	0.273	3.665
Government Business Support	0.443	2.255
External Environment	0.630	1.587

Descriptive Statistics of the Latent Variables

Variables	Mean Statistics	Std. Deviation Statistics
SMEs Performance	3.732	0.578
Entrepreneurial Competencies	3.761	0.652
Entrepreneurial Orientation	4.067	0.650
Information and Communication Technology	4.092	0.646
Entrepreneurial Network	4.176	0.663
Government Business Support	3.950	0.788
External Environment	3.967	0.684

Source: Field Survey

The second method of detecting multicollinearity is by tolerance and variance inflation factor (Hair et al., 2017; Hair et al., 2011; Hair, William, Babin, & Anderson, 2014; Peng & Lai, 2012). VIF is defined as the reciprocal of the tolerance level among variables (Hair et al., 2017). Therefore, the tolerance level of 0.20 and higher, while VIF values of below 5 indicate the absence of multicollinearity (Hair et al., 2017; Sarstedt, Ringle, Smith, Reams, & Hair, 2014). Table 4 present the tolerance and VIF values of the independent variables.

The results from table 4 indicated the absence of multicollinearity. The levels of tolerance of all the independent variables are higher than 0.20, and the values of VIF are lower than 5 for all the variables in the study. In summary, the results of the correlation matrix, tolerance and VIF revealed that all the exogenous latent variables in the study show no multicollinearity problem. As a result, multicollinearity is not a problem in the study.

Descriptive Statistics of the Latent Variables

According to Pallant (2010b, 2011), descriptive statistics analysis is essential and has several advantages which include: describing the characteristics of the sample as it might exist at the time of the research, checking the research variables for any violation of the assumptions underlying a statistical techniques used in research and is also used to address specific research objectives and questions. Therefore, this section presents the descriptive statistics of the latent constructs used in the study. The results show the mean and the standard deviation of the computed constructs to determine the descriptive characteristics of the study.

As discussed earlier, the study employed 5 points Likert scale ranging from strongly disagree to strongly agree. Therefore, the descriptive results with the mean value of less than 2.34 are considered as low, from the value of 2.34 to 3.66 are considered as moderate, while mean value of 3.67 and beyond are considered as high (Nunnally & Bernstein, 1994).

From table 4, the results indicated that the mean score of SMEs performance is 3.7320 and has a standard deviation of 0.57852. The mean score of entrepreneurial competencies is 3.7614, and the standard deviation is 0.65289. Entrepreneurial orientation has a mean score of 4.0675 and standard deviation of 0.65019. Similarly, information and communication technology has a mean score of 4.0922 and standard deviation of 0.64644. Furthermore, the entrepreneurial network has a mean score of 4.1768 and standard deviation of 0.66304. Government business support has a mean score of 3.9506 and standard deviation of 0.78878.

Last but not least, the external environment has a mean score of 3.9672 and standard deviation of 0.68475. Going by the criteria of Muhammad and Taib (2010) as employed by Naala (2016), the mean scores of all the variables in the study are considered as high, since all the mean scores of the variables have exceeded 3.6.

CONCLUSION AND RECOMMENDATIONS

Despite the importance of conducting data screening, cleaning and preliminary analysis, it was found that many studies don not perform and report this stage of statistical analysis in research works. Therefore, this study provided detailed steps and procedure for performing data screening, cleaning and preliminary analysis. This is achieved by developing a framework and empirically testing and treating a data collected from SMEs. Therefore, it is concluded that the data has satisfied the assumptions and requirement for conducting multivariate data analysis. Thus, it is strongly recommended that researchers in entrepreneurship, social sciences, management and other related disciplines should follows the procedure for data screening, cleaning and preliminary analysis to enhance the reliability and accuracy of research results and findings the their studies.

ACKNOWLEDGMENT

Single Grant
This work was supported by University of Maiduguri, Nigeria (through the Federal Government of Nigeria Needs Assessment Grant) for 2016.

Services and Facilities Involved
The author gratefully acknowledge the use of service and facilities of the Sultanah Bahiyah Library and facilities of School of Business Management, Universiti Utara Malaysia from February, 2016 to June, 2019.

REFERENCES

Abduwahab, L., Dahalin, Z., & M.B., G. (2011). Data screening and Preliminary Analysis of the Determinants of User Acceptance of Telecentre. Journal of Information Systems: New Paradigms, 1(1), 11–23.

Acock, A. C. (2005). Working with Missing Values. Journal of Mariage and Family, 67(4), 1012–1028.

Alin, A. (2010). Multicollinearity. Wiley Interdisciplinary Reviews: Computational Statistics, 2(3), 370–374. https://doi.org/10.1002/wics.84

Aminu, I. M. (2015a). Mediating role of access to finance and moderating role of business environment on the relationship between strategic orientation attributes and performance of small and medium enterprises in Nigeria (Doctoral dissertation, Universiti Utara Malaysia). PhD Thesis. Universiti Utara Malaysia.

Aminu, I. M. (2015b). Meditiating role of access to finance moderating role of business environment relationship between strategic orientation attributes and performance of SMEs in Nigeria, 17(3).

Armstrong, J. S., & Overton, T. S. (1977). Estimating Nonresponse Bias in Mail Surveys The Wharton School , University of Pennsylvania. Journal of Marketing, 14(3), 396–402. https://doi.org/10.2307/3150783

Babbie, E. (2007). The practice of social researach (11th ed.). California: Wadsworth: Belmont.

Badara, A. K. M. (2015). Leadership Succession , Organizational Climate , Trust and Individual Performance in Nigerian Commercial Banks (Doctoral dissertation, Universiti Utara Malaysia).

Barroso, C., Carri´on, G. C., & Rold´an, J. L. (2010). Applying Maximum Likelihood and PLS on Different Sample Sizes: Studies on SERVQUAL Model and Employee Behavior Model. In Springer Berlin Heidelberg. (p. 627). https://doi.org/10.1007/978-3-642-16345-6

Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational Research: Determining Appropriate Sample

Size in Survey Research. Information Technology, Learning, and Performance Journal, 19(1), 43–50. https://doi.org/10.1109/LPT.2009.2020494

Baruch, Y. (1999). Response Rate in Academic Studies-A Comparative Analysis. Human Relations, 52(4), 421–438. https://doi.org/10.1177/001872679905200401

Baruch, Y., & Holtom, B. C. (2008). Survey response rate levels and trends in organizational research. Human Relations, 61(8), 1139–1160. https://doi.org/10.1177/0018726708094863

Bayo-Moriones, A., Billon, M., & Lera-Lopez, F. (2013). Perceived performance effects of ICT in manufacturing SMEs. Industrial Management & Data Systems, 113(1), 117–135. https://doi.org/http://dx.doi.org/10.1108/02635571311289700

Belsley, D. (1991). Conditioning diagnostics collinearity and weak data in regression. New York: John Wiley & Sons.

Ben-gal, I. (2005). Outlier Detection. In Data Mining and Knowledge Discovery Handbook (pp. 131–146). https://doi.org/10.1007/0-387-25465-x_7

Bennett, D. A. (2001). How can I deal with missing data in my study? Australian and New Zealand Journal of Public Health, 25(5), 464–469. https://doi.org/10.1111/j.1467-842X.2001.tb00294.x

Blischke, W. R., Karim, M. R., & Murthy, D. N. P. (2011). Warranty data collection and analysis. Springer Series in Reliability Engineering. https://doi.org/10.1007/978-0-85729-647-4

Caroni, C., Karioti, V., Economou, P., Pierrakou, C., & Sciences, P. (2005). The Analysis of Outliers in Statistical Data. Thales Project, (xxxx).

Chang, S.-J., van Witteloostuijn, A., & Eden, L. (2010). From the Editors: Common method variance in international business research. Journal of International Business Studies, 41(2), 178–184. https://doi.org/10.1057/jibs.2009.88

Chatterjee, S., & Hadi, A. . (2006). Regression Analysis byExample. 4th ed. New York: John Wiley&Sons.

Chi, T. (2006). A Study of the Relationships between Business Environment Characteristics, Competitive Priorities, Supply Chain Structures, and Firm Performance in the U.S. Technical Textile Industry. A Dissertation Submitted to the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Chiang, C.-Y., Kocabasoglu-Hillmer, C., & Suresh, N. (2012). An empirical investigation of the impact of strategic sourcing and flexibility on firm's supply chain agility. International Journal of Operations & Production Management, 32(1), 49–78.

Coakes, S. J. (2013). SPSS version 20.0 for windows: Analysis without anguish. Australia: Wiley.

Conway, J. M., & Lance, C. E. (2010). What reviewers should expect from authors regarding common method bias in organizational research. Journal of Business and Psychology, 25(3), 325–334. https://doi.org/10.1007/s10869-010-9181-6

Covin, J. G., & Slevin, D. P. (1991). A conceptual model of entrepreneurship as firm behavior. Entrepreneurship: Critical Perspectives on Business and Management, 3, 5-28.

Covin, Jeffrey G, & Slevin, D. P. (1989). Strategic Management of Small Firms in Hostile and Benign Environments. Strategic Management Journal, 10(1), 75–87.

Craighead, C. W., Ketchen, D. J., Dunn, K. S., & Hult, G. T. M. (2011). Addressing Common Method Variance : Guidelines for Survey Research on Information Technology , Operations , and Supply Chain Management. 578 IEEE TRANSACTIONS ON ENGINEERINGMANAGEMENT, 58(3), 578–588.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. Psychological Methods, 1(1), 16–29. https://doi.org/10.1037/1082-

989X.1.1.16

Economics, I. J. P., Han, J. H., Wang, Y., & Naim, M. (2017). Reconceptualization of information technology fl exibility for supply chain management : An empirical study. Intern. Journal of Production Economics, 187(February), 196–215. https://doi.org/10.1016/j.ijpe.2017.02.018

Fidell, L. S., & Tabachnick, B. G. (2003). Preparatory data analysis. Handbook of Psychology: Volume 2 Research Methods in Psychology. https://doi.org/10.1002/0471264385.wei0205

Field, A. (2009). Discovering Statistics using SPSS (3rd ed.). London: Sage Publication.

Gorondutse, A. H. (2014). Effect of business social responsibity (BSR) on Performance of SMES in Nigeria (Doctoral dissertation, Universiti Utara Malaysia).

Groves, R. (2006). Nonresponse rates and nonresponse bias in household surveys: What Do We Know about the Linkage between Nonresponse Rates and Nonresponse Bias? Public Opinion Quarterly, 70(5), 646–675. https://doi.org/10.1093/poq/nfl033

Grubbs, F. E. (1969). procedures for detecting outlying observations in samples. Technometrics, 11(1), 1–21. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/781499.pdf

Hair, J. F., Anderson, R. E., Babin, B. J., & Black, W. C. (2010). Multivariate data analysis: A global perspective (Vol. 7). Upper Saddle River, NJ: Pearson.

Hair, J. F., Wolfinbarger, M. F., Ortinau, D. J., & Bush, R. P. (2010). Essentials of Marketing Research: New York: McGraw-Hill/Irwin.

Hair, J.F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). A Primer on Partial Least Squares Structural Equation Modelling (PLS-SEM). 2nd Edition, SAGE Publishers.

Hair, Joe F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. The Journal of Marketing Theory and Practice, 19(2), 139–152. https://doi.org/10.2753/MTP1069-6679190202

Hair, Joseph F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). SAGE Publications, Inc (Second Edi). Melbourne.

Hair, Joseph F., Ringle, C. M., & Sarstedt, M. (2013). Partial Least Squares Structural Equation Modeling: Rigorous Applications, Better Results and Higher Acceptance. Long Range Planning, 46(1–2), 1–12. https://doi.org/10.1016/j.lrp.2013.01.001

Hair Jr, J. F., William, C., Babin, B. J., & Anderson, R. E. (2014). Multivariate Data Analysis Joseph F . Hair Jr . William C . Black Seventh Edition. Pearson Education Limited.

Hodge, V. J., & Austin, J. (2004). A Survey of Outlier Detection Methodoligies. Artificial Intelligence Review, 22(1969), 85–126. https://doi.org/10.1007/s10462-004-4304-y

Ibrahim, M. A., & Shariff, M. N. M. (2014). Strategic Orientation , Access to Finance , Business Environment and SMEs Performance in Nigeria : Data Screening and Preliminary Analysis Strategic Orientation , Access to Finance , Business Environment and SMEs Performance in Nigeria : Data Screening an. European Journal of Business and Management, 6(35).

Jabeen, R. (2014). Moderating Effect of External Environment on Performance of SMEs in Pakistan. Universiti Utara Malaysia.

Jakobsen, M., & Jensen, R. (2015). Common method bias in public management studies. International Public Management Journal, 18(1), 3–30. https://doi.org/10.1080/10967494.2014.997906

Kaur, H., & Bains, A. (2013). Understanding The Concept Of Entrepreneur Competency. Journal of Business Management & Social Sciences Research (JBM&SSR), 2(11), 31–33.

Kline, R. B. (2015). principles and practice of structural equation modelling. guilford publications. fourth edition.

Krejcie, R. V, & Morgan, D. W. (1970). Determining Sample Size for Research Activities Robert. Educational and Psychological Measurement, 38(1), 607–610. https://doi.org/10.1177/001316447003000308

Kura, K. M. (2014). Organisational Formal Controls , Group Norms And Workplace Deviance : The Moderating Role Of Self-Regulatory Efficacy. Doctor Of Philosophy Universiti Utara Malaysia.

Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. Journal of Applied Psychology, 86(1), 114–121. https://doi.org/10.1037//0021-9010.86.1.114

Lindner, James R. & Wingenbach, G. J. (2002). Communicating the Handling of Nonresponse Error in Journal of Extension Research in Brief Articles. Journal of Extension, 40(6), 1–5. Retrieved from http://www.joe.org/joe/2002december/rb1.php

Little, R. J. A., & Rubin, D. B. (1987). StatisticalAnalysis with Missing Data. New York: John Wiley & Sons, Inc.

Lowry, P. B., & Gaskin, J. (2014). Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. IEEE Transactions on Professional Communication, 57(2), 123–146. https://doi.org/10.1109/TPC.2014.2312452

Man, T. W. Y. (2001). Entrepreneurial Competencies and the Performance of Small and Medium Enterprises in the Hong Kong Services Sector.

McInnis, E. D. (2006). Nonresponse bias in student assessment surveys: a comparison of respondents and non-respondents of the national survey of student engagement at an independent comprehensive Catholic University.

Muhammad, N. M. N., & Taib, M. J. and F. M. (2010). Moderating Effect of Information Processing Capacity to Investment Decision Making and Environmental Scanning. BMQR, 1(1).

Naala, M. N. I. (2016). Moderating and Mediating Roles of Human Capital and Competitive Advantage on Entrepreneurial Orientation, Social Network, and Performance of SMEs in Nigeria. A hD Thesis Submitted to Universiti Utara Malysia.

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric Theory, (3rd edn), Mcgraw-Hill: New York.

Otache, I., & Mahmood, R. (2015). Corporate Entrepreneurship and Business Performance: The Role of External Environment and Organizational Culture: A Proposed Framework. Mediterranean Journal of Social Sciences, 6(4). https://doi.org/10.5901/mjss.2015.v6n4s3p524

Pallant, J. (2010a). SPSS survival manual, 4th edtion. England: McGraw-Hill Education. London.

Pallant, J. (2010b). SPSS survival manual: A step by step guide to data analysis using.

Pallant, J. (2011). SPSS survival manual: A step by step guide to data analysis using SPSS (4th ed.). New York: Open University Press.

Peng, C. Y. J., Harwell, M., Liou, S. M., & Ehman, L. H. (2006). Advances in missing data methods and implications for educational research. In S. Sawilowsky (Ed.), Real data analysis (pp. 31–78). Greenwich, CT: Information Age.

Peng, D. X., & Lai, F. (2012). Using partial least squares in operations management research: A practical guideline and summary of past research. Journal of Operations Management, 30(6), 467–480. https://doi.org/10.1016/j.jom.2012.06.002

Pigott, T. D. (2001). A Review of Methods for Missing Data. Educational Research and Evaluation, 7(4), 353–383. https://doi.org/10.1076/edre.7.4.353.8937

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), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of Method Bias in Social Science Research and Recommendations on How to Control It. Annual Review of Psychology, 63(1), 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Podsakoff, P. M., & Organ, D. W. (1986). Self-Reports in Organizational Research: Problems and Prospects. Journal of Management, 12(4), 531–544. https://doi.org/10.1177/014920638601200408

Pulka, B. M. (2019). The Determinants of SMEs Performance in Nigeria; The Moderating Role of External Environment. PhD Thesis Submittted to Universiti Utara Malaysia.

Rahm, E., & Do, H. H. (2000). Data Cleaning: Problems and Current Approaches. Bulletin of the Technical Committee on Data Engineering, 23(4), 3–13. https://doi.org/10.1145/1317331.1317341

Raymond, M. R. (1986). Missing data in evaluation research. Evaluation & the Health Professions, 9, 395–420.

Richardson, H. A., Simmering, M. J., & Sturman, M. C. (2009). A Tale of Three Perspectives : Examining Post Hoc Statistical Techniques for Detection and Corrections of Common Method Variance. M Cornell University, School of Hospitality Administration. https://doi.org/10.1177/1094428109332834.

Rodrigues, A. P., & Carlos Pinho, J. (2012). The impact of internal and external market orientation on performance in local public organisations. Marketing Intelligence & Planning, 30(3), 284–306. https://doi.org/10.1108/02634501211226276

Rousseeuw, P. J., & Hubert, M. (2011). Robust statistics for outlier detection. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1(1), 73–79. https://doi.org/10.1002/widm.2

Rubin, A., & Babbie, E. R. (2015). Research Methods for Social Work.

Salkind, N. J. (1997). Exploring research (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Samson, A. T. (2015). Strategic Orientations , Reconfiguring Capability , Environmental Turbulence and Export Performance of Smes in Nigeria Doctor of Philosophy Universiti Utara Malaysia. Universiti Utara Malaysia.

Santos, J. B., & Brito, L. A. L. (2012). Toward a subjective measurement model for firm performance. BAR - Brazilian Administration Review, 9(SPL. ISS), 95–117. https://doi.org/10.1590/S1807-76922012000500007

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. https://doi.org/10.1016/j.jfbs.2014.01.002

Schafer, J. L. (1999). Multiple imputation: A primer. Statistical Methods in Medical Research, 8, 3–15.

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. Journal of Counseling Psychology, 57(1), 1–10. https://doi.org/10.1037/a0018082

Sekaran, U., & Bougie, R. (2010). Research methods for business. A skill building approach (5th ed.). UK: John Willey.

Sekaran, U., & Bougie, R. (2013). 6th edition. Research Methods for Business.

Shamsuddin, K., Yeng, K., & Hassan, H. (2016). The Mediatory Role of Access to Finance between Finance Awareness and SMEs Performance in Nigeria. International Business Management, 10(18), 4304–4310.

Shamsudeen, K., Keat, O. Y., & Hassan, H. (2016). Assessing the Impact of Viable Business Plan on the Performance of Nigerian SMEs : A Study among Some Selected SMEs Operators in, 1, 18–25.

Sharma, R., Yetto, P., & Crawford, J. (2009). Estimating the Effect of Common Method Variance: The Method-Method Pair Technique with an Illustration from A Research. MIS Quarterly, 33(3), 473–490.

Shehu, A. M. (2014). Market orientation, knowledge management, entrepreneurial orientation and performance of Nigeria SMEs (Doctoral dissertation, Universiti Utara Malaysia).

Siemsen, E., Roth, A., & Oliveira, P. (2010). Common Method Bias in Regression Models With Linear, Quadratic, and Interaction Effects. Organizational Research Methods, 13(3), 456–476. https://doi.org/10.1177/1094428109351241

SMEDAN. (2012). Smedan 2012 annual report. Retrieved from http://www.smedan.gov.ng/images/SMEDAN 2012 ANNUAL REPORT.pdf

SMEDAN, & NBS. (2013). Smedan and National Bureau of Statistics Collaborative Survey : Selected Findings.

Streiner, D. L. (2002). The case of the missing data: Methods of dealing with dropouts and other research vagaries. Canadian Journal of Psychiatry, 47(1), 68–75.

Tabachnick, B. G., & Fidell, L. S. (2007). Using multivariate statistics (5th ed.). Boston, MA: Allyn & Bacon/Pearson Education.

Tabachnick, B. G., & Fidell, L. S. (2013). Using multivariate statistics (6th ed.). New Jersey: Pearson Education Inc.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y. M., & Lauro, C. (2005). PLS path modeling. Computational Statistics and Data Analysis, 48(1), 159–205. https://doi.org/10.1016/j.csda.2004.03.005

Vagias, W. (2006). Likert-type scale response anchors. Clemson International Institute for Tourism and Research Development, Department of Parks, Recreation and Tourism Managment., 3–4. https://doi.org/10.1525/auk.2008.125.1.225

Verardi, V., Croux, C., Verardi, V., & Croux, C. (2009). Robust Regression in Stata. The Strata Journal, 9(3), 439–453.

Vink, J. M., & Boomsma, D. I. (2008). A Comparison of Early and Late Respondents in a Twin – Family Survey Study. Twin Research and Human Genetics, 11(2), 165–173.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 56-75). Thousand Oaks, CA: Sage Publications.

Yusuf, A. A. (2013). Impact of ict on smes — Case Rwanda. Turku University Of Applied Sciences.

Zakaria, N. (2016). Entrepreneurial opportunity recognition among Bumiputera SMEs entrepreneurs in Malaysia: The influential factors of social network, entrepreneurial alertness and creativity (Doctoral dissertation, Universiti Utara Malaysia).

Zikmund, W. G., Babin, B. J., & Griffin, M. (2010). Business Research Methods. Mason, Ohio, South-Western.