

## CHAPTER 4

### DATA ANALYSIS

#### 4.1 Introduction

Data analysis is a critical step in research that involves organising, cleaning, transforming, examining, and interpreting data to extract meaningful insights and draw conclusions. There are various methods available for data analysis, including descriptive statistics, inferential statistics, qualitative analysis, and mixed-methods analysis. This chapter highlights and presents the analysis of the collected data in this study. The selected statistical tests, which were noted in the previous chapter, were performed. All results are presented in tabular forms accompanied by the interpretation of results for better understanding.

#### 4.2 Data Preparation

In this section, the researcher presented selected instruments to measure the variables contributing to the study's conceptual model. A distinction was made between independent and dependent variables. Additionally, appropriate measurement scales were chosen based on the study's context.

#### 4.3 Data Collection Method

There are various survey methods available, which include interview (either face-to-face or telephone), questionnaire survey (e.g., administered personally, by mail, email, or online), and panel discussion group methods. Babbie (1989) and Miller (1991)

discussed the relative advantages and disadvantages of these methods in social sciences, especially mail survey and personal and telephone interview methods. In general, personal interview method is superior to the other methods, but this method is the most expensive and can be rather challenging to secure information (Miller, 1991; Kerlinger & Lee, 2000). On the contrary, the mail survey method has considerable advantages in terms of cost and ease of getting information (Miller, 1991). Arguably, each of these data collection methods has its advantages and disadvantages. However, evidence shows that surveys are more often carried out by mail (Dillman, 1991; Bartholomew & Smith, 2006). As a general guide, it is recommended that researchers weigh up the advantages and disadvantages of the available methods in relation to their research context and resources (Babbie, 1989; Miller, 1991; Sandford, 1995; Evans et al., 1996).

In the context of small businesses, Forsgren (1989) recommended using altruistic appeal in the cover letter considering the proper design of the questionnaire survey, and ensuring anonymity for a sensitive study as further measures can increase the response rate. In another recent development, Dennis (2003) argued that not all measures can improve the response rate and recognised only four specific measures that improve the response rate, namely follow-ups and reminders, monetary incentives, university sponsorship, and stamped return envelopes.

This study's questionnaire survey was administered among several groups via Google Forms. A total of 450 questionnaire sets were distributed among randomly selected respondents through personal visits and emails. A total of 384 useable questionnaire sets were returned, which represented an overall response rate of 75% of the net sampling frame. The survey was conducted from September 2021 to February 2022. After further screening, 384 questionnaire sets were deemed valid for analysis.

The demographic characteristics of respondents included the type of business activity,

current position, the average number of staff, accounting records, and registration for VAT. Details of the study's constructs, dimensions, and indicators are given in Table 4.1.

**Table 4.1:** Constructs, Dimensions, and Indicators

Variables	Indicators
Compliance cost	COMC
Tax rate	TR
Probability of detection	POF
Complexity	CMP
Penalty imposition	PEN
Tax incentive	TINC
VAT compliance behaviour	VATCB

#### 4.4 Demographic Profile of Respondents

Descriptive analysis was conducted to obtain the frequencies and percentages of the respondents' demographic characteristics. Many respondents who were associated with SMEs in the UAE responded to the survey, reflecting their interest and voluntary participation in this study. The response rate of 75% was deemed ideal and consistent with the argumentation (Bista, 2018) for a survey-based study. Referring to Table 4.2, in regards to the types of business activities, 27.3% (N = 105) of the total respondents were from trading companies, followed by those in the services sector (25.3%, N = 97) and industrial companies (23.7%, N = 91) and other sectors (23.7%, N = 91).

**Table 4.2:** Demographic Profile of Respondents

<b>Constructs</b>	<b>Characteristics</b>	<b>N</b>	<b>%</b>
Business Activity	Trade	105	27.3
	Industrial	91	23.7
	Services	97	25.3
	Others	91	23.7
	<b>Total</b>	<b>384</b>	<b>100.0</b>
Current Position	Chief Executive Officer	85	22.1
	Managing Director	82	21.4
	Finance Manager	69	18.0
	Accountant	70	18.2
	Others	78	20.3
	<b>Total</b>	<b>384</b>	<b>100.0</b>
Average Number of Staff	10 or less	91	23.7
	11–50	99	25.8
	51–100	95	24.7
	101 or more	99	25.8
	<b>Total</b>	<b>384</b>	<b>100.0</b>
Registration for VAT	Yes	122	31.8
	No	143	37.2
	Not sure	119	31.0
	<b>Total</b>	<b>384</b>	<b>100.0</b>
Accounting Records	Manual	80	20.8
	In a partially computerised	96	25.0

system		
In a fully computerised system	104	27.1
Using external accountant	104	27.1
<b>Total</b>	<b>384</b>	<b>100.0</b>

Furthermore, in regards of the current position (designation) of the respondents, most of the respondents (22.1%, N = 85) indicated holding the position of Chief Executive Officer (CEO), which was followed by those in the position of Managing Director (MD) (21.4%, N = 82) and other positions (20.3%, N = 78). Different individuals, including CEOs, MDs, Finance Managers, accountants, and others, participated in this study's survey. Thus, the majority of the respondents in this study held a prominent position within the companies. Concerning the average number of staff, most of the participating SMEs in this study had either 11 to 50 staff members (25.8%, N = 99) or more than 101 staff members (25.8%, N = 99). Meanwhile, 24.5% (N = 95) of the total respondents indicated having 51 to 100 staff members, and another 23.7% (N = 91) indicated having 10 staff members or fewer.

Similarly, in regards of the registration for VAT, 37.2% (N = 143) of the total respondents indicated not registered for VAT. About 31.8% (N = 122) of the total respondents indicated registered. Meanwhile, 31% (N = 119) of the total respondents reported not sure. Finally, according to the accounting records, 27.1% (N = 04) of the total respondents indicated to use either a fully computerised system or an external account. The remaining 25% (N = 96) of the total respondents indicated using a partially computerised system. Only 20.8% (N = 80) of the total respondents reported using the manual approach.

#### 4.4.1 Compliance Cost

Compliance cost encompasses time cost and monetary cost. Section 3.4.1 explained the measurement of total compliance cost. Following the estimation of compliance cost, several statistical tests were conducted to examine the influence of e-filing on compliance cost (by comparing the e-filing and manual filing compliance costs). Prior to any statistical analysis, it was crucial to examine the condition of the data, specifically on whether the data were normally distributed and to determine the appropriate type of analysis (parametric or non-parametric tests). Referring to the results of normality test, most of the data for compliance cost estimation was not normally distributed ( $p < 0.05$ ). Therefore, non-parametric tests were employed for inferential statistics. Field (2009, p. 790) defined non-parametric tests as a family of statistical procedures that do not rely on the restrictive assumptions of parametric tests. In particular, non-parametric tests do not assume that the sampling distribution is normally distributed. In this study, Mann-Whitney U and Kruskal-Wallis tests were performed to analyse two independent groups and several independent groups, respectively. Pearson's chi-square was used to ascertain the presence of a relationship between two categorical variables.

#### 4.4.2 Internal Costs

Internal cost bias was tested in this study, as the reasons of why certain respondents did not return the questionnaire survey were not known (Benke & Street, 1992). In this study, the non-response bias was tested for the determinants of e-filing usage behaviour and compliance cost results. T-test was performed to compare the responses before reminders were sent (early responses) with the responses generated after the reminders (late responses) to ascertain the presence of any non-response bias

in the samples (Benke & Street, 1992). The responses generated after the reminders served as proxies for non-respondents. Selected items of perceptions towards the internal cost system were chosen. The means for these items were compared between the early response and late response samples to determine whether there is evidence of non-response bias affecting the determinants of internal cost usage behaviour.

Overall, there were 155 cases from the accounting staff group and 54 cases from the manager/accountant group. Based on the obtained p-value of two-tailed test, this study determined whether the groups were significantly different from each other.

As for the non-response bias in regards to internal cost compliance, an independent sample Mann-Whitney U test was performed to ascertain the differences between early and late response groups. This non-parametric test was used because the data for compliance costs were not normally distributed. Four major compliance cost items, specifically total time, value of time, total monetary cost, and total compliance cost, were included in the analysis. The results in Table 4.3 suggested similar distribution in the early and late response groups for these four items. Therefore, non-response bias was not a concern in this study for the analysis of the determinants of internal cost usage behaviour and compliance cost.

**Table 4.3:** Compliance Cost Items

Items	Early Response	Late Response	P-Value
	Group (N = 155)	Group (N = 54)	
Total time	4.83	4.00	0.71
Value of time	20.00	23.00	0.34
Total monetary cost	8.00	10.00	0.27
Total compliance cost	166.94	256.24	0.44

Notes: \* denotes 0.05 level of significance. A p-value of greater than 0.05 indicates no significant difference between both groups.

The formula used for the analysis is the t-test for independent samples. This statistical test compares the means of compliance cost items between two groups: the early responses group and the late responses group. Each compliance cost item, such as total time, value of time, total monetary costs, and total compliance cost, has corresponding mean values for both the early and late response groups. The t-test calculates a p-value for each item, which indicates whether there is a significant difference between the means of the two groups. If the p-value is greater than 0.05, it suggests that there is not a significant difference between the early and late responses for that particular compliance cost item. Conversely, if the p-value is less than or equal to 0.05, it indicates a significant difference between the means of the two groups for that item. Based on the responses provided by 384 SMEs, the compliance cost items commonly faced by SMEs and the obtained standard deviations are shown in Table 4.4.

**Table 4.4:** Test for Compliance Cost Items

Items	Early Responses (N = 384)	P-Value*
Staff time and training	71%	0.15
Equipment and technology upgrades	62%	0.17
Data management and reporting	51%	0.19
Auditing and monitoring	39%	0.14
Legal fees	34%	0.13
Insurance costs	26%	0.10
Opportunity costs (diverted resources)	19%	0.8

Notes: \* denotes 0.05 level of significance. A p-value of greater than 0.05 indicates no significant difference between both groups.

Standard deviation indicates the degree of variation or dispersion of the data from the mean. A higher standard deviation implies a wide dispersion of data that are less tightly clustered around the mean. For example, the standard deviation of staff time and training was 15% (P-Value), indicating a significant variation in the percentage of SMEs facing this particular item. The obtained results would provide better understanding on the range of compliance cost SMEs may encounter for the development of more accurate budgets and resource allocation plans. Through a data-driven approach to compliance, SMEs can minimise their exposure to compliance-related risks and improve their overall financial performance.

### 4.4.3 External Costs

The analysis revealed that nearly 95% of SMEs opted for professionals to manage their corporate tax affairs for the assessment year of 2022. It is notable that the maximum working hours for employees in the UAE during the month of Ramadan are six hours per day. This translates to a weekly limit of 36 hours for a six-day working week or 30 hours for a five-day working week. As of 5 April 2021, this arrangement was deemed effective (SHRM, 2022). External costs of non-compliance are straightforward monetary cost charged by external tax professionals solely on tax activities. It is common in the UAE to have tax fees charged separately from financial audits or other fees. However, certain SMEs may refuse to indicate their actual external costs. Thus, respondents were required to provide or estimate their external costs for the financial year of 2022.

As summarised in Table 4.5, in 2022, SMEs paid an average tax fee of AED25,000, with the lowest and highest external costs being AED150,000 and AED15,000, respectively (European Commission, Directorate-General for Taxation and Customs Union et al., 2022; Government of UAE, 2022). In other words, external costs can vary significantly, depending on the nature and size of the business, industry, location, and other factors. These external costs include legal fees, consultant fees, permit and license fees, and costs related to environmental compliance. Based on the obtained responses (n = 384), it can be concluded that external costs can be a significant burden for SMEs. SMEs should carefully manage and monitor their external costs to remain competitive and profitable. It is also important for SMEs to identify areas where they can potentially reduce costs, such as negotiating with suppliers, outsourcing non-core functions, or adopting more efficient processes and technologies.

**Table 4.5:** External Costs of SMEs in the UAE for the Financial Year of 2022

<b>Cost Category</b>	<b>Mean</b>	<b>Minimum<sup>a</sup></b>	<b>Maximum</b>	<b>Standard Deviation</b>
External Costs (AED)	25,000	15,000	150,000	15,000

Note: <sup>a</sup> denotes the exclusion of 10 SMEs with no component of external costs.

Table 4.6 shows legal fees as one of the most common external costs that SMEs in the UAE encounter, which may be due to legal advice required to navigate complex regulations, compliance with labour laws, or drafting legal documents like contracts. These legal fees ranged from AED5,000 to AED100,000, with an average cost of AED 25,000. Consultant fees are also part of the common external costs for SMEs (Government of UAE, 2022). Consultants can be hired to help SMEs navigate regulatory compliance requirements or to provide specialised expertise. Based on the provided responses, consultant fees can range from AED10,000 to AED75,000, with an average cost of AED30,000. Permit and license fees are also significant external costs for SMEs. These fees are incurred when SMEs need to obtain permits or licenses to operate their businesses. Based on the obtained responses, permit and license fees can range from AED2,500 to AED25,000, with an average cost of AED10,000 (European Commission, Directorate-General for Taxation and Customs Union et al., 2022). Costs related to environmental compliance are also a significant concern for SMEs. These costs can include the cost of implementing environmentally friendly practices, obtaining certifications, and complying with regulations. Based on the valid responses obtained, environmental compliance costs can range from AED5,000 to AED75,000, with an average cost of AED20,000.

**Table 4.6:** A Summary of External Costs

<b>External Costs</b>	<b>Mean</b>	<b>Standard Deviation</b>
15,000–50,000	22,500	9,250
20,001–100,000	75,000	15,750
100,001–150,000	125,000	22,500
150,001–200,000	175,000	23,000
200,001–250,000	225,000	26,000
250,001–300,000	275,000	28,250

#### **4.5 Descriptive Statistics**

This section presents the results of preliminary analysis in this study. Descriptive statistics were applied to identify and manage the data statistically. All data were verified using the mean, standard deviation, variance, skewness, and kurtosis. The subsequent tables present descriptive statistics (item-wise) that constitute item code, sample size, minimum and maximum values, standard deviation, variance, skewness, and kurtosis.

##### **4.5.1 Independent Variables**

Table 4.7 presents descriptive statistics (item-wise) for six independent variables, namely compliance cost, POD, tax rate, complexity, penalty, and incentive tax. The sample size (N) was consistent at 384. The mean value indicates the average score, while standard deviation reflects the dispersion of data around the mean value (variability). Overall, the mean values ranged from 3.61 (POD) to 4.65 (tax rate), and the values of standard deviation ranged from 0.961 (POD) to 2.568 (incentive tax). In particular, the recorded mean values for compliance cost ranged from 4.23 to 4.33. Tax

rate recorded mean values ranging from 4.21 to 4.65, while POD recorded mean values ranging from 3.61 to 3.90. The mean values for complexity varied from 4.42 to 4.35, while the mean values for penalty varied from 3.67 to 3.94. Lastly, the mean values for incentive tax ranged from 4.01 to 4.46. Meanwhile, variance, skewness, and kurtosis provide insights on the distribution of data. The values for skewness and kurtosis were found to be not within the recommended interval range of  $\pm 3$ , the cut-off point proposed (Ringle & Sarstedt, 2016). The recorded minimum and maximum values varied for each variable. These descriptive statistics offered a snapshot of the central tendency, variability, and distribution of each variable in the dataset, aiding in further analysis and interpretation.

**Table 4.7:** Descriptive Statistics (Item-Wise) for Independent Variables

	<b>N</b>	<b>Min.</b>	<b>Max.</b>	<b>Mean</b>	<b>SD</b>	<b>Variance</b>	<b>Skewness</b>	<b>Kurtosis</b>
<i>Compliance Cost</i>								
COMC1	384	1	7	4.33	1.635	2.672	-0.027	0.887
COMC2	384	1	7	4.23	1.782	3.174	-0.52	1.077
COMC3	384	1	7	4.23	1.675	2,809	-0.125	1.023
<i>Tax Rate</i>								
RT1	384	1	7	4.65	1.579	2.493	0.079	1.244
RT2	384	1	7	4.21	1.588	2.522	0.043	0.893
RT3	384	1	7	4.50	1.818	3.305	0.037	1.288
<i>Probability of Detection</i>								
POD1	384	1	7	3.61	1.178	3.027	-0.452	-0.785
POD2	384	1	7	3.78	1.262	2.920	-0.427	-0.887
POD3	384	1	7	3.90	1.146	3.419	-2.970	1.134
POD4	384	1	7	3.68	0.961	3.201	-0.125	1.050

	N	Min.	Max.	Mean	SD	Variance	Skewness	Kurtosis
<i>Complexity</i>								
COMP1	384	1	7	4.42	1.055	1.113	-1.297	1.213
COMP2	384	1	7	4.35	1.697	0.976	-1.267	1.464
COMP3	384	1	7	4.35	1.586	0.772	-1.387	2.201
COMP4	384	1	7	4.52	1.593	0.775	-1.456	3.556
<i>Penalty</i>								
PEN1	384	1	7	3.94	1.134	1.286	-1.124	0.508
PEN2	384	1	7	3.78	1.262	2.920	-0.427	-0.887
PEN3	384	1	7	3.68	0.961	3.201	-0.125	1.050
PEN4	384	1	7	3.67	1.431	2.411	0.202	0.887
<i>Incentive Tax</i>								
INCET1	384	1	7	4.05	1.886	3.812	-0.224	1.141
INCET2	384	1	7	4.31	1.503	2.337	-0.140	0.868
INCET3	384	1	7	4.01	1.690	2.666	0.350	1.026

	<b>N</b>	<b>Min.</b>	<b>Max.</b>	<b>Mean</b>	<b>SD</b>	<b>Variance</b>	<b>Skewness</b>	<b>Kurtosis</b>
INCET4	384	1	7	4.35	1.581	2.500	0.181	0.957
INCET5	384	1	7	4.35	2.522	2.442	0.125	1.253
INCET6	384	1	7	4.46	2.568	2.548	0.107	0.910
Valid N	384							

#### 4.5.2 Dependent Variable

Table 4.8 presents descriptive statistics (item-wise) for VAT compliance behaviour (dependent variable), which includes information on the minimum and maximum values, measures of central tendency and variability (mean and standard deviation), and variance. Accordingly, standard deviation reveals the extent of variability or dispersion of data around the mean value. Variance provides a complementary perspective on the spread of data points. Skewness and kurtosis values offer insights on the distributional characteristics of the data. Skewness values approaching zero suggest symmetry, while positive or negative values indicate skewness towards the right or left, respectively. Similarly, kurtosis values approaching zero indicate normal distribution, with positive values indicating peakedness, and negative values indicating flatness.

**Table 4.8:** VAT Compliance Behaviour

	<b>Min.</b>	<b>Max.</b>	<b>Mean</b>	<b>SD</b>	<b>Variance</b>	<b>Skewness</b>	<b>Kurtosis</b>
VAT1	1	7	4.43	1.851	3.425	0.028	-1.315
VAT2	1	7	3.67	1.721	2.963	0.412	-0.819
VAT3	1	7	3.84	1.686	2.842	0.398	0.910
VAT4	1	7	3.97	1.814	3.289	0.261	1.142

The minimum score of 1 indicated the lowest level of compliance, while the maximum score of 7 represented the highest level of compliance.

The mean values for VAT1, VAT2, VAT3, and VAT4 ranged from 3.67 to

4.43, reflecting the average level of compliance across different aspects of VAT compliance. Additionally, the values for skewness and kurtosis were not within the recommended interval range of  $\pm 3$ , the cut-off point proposed (Lei & Lomax, 2005 Wilson, 2010). The results also reaffirmed the absence of missing data for all items.

Based on these results, it was evident that SMEs in the UAE spent significant amount of money to comply with VAT regulations. This study estimated total cost for SMEs to comply with VAT at AED 12.8 billion. This generally includes expenses like registration, filing taxes, and implementing necessary procedures to ensure compliance. This study also looked at different aspects of compliance cost, such as the time and money spent, and found similar distribution of these costs among SMEs, regardless of whether they responded prior or after the reminders.

Interestingly, the study observed higher VAT compliance behaviour among SMEs following the increase in compliance cost. When SMEs face higher costs to comply with VAT regulations, they tend to adhere more closely to the regulations. This highlighted the financial burden of VAT on SMEs and underscored the importance of policymakers to understand and manage these costs effectively. The obtained results were specific to SMEs in the UAE and their compliance with VAT regulations. However, these results offered valuable insights that may be relevant to other contexts. The general principle on the relationship between higher compliance cost and increased compliance behaviour may hold across different regions and regulatory frameworks, suggesting the need for policymakers and business owners in various countries to learn from the UAE's experience and consider the

implications of compliance cost on businesses and the country's overall economy.

#### **4.6 Assessment of Measurement Model**

The assessment of measurement model encompasses composite reliability for the determination of internal consistency, outer loadings of items (individual item's reliability) and Cronbach's  $\alpha$  (overall dimension's reliability), average variance extracted (AVE) for the determination of convergent validity, and Fornell-Larcker criterion and cross-loadings for the determination of discriminant validity. The collected data were analysed using SmartPLS 4. Exploratory analysis was first conducted to assess the reliability and validity of the measurement scales used in this study and to identify all the purified items. In the second step, confirmatory analysis was carried out using a path weighting scheme, a maximum of 500 iterations, a stop criterion of 0.0000001 ( $1 \times 10^{-7}$ ), and equal indicator weights for the initialisation (Wilson, 2010; Ting et al., 2021). This analysis assessed the reliability and validity of the measurement model and to associate each latent variable with the corresponding indicator.

##### **4.6.1 Reliability Analysis**

The results of reliability for multiple indicators are tabulated in Table 4.9. All items recorded outer loadings of above 0.5, which demonstrated the items' quality representation of the data. Moreover, the results of reflective measurement model specified strong internal consistency. Compliance cost (0.855), tax rate (0.803), POD (0.879), complexity (0.825), penalty (0.872), incentive tax (0.894), and VAT compliance behaviour (0.879) also recorded high CR. Likewise, compliance cost (0.865), tax rate (0.830), POD (0.883),

complexity (0.865), penalty (0.878), incentive tax (0.908), and VAT compliance behaviour (0.873) recorded high Cronbach's  $\alpha$ , which exceeded the threshold value of 0.7. Thus, these constructs and items were deemed statistically reliable (Hair et al., 2021). Likewise, the results of AVE for compliance cost (0.781), tax rate (0.718), POD (0.735), complexity (0.659), penalty (0.727), incentive tax (0.660), and VAT compliance behaviour (0.720) exceeded 0.5, which corresponded to Fornell and Larcker's (1981) recommended threshold value for acceptable convergent validity. In this regard, this study's results of confirmatory analysis confirmed the validity of the instrument. This study empirically validated that each latent variable effectively explained more than 50% of the variance of its corresponding indicator.

**Table 4.9: Reliability and Validity Results**

Constructs	Items	Loadings	C $\alpha$	rho_A	CR	AVE
Compliance cost			0.855	0.865	0.915	0.781
	COMC1	0.803				
	COMC2	0.687				
	COMC3	0.845				
Tax rate			0.803	0.830	0.884	0.718
	RT1	0.673				
	RT2	0.778				
	RT3	0.712				

Constructs	Items	Loadings	Ca	rho_A	CR	AVE
Probability of detection			0.879	0.883	0.917	0.735
	POD1	0.630				
	POD2	0.754				
	POD3	0.794				
	POD4	0.762				
Complexity			0.825	0.865	0.885	0.659
	COMP1	0.718				
	COMP2	0.624				
	COMP3	0.767				
	COMP4	0.529				
Penalty			0.872	0.878	0.914	0.727
	PEN1	0.773				
	PEN2	0.715				
	PEN3	0.669				
	PEN4	0.746				
Incentive tax			0.894	0.908	0.921	0.660
	INCET1	0.703				
	INCET2	0.737				
	INCET3	0.549				
	INCET4	0.753				
	INCET5	0.582				
	INCET6	0.739				

Constructs	Items	Loadings	Ca	rho_A	CR	AVE
VAT compliance			0.879	0.873	0.911	0.720
behaviour	VAT1	0.719				
	VAT2	0.685				
	VAT3	0.684				
	VAT4	0.794				

The indicator of rho\_A for all constructs exceeded the proposed threshold value of 0.7. Thus, the results indicated that the reflective measurement model satisfied the relevant assessment criteria. Variance inflation factor (VIF) and t-statistics results provided for the assessment of measurement model and confirmed the model's robustness. However, the flexible approach can reach up to according to the researchers' established criteria. As all values of VIF in this study were lower than the minimum cut-off value ( $p > 3$ ), it can be concluded that all constructs of the conceptual model were reliable and valid at a statistical level (see Table 4.10).

**Table 4.10:** Variance Inflation Factor, T-Statistics, and Significance Results

First-Order Constructs	Items	VIF	T-Statistics	P-Value
Compliance cost	COMC1	2.226	41.388	0.000
	COMC2	1.681	18.033	0.000
	COMC3	2.511	41.142	0.000
Tax rate	RT1	1.439	13.565	0.000
	RT2	2.649	39.082	0.000

First-Order Constructs	Items	VIF	T-Statistics	P-Value
	RT3	1.760	24.200	0.000
Probability of detection	POD1	2.973	44.943	0.000
	POD2	2.762	43.395	0.000
	POD3	2.167	33.929	0.000
	POD4	1.939	21.443	0.000
Complexity	COMP1	2.536	48.420	0.000
	COMP2	2.444	21.955	0.000
	COMP3	2.736	45.690	0.000
	COMP4		53.444	0.000
Penalty	PEN1	2.569	40.663	0.000
	PEN2	2.017	13.234	0.000
	PEN3	1.765	31.636	0.000
	PEN4	2.500	34.244	0.000
Incentive tax	INCET1	2.814	59.426	0.000
	INCET2	1.733	37.825	0.000
	INCET3	2.966	18.962	0.000
	INCET4	1.646	35.062	0.000
	INCET5	1.935	13.901	0.000
	INCET6	1.999	44.993	0.000
VAT compliance behaviour	VAT1	2.162	31.051	0.000
	VAT2	1.935	23.787	0.000
	VAT3	1.999	58.986	0.000
	VAT4	2.683	53.325	0.000

#### 4.6.2 Discriminant Validity

Discriminant validity is assumed when the indicators that are intended to measure different phenomena are weakly correlated with one another (Cheung & Wang, 2017). According to Hair et al. (2021), discriminant validity is obtained when the variance extracted for each dimension exceeds the squared value of the correlation between the dimensions. In other words,  $R^2$  of each factor must exceed the squared value of its correlation with the other individual factor. Based on the results of Fornell-Larcker criterion in Table 4.11, the highlighted values for all constructs reflected adequate discriminant validity. Referring to Table 4.12, all indicators' loadings on their corresponding constructs exceeded the recommended threshold value of 0.6 and were lower on other constructs. The square root of AVE for each construct exceeded the inter-constructed correlations. These results proved the discriminant validity of the measurement model.

**Table 4.11:** Fornell-Larcker Criterion

	<b>COMP</b>	<b>COM</b>	<b>PEN</b>	<b>POD</b>	<b>TR</b>	<b>TICEN</b>	<b>VAT</b>
<b>COMP</b>	0.744						
<b>COMC</b>	1.061	0.822					
<b>PEN</b>	0.928	0.912	0.862				
<b>POD</b>	0.930	0.928	0.611	0.806			
<b>TR</b>	1.007	0.802	0.464	0.895	0.768		
<b>TICEN</b>	1.091	0.966	0.370	0.971	1.035	0.772	
<b>VAT</b>	0.939	0.860	0.360	1.072	0.994	0.996	0.792

**Table 4.12:** Summary of Cross-Loadings

	<b>COMP</b>	<b>COM</b>	<b>PEN</b>	<b>POD</b>	<b>TR</b>	<b>TICEN</b>	<b>VAT</b>
<b>COMP1</b>	0.792	0.891	0.716	0.770	0.722	0.872	0.730
<b>COMP2</b>	0.689	0.791	0.642	0.620	0.699	0.761	0.665
<b>COMP3</b>	0.852	0.892	0.782	0.797	0.766	0.895	0.805
<b>COMP4</b>	0.622	0.547	0.612	0.555	0.843	0.707	0.575
<b>COM1</b>	0.886	0.813	0.742	0.793	0.655	0.819	0.696
<b>COM1</b>	0.816	0.768	0.686	0.661	0.633	0.719	0.646
<b>COM1</b>	0.914	0.881	0.817	0.827	0.689	0.841	0.774
<b>PEN1</b>	0.801	0.778	0.831	0.818	0.917	0.868	0.797
<b>PEN2</b>	0.760	0.765	0.809	0.791	0.896	0.762	0.871
<b>PEN3</b>	0.619	0.615	0.724	0.833	0.661	0.651	0.862
<b>PEN4</b>	0.773	0.746	0.824	0.926	0.758	0.794	0.873
<b>POD1</b>	0.625	0.630	0.853	0.743	0.654	0.656	0.859

	COMP	COM	PEN	POD	TR	TICEN	VAT
<b>POD2</b>	0.772	0.753	0.897	0.826	0.747	0.799	0.877
<b>POD3</b>	0.776	0.791	0.853	0.849	0.753	0.859	0.928
<b>POD4</b>	0.817	0.809	0.795	0.801	0.727	0.804	0.790
<b>TR1</b>	0.761	0.508	0.603	0.537	0.698	0.706	0.569
<b>TR2</b>	0.809	0.661	0.867	0.760	0.826	0.893	0.803
<b>TR3</b>	0.771	0.666	0.845	0.744	0.855	0.786	0.884
<b>TINCEN1</b>	0.785	0.701	0.807	0.908	0.798	0.893	0.945
<b>TINCEN2</b>	0.929	0.859	0.707	0.770	0.740	0.736	0.737
<b>TINCEN3</b>	0.841	0.786	0.643	0.623	0.701	0.670	0.673
<b>TINCEN4</b>	0.951	0.887	0.783	0.801	0.768	0.814	0.814
<b>TINCEN5</b>	0.767	0.541	0.612	0.558	0.845	0.609	0.583
<b>TINCEN6</b>	0.821	0.714	0.893	0.791	0.962	0.871	0.818
<b>VAT1</b>	0.773	0.730	0.867	0.770	0.915	0.786	0.751
<b>VAT2</b>	0.623	0.558	0.813	0.802	0.677	0.664	0.747
<b>VAT3</b>	0.783	0.698	0.875	0.899	0.777	0.815	0.832
<b>VAT4</b>	0.790	0.735	0.821	0.918	0.787	0.884	0.835

#### 4.7 Assessment of Structural Model

Based on the exploratory and confirmatory results, the psychometric specificities of this study's measurement model were found satisfactory. As suggested by Myers et al. (2010), this study proceeded to assess the structural relations between the constructs. In order to examine the adjustment quality of the structural model, the values and significance of  $R^2$  were obtained. Hair et al. (2021) described  $R^2$  as the amount of the explained variance in

endogenous latent variables explained by its predictors in the structural model. The higher the  $R^2$  value, the better the construct is explained by the latent variables in the structural model that point at it via structural model path relationships. Besides that, Nakagawa et al. (2018) explained  $R^2$  as the predictive power of the exogenous variable. With that, PLS algorithm statistics were run using the path setup, yielding the results of  $R^2$  and  $R^2_{\text{adjusted}}$ . Table 4.12 shows what we found when we conducted that analysis.

#### 4.7.1 Predictive Relevance ( $Q^2$ )

Table 4.13 presents the obtained coefficients of determination and predictive relevance of all constructs in the structural model. Overall, the structural model demonstrated good quality of adjustment. The adjusted  $R^2$  and  $O^2$  clearly exceeded the threshold value of 0.5 for all constructs, which they have a very high value of  $R$ . Likewise, the values of  $Q^2$  exceeded zero for all constructs, reflecting the path model's predictive relevance (Hair et al., 2014). However, in our case, we have a positive and ( $Q^2 > 0.35$ ), which confirmed sufficient predictive relevance and significance of the model (Hair et al., 2014).

**Table 4.13:** Coefficient of Determination and Predictive Relevance

Exogenous Constructs	R <sup>2</sup>	Q <sup>2</sup>
Complexity	0.866	0.865
Penalty	0.861	0.861
Probability of Detection	0.801	0.801
Tax rate	0.643	0.642
Tax incentive	0.934	0.934
Vat compliance behaviour	1.089	1.091

#### 4.7.2 Effect Size ( $f^2$ )

According to Forster (2008),  $f^2$  is an important statistical analysis that ensures generalisable treatment effects to items from where these items are drawn that further indicate that the effect is suitably general among all the items. A larger effect size implies greater deviation from the null hypothesis, and vice versa. The null hypothesis corresponds to the nullity of effect size ( $ES = 0$ ). Dividing the difference between the two means by the standard deviation is intended to neutralise the effect of change in units. Following this explication, this study's results confirmed that our data present a very strong impact of one exogenous variable explaining a specific endogenous variable in  $R^2$ . In other words, the analysis reveals whether the deletion of a variable from the model exerts significant influence on the endogenous variable.

#### 4.7.3 Collinearity Statistics

This study also examined the potential presence of multicollinearity between the independent variables. Collinearity statistics determine whether

and to what extent the independent variables are correlated (Asthana, 2018; Shrestha, 2020). As noted by Daoud (2017), regression analysis-based studies are aimed at determining the correlation of the variables; however, this correlation is based on the independent variables and dependent variables. Correlation between independent variables in such studies is not favoured, as it further affects the generalisability of results. Notably, the presence of multicollinearity between two or more predictor variables raises the chances of standard error of coefficients (Kim, 2019), affecting different factors and reducing the model's reliability. The tabulated results in Table 4.14 demonstrated no correlations between the factors, suggesting no multicollinearity. Thus, this study encountered no collinearity problem in the data used.

**Table 4.14:** Variance Inflation Factor

	Effect Size						Variance Inflation Factor						
	COMP	PEN	POD	TR	TICEN	VAT	PEN	COMP	PEN	POD	TR	TICEN	VAT
<b>COMP</b>		1.752				0.003		1.000					-3.358
<b>COM</b>				0.857		0.092				1.000			-4.073
<b>PEN</b>						0.094							-9.329
<b>POD</b>	1.806				2.908	0.323		1.000			1.000		-8.730
<b>TR</b>						0.000							-1.858
<b>TICEN</b>			1.421			0.053				1.000			-13.159
<b>VAT</b>													

#### 4.7.4 Testing of Variables

This study proceeded to test the proposed variables to ascertain the proposed relationships in the conceptual framework (Habes et al., 2023). Path analysis, including regression weights and intercepts (t-value, bootstrapping, and statistical significance in terms of p-value), was conducted. This section summarises the results of path analysis. The main results are further discussed for better understanding.

The obtained statistical results provided valuable insights on the relationships between different constructs and their significance in influencing VAT compliance behaviour among SMEs. Statistical significance reflects the strength and reliability of the observed relationships. In this study, statistical significance was primarily determined by the obtained p-values of the variables relationships. A p-value of below the predetermined threshold value, which is often set at 0.05 (5%), indicates a statistically significant relationship. In contrast, a p-value that exceeds the threshold value suggests that the relationship is not statistically significant and may be due to random chance. The correlations between the constructs were assessed through the bootstrapping procedure, which provided estimation by examining T (standard deviation), which must be superior to 2 and the significance of the links of correlation (p), justified when the values (p) are lower than the risk of 5%.

This study first examined the direct relationships between the constructs, as listed in Table 4.15. The statistical results revealed significant correlations between the constructs ( $p < 0.05$ ), except for the correlation between tax rate and VAT compliance behaviour ( $5\% < 0.950$ ) and the

correlation between complexity and VAT compliance behaviour ( $5\% < 0.595$ ). In other words, all the other variables were supported, confirming the presence of direct relationships with VAT compliance behaviour. This study demonstrated the results of direct effect analysis, which confirmed the positive and significant influence of compliance cost ( $\beta = 0.798$ ,  $t = 42.603$ ,  $p < 0.000$ ), penalty ( $\beta = 0.351$ ,  $t = 3.033$ ,  $p < 0.001$ ), tax incentive ( $\beta = 0.289$ ,  $t = 2,361$ ,  $p < 0.018$ ), and POD ( $\beta = 0.521$ ,  $t = 6,880$ ,  $p < 0.000$ ) on VAT compliance behaviour. In conclusion, H1, H3, H4, H5, H6 H7, H8, H9, and H11 were supported, whereas H2 and H10 were not supported (see Table 4.15).

**Table 4.15:** Results of Direct Relationships

Variables Path	Path Coefficient	T-Statistics	P-Value	Decision
Compliance cost → VAT compliance behaviour	0.894	13.405	0.000	Accepted
Penalty → VAT compliance behaviour	0.756	9.735	0.000	Accepted
Probability of detection → VAT compliance behaviour	0.447	12.001	0.000	Accepted
Tax rate → VAT compliance behaviour	3.601	5.262	0.000	Accepted
Tax incentive → VAT compliance behaviour	0.807	6.721	0.000	Accepted

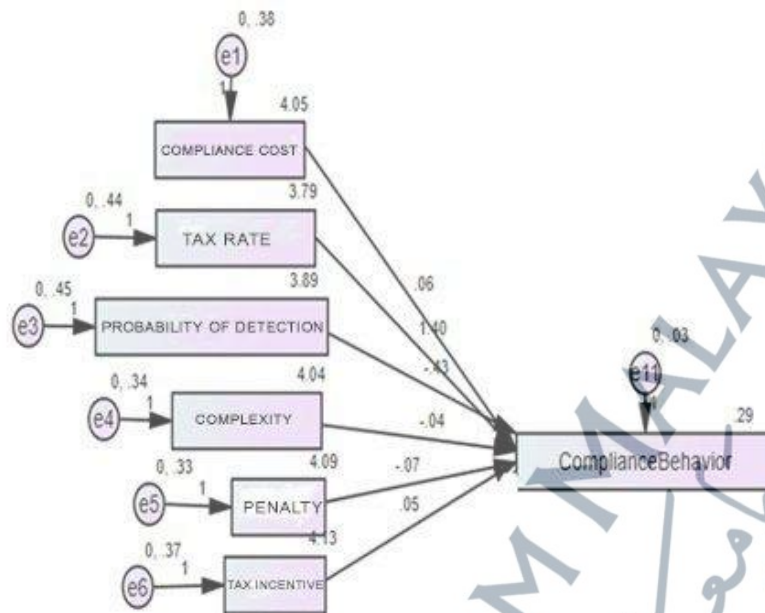
#### 4.7.5 Regression Analysis

Following the application of PLS-SEM, this study proceeded to multiple regression analysis to reaffirm the results obtained from the prior analyses. According to Groß (2003), SEM and regression analysis are statistical techniques that analyse the underlying relationships among the constructs of interest. SEM allows for examining complex relationships between latent constructs and observed variables, while regression analysis focuses on understanding the relationship(s) of one or more independent variables with a dependent variable.

The results of multiple regression analysis in Table 4.16 revealed crucial insights on the factors influencing VAT compliance behaviour among SMEs. The results demonstrated the significant increase in penalty with increased complexity in tax regulations ( $\beta = 5.536$ ,  $t = 21.378$ ,  $p = 0.000$ ). However, complexity did not exert direct influence on VAT compliance behaviour ( $\beta = -0.0025$ ,  $t = 0.630$ ,  $p = 0.531$ ). SMEs appeared to demonstrate higher level of compliance when they encountered higher compliance cost ( $\beta = 2.737$ ,  $t = 13.777$ ,  $p = 0.000$ ) and were apprehensive of the penalty ( $\beta = 1.893$ ,  $t = 54.674$ ,  $p = 0.000$ ). Moreover, the likelihood of being detected for non-compliance ( $\beta = 0.912$ ,  $t = 7.732$ ,  $p = 0.000$ ) and the availability of tax incentives ( $\beta = 2.901$ ,  $t = 6.424$ ,  $p = 0.000$ ) significantly influenced SMEs' adherence to VAT regulations. Interestingly, changes in tax rates did not seem to directly influence VAT compliance behaviour ( $\beta = 0.591$ ,  $t = 0.362$ ,  $p = 1.463$ ). These statistical results underscored the multifaceted nature of VAT compliance and offered valuable insights for policymakers and businesses alike in fostering greater adherence to tax regulations. Addressing the specific objective of determining the relationship between compliance cost and compliance behaviour of VAT among SMEs across diverse industries. The results of regression analysis revealed statistical significance ( $p = 0.000$ ).

**Table 4.16:** Results of Multiple Regression Analysis

<b>Variables Paths</b>	<b>Coefficient</b>	<b>T-Value</b>	<b>P-Value</b>
Complexity → Penalty	5.536	21.378	0.000
Complexity → VAT compliance behaviour	-0.0025	0.630	0.531
Compliance cost → VAT compliance behaviour	2.737	13.777	0.000
Compliance cost → VAT compliance behaviour	0.926	10.541	0.000
Penalty → VAT compliance behaviour	1.893	54.674	0.000
Penalty → VAT compliance behaviour	3.372	25.640	0.000
Probability of detection → VAT compliance behaviour	0.912	7.732	0.000
Probability of detection → VAT compliance behaviour	0.428	5.730	0.000
Probability of detection	0.529	2.983	0.000
Tax rate → VAT compliance behaviour	0.591	0.362	1.463
Tax incentive → VAT compliance behaviour	2.901	6.424	0.000



**Figure 4.1: Path Coefficients**

The obtained p-value of 0.000 indicated a highly significant relationship between compliance cost and compliance behaviour of VAT among SMEs across different industries. In other words, the variations in compliance cost can significantly predict the variations in compliance behaviour within each industry segment. The results of regression analysis allowed for detailed examination and interpretation of how compliance cost can influence compliance behaviour within specific industries, providing valuable insights on the dynamics of VAT compliance across different industries. The obtained results highlighted the importance of considering industry-specific factors when it comes to analysing VAT compliance practices among SMEs.

#### **4.8 Chapter Summary**

This chapter discussed the analysis of the obtained data and provided the main results. Notably, the results were tabulated, discussed, and interpreted for better

understanding. This study first descriptively analysed the collected data. Following that, the data were subjected to PLS-SEM and regression analysis. The results of the testing of all variables were presented in this chapter.

