

SPECIFYING THE PATH MODEL AND EXAMINING DATA

CASE STUDY ILLUSTRATION—SPECIFYING THE PLS-SEM MODEL

The most effective way to learn how to use a statistical method is to apply it to a set of data. Throughout this book, we use a single example that enables you to do that. We start the example with a simple model, and in Chapter 5, we expand that same model to a much broader, more complex model. For our initial model, we hypothesize a path model to estimate the relationships between corporate reputation, customer satisfaction, and customer loyalty. The example will provide insights on (1) how to develop the structural model representing the underlying concepts/theory, (2) the setup of measurement models for the latent variables, and (3) the structure of the empirical data used. Then, our focus shifts to setting up the SmartPLS 4 software (Ringle, Wende, & Becker, 2022) for PLS-SEM.

Application of Stage 1: Structural Model Specification

To specify the structural model, we must begin with some fundamental explications about theoretical models. The corporate reputation model by Eberl (2010) is the basis of our theory. The goal of the model is to explain the effects of corporate reputation on customer satisfaction (*CUSA*) and, ultimately, customer loyalty (*CUSL*). Corporate reputation represents a company's overall evaluation by its stakeholders (Helm, Eggert, & Garnefeld, 2010). It

is measured using two dimensions. One dimension represents cognitive evaluations of the company, and the construct is the company's competence (COMP). The second dimension captures affective judgments, which determine the company's likeability (LIKE). This two-dimensional approach to measure reputation was developed by Schwaiger (2004). It has been validated in different countries (e.g., Eberl, 2010; Zhang & Schwaiger, 2012) and applied in various research studies (e.g., Eberl & Schwaiger, 2005; Radomir & Moisescu, 2019; Radomir & Wilson, 2018; Raithel & Schwaiger, 2015; Raithel, Wilczynski, Schloderer, & Schwaiger, 2010; Sarstedt & Schloderer, 2010; Schloderer, Sarstedt, & Ringle, 2014; Schwaiger, Raithel, & Schloderer, 2009; Yun, Kim, & Cheong, 2020). Research also shows that the approach performs favorably (in terms of convergent validity and predictive validity) compared with alternative reputation measures (Sarstedt, Wilczynski, & Melewar, 2013).

Building on a definition of corporate reputation as an attitude-related construct, Schwaiger (2004) further identified four antecedent dimensions of reputation-quality, performance, attractiveness, and corporate social responsibility-measured by a total of 21 formative indicators. These driver constructs of corporate reputation are components of the more complex example we will use in the book and will be added in Chapter 5. Likewise, we do not consider more complex model setups such as mediation or moderation effects yet. These aspects will be covered in the case studies in Chapter 7.

In summary, the simple corporate reputation model has two main theoretical components: (1) the target constructs of interest-namely, CUSA and CUSL (endogenous constructs)—and (2) the two corporate reputation dimensions COMP and LIKE (exogenous constructs), which represent key determinants of the target constructs. Exhibit A2.1 shows the constructs and their relationships, which represent the structural model for the PLS-SEM case study.

To propose a theory, researchers usually build on existing research knowledge. When PLS-SEM is applied, the structural model displays the theory with its key elements (i.e., constructs) and cause-effect relationships (i.e., paths). Researchers typically develop hypotheses for the constructs and their path relationships in the structural model. For example, consider Hypothesis 1 (H₁): Customer satisfaction has a positive effect on customer loyalty. PLS-SEM enables statistically testing the significance of the hypothesized relationship (Chapter 6). When conceptualizing the theoretical constructs and their hypothesized structural relationships for PLS-SEM, it is important to make sure the model has no circular relationships (i.e., causal loops). A circular relationship would occur if, for example, we reversed the relationship between COMP and CUSL as this would yield the causal loop COMP \rightarrow CUSA \rightarrow $CUSL \rightarrow COMP$.





Application of Stage 2: Measurement Model Specification

Since the constructs are not directly observed, we need to specify a measurement model for each construct. The specification of the measurement models (i.e., multi-item vs. single-item measures and reflective vs. formative measures) draws on prior research studies by Schwaiger (2004) and Eberl (2010).

In our simple example of a PLS-SEM application, we have three constructs (*COMP, CUSL*, and *LIKE*) measured by multiple items (Exhibit A2.2). All three constructs have reflective measurement models as indicated by the arrows pointing from the construct to the indicators. For example, *COMP* is measured by means of the three reflective items *comp_1, comp_2*, and *comp_3*, which relate to the following survey questions (Exhibit A2.3): "[The company] is a top competitor in its market," "As far as I know, [the company] is recognized worldwide," and "I believe that [the company] performs at a premium level." Respondents had to indicate the degree to which they (dis)agree with each of the statements on a 7-point scale from 1 = fully disagree to 7 = fully agree.

Different from COMP, CUSL, and LIKE, the customer satisfaction construct (CUSA) is operationalized by a single item (cusa) that is related to the following question in the survey: "If you consider your experiences with [company], how satisfied are you with [company]?" The single indicator is measured with a 7-point scale indicating the respondent's degree of satisfaction (1 = very dissatisfied; 7 = very satisfied).

4 A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)

Exhibit A2.2 Types of Measurement Models in the Simple Model



Exhibit A	2.3 Indicators for Reflective Measurement Model Constructs
	Competence (COMP)
comp_1	[The company] is a top competitor in its market.
comp_2	As far as I know, [the company] is recognized worldwide.
comp_3	I believe that [the company] performs at a premium level.
	Likeability (LIKE)
like_1	[The company] is a company that I can better identify with than other companies.
like_2	[The company] is a company that I would regret more not having if it no longer existed than I would other companies.
like_3	I regard [the company] as a likeable company.
	Customer Loyalty (CUSL)
cusl_1	I would recommend [company] to friends and relatives.
cusl_2	If I had to choose again, I would choose [company] as my mobile phone services provider.
cusl_3	I will remain a customer of [company] in the future.

Note: For data collection, the actual name of the company was inserted in the bracketed space that indicates company.

The single item has been used due to practical considerations in an effort to decrease the overall number of items in the questionnaire. As customer satisfaction items are usually highly homogeneous, the loss in predictive validity compared with a multi-item measure is not considered severe. As *cusa* is the only item measuring customer satisfaction, construct and item are equivalent (as indicated by the fact that the relationship between construct and single-item measure is always one in PLS-SEM). Therefore, the choice of the measurement perspective (i.e., reflective vs. formative) is of no concern and the relationship between construct and indicator is undirected.

Application of Stage 3: Data Collection and Examination

To estimate the PLS-SEM, data were collected using computer-assisted telephone interviews (Sarstedt & Mooi, 2019) that asked about the respondents' perception of and their satisfaction with four major mobile network providers in Germany's mobile communications market. Respondents rated the questions on 7-point Likert scales, with higher scores denoting higher levels of agreement with a particular statement. In the case of *cusa*, higher scores denote higher levels of satisfaction. Satisfaction and loyalty were measured with respect to the respondents' own service providers. The data set used in this book is a subset of the original set and has a sample size of 344 observations. The data have been collected using a quota sampling approach (Sarstedt, Bengart, Shaltoni, & Lehmann, 2018) by a professional market research company in the German market. The resulting sample is representative of the German population.

Exhibit A2.4 shows the data matrix for the model. The 10 columns represent a subset of all variables (i.e., specific questions in the survey as described in the previous section) that have been surveyed, and the 344 rows (i.e., cases) contain the answers of every respondent to these questions. For example, the first row contains the answers of Respondent 1 while the last row contains the answers of Respondent 1 while the last row contains the answers of Respondent 344. The columns show the answers to the survey questions. Data in the first nine columns are for the indicators associated with the three constructs, and the tenth column includes the data for the single indicator for *CUSA*. The data set contains further variables that relate to, for example, the driver constructs of *LIKE* and *COMP*. We will cover these aspects in Chapter 5.

If you are using a data set in which a respondent did not answer a specific question, you need to insert a number that does not appear otherwise in the responses to indicate the missing values. Researchers commonly use -99 to indicate missing values, but you can use any other value that does not normally occur in the data set. In the following, we will also use -99 to indicate missing values. If, for example, the first data point of *comp_1* were a missing value, the -99 value would be inserted into the space as a missing value space holder instead of the value of 6 that you see in Exhibit A2.4. Missing value

treatment procedures (e.g., mean replacement) could then be applied to these data (e.g., Hair, Black, Babin, & Anderson, 2019). Again, if the number of missing values in your data set per indicator is relatively small (i.e., less than 5% missing per indicator), we recommend mean value replacement instead of casewise deletion to treat the missing values when running PLS-SEM. Furthermore, we need to ascertain that the number of missing values per observation and per indicator does not exceed 15%. If this was the case, the corresponding observation should be eliminated from the data set.

Exhibi	t A2.4	Data	a Matrix	for tł	ne Indi	cator	Variat	oles			
Case				١	/ariable	Name					
Number	comp_1	comp_2	comp_3	like_1	like_2	like_3	cusl_1	cusl_2	cusl_3	cusa	
1	6	7	6	6	6	6	7	7	7	7	
2	4	5	6	5	5	5	7	7	5	6	
344	6	5	6	6	7	5	7	7	7	7	7

The data example shown in Exhibit A2.4 (and in the book's example) has only very few missing values. More precisely, *cusa* has one missing value (0.29%), *cusl_1* and *cusl_3* have three missing values (0.87%), and *cusl_2* has four missing values (1.16%). Since the missing values per indicator are less than 5%, mean value replacement can be used. Furthermore, none of the observations and indicators has more than 15% missing values, so we can proceed analyzing all 344 respondents.

To run outlier diagnostics, we compute a series of box plots using IBM SPSS Statistics—see Chapter 5 in Sarstedt and Mooi (2019) for details on how to run these analyses in IBM SPSS Statistics. The results indicate some influential observations but no outliers. Moreover, nonnormality of data regarding skewness and kurtosis is not an issue. The kurtosis and skewness values of all the indicators are within the -2 and +2 range.

PATH MODEL CREATION USING THE SMARTPLS SOFTWARE

The SmartPLS 4 software (Ringle, Wende, & Becker, 2022) is used to execute all the PLS-SEM analyses in this book. The discussion includes an overview of the software's functionalities. The student version of the software is available free of charge at https://www.smartpls.com. The student version offers practically all functionalities of the full version but is restricted to data sets with a maximum of 100 observations. However, as the data set used in this book has more than 100 observations (344 to be precise), you should use the professional version of SmartPLS, which is available as a 30-day trial version at https://www.smartpls.com. After the trial period, a license fee applies. Licenses are available for different periods of time (e.g., 1 month, 1 year, 2 years etc.) and can be purchased through the SmartPLS website. The SmartPLS website includes a download area and many additional resources such as short explanations of PLS-SEM and software related topics, a list of recommended literature, answers to frequently asked questions, tutorial videos for getting started using the software, and the SmartPLS forum, which allows you to discuss PLS-SEM topics with other users.

SmartPLS has a graphical user interface that enables the user to estimate the PLS path model. Exhibit A2.8 at the end of this section shows the graphical interface for the SmartPLS software, with the simple model already drawn. In the following paragraphs, we describe how to set up this model using the SmartPLS software. Before you draw your model, you need to have data that serve as the basis for running the model. SmartPLS 4 supports data imported from various file formats, such as Microsoft Excel (.xls or xlsx), SPSS (.sav), comma-separated values (.csv), and text (.txt). The only aspect we have to pay attention to is that the first row contains the variable names in text format and otherwise only numerical values (no text or special characters; also, no numerical values in scientific format, e.g., 10 E-7). For example, SmartPLS interprets single dots (such as those produced by IBM SPSS Statistics in case an observation has a system-missing value) as string elements.

The data we will use with the reputation model can be downloaded either as comma-separated value (.csv) or text (.txt) data sets in the download section of this book's webpage at the following URL: https://www.pls-sem.net. Now run the SmartPLS software by clicking on the desktop icon that is available after the software installation on your computer device. Alternatively, go to the folder where you installed the SmartPLS software on your computer. Click on the icon that runs SmartPLS to start the software.

To create a new project after running SmartPLS, use the *New project* option in the toolbar. You can also click on *Files* \rightarrow *Create new* \rightarrow *New project*. First type a name for the project into the *Name* box (e.g., *Corporate reputation*). After clicking *Create*, the new project is created and appears in the *Workspace* window that is in the left below the menu bar. All previously created SmartPLS projects also appear in this window. On top of the *Workspace* window, you see the path to your *Workspace* folder (e.g., on your computer) where your SmartPLS projects are saved. You can select or create a new *Workspace* (e.g., on your computer device or on a cloud drive) by clicking the *Switch* button and, then, by selecting a previously used workspace from the list or by clicking on the link with the label *switch or create another workspace*.

Next, you need to assign a data set to the project, in our case, *Corporate* reputation data.txt (or whatever name you gave to the data you downloaded).

To do so, click on the information button labeled Import data file below the project you just created, find and highlight your data folder, and click Open. A window opens (Exhibit A2.5) that offers a general overview of the data set, including a list of all variables, their measurement scales as identified by the SmartPLS software, as well as the observed minimum and maximum values, you can specify the *Delimiter character* to determine the separation of the data values in your data set (i.e., comma, semicolon, tabulator, space), the Escape character (i.e., none, single quote, double quote) in case the values use quotations (e.g., "7"), and the Locale (i.e., United States with a dot as decimal separator or Europe with a comma as decimal separator). SmartPLS automatically reads the minimum and the maximum, but respondents may have not made use of the full scale (e.g., the actual minimum value is 2 instead of 1 or the maximum value is 6 instead of 7). You can change the minimum and the maximum value for each variable manually by entering the correct value under Min and/or Max in the Change settings window. For simultaneous changes you can use the Bulk change option. Under Missing value treatment (i.e., the yellow box at the bottom), SmartPLS indicated that it did not identify any missing values in this dataset. Type -99 into the text field and click on the Apply marker button. Thereafter, the color of the box changes to green and the message that SmartPLS identified 11 missing values is displayed.

roie	et			File name			
Cor	porate reputation		-	Corporate reput	ation data		
sv	file settings						
Delin	liter character	Escape character		Locale		Encoding	
Sen	nicolon	▼ None	•	US (example: 1,	000.23) 👻	UTF-8	
Cho	ose data to import (-1 cases and 41 indicat	ors)				
\checkmark	Name	Missing	Scale (Bulk	change)	N	lin	Max
\checkmark	serviceprovider	0	Metric	•	1.000	0	4.0000
\checkmark	servicetype	0	Binary	•	1.000	0	2.0000
\checkmark	comp_1	0	Metric	•	1.000	0	7.0000
\checkmark	comp_2	0	Metric	•	1.000	0	7.0000
\checkmark	comp_3	0	Metric	•	1.000	0	7.0000
~	like_1	0	Metric	•	1.000	0	7.0000
\checkmark	like_2	0	Metric	•	1.000	0	7.0000

8

2 Indicators													
Name	No.	Type	Missings	Mean	Median	Scale min	Scale max	Observed min	Observed max	Standard deviation	Excess kurtosis	Skewness	Cramér-von Mises p value
serviceprovider	-	CAT	0	2.000	2.000	1.000	4.000	1.000	4.000	1.003	-0.513	0.747	0.000
ervicetype	2	011	0	1.637	2.000	1.000	2.000	1.000	2.000	0.481	-1.684	-0.571	0.000
:omp_1	e	MET	0	4.648	5.000	1.000	7.000	1.000	7.000	1.433	-0.324	-0.264	0.000
tomp_2	4	MET	0	5.424	6.000	1.000	7.000	1.000	7.000	1.375	-0.616	-0.566	0:000
tomp_3	5	MET	0	5.221	6.000	1.000	7.000	1.000	7.000	1.458	-0.188	-0.677	0:000
ike_1	9	MET	0	4.584	5.000	1.000	7.000	1.000	7.000	1.547	-0.399	-0.405	0:000
ike_2	7	MET	0	4.250	4.000	1.000	7.000	1.000	7.000	1.848	-0.901	-0.312	0.000
ike_3	80	MET	0	4.480	5.000	1.000	7.000	1.000	7.000	1.871	-0.941	-0.325	0000
tusi_1	6	MET	3	5.129	5.000	1.000	7.000	1.000	7.000	1.513	0.268	-0.792	0.000
tusl_2	10	MET	4	5.276	6.000	1.000	7.000	1.000	7.000	1.744	0:040	-0.951	0.000
susi_3	5	MET	3	5.651	6.000	1.000	7.000	1.000	7.000	1.655	0.930	-1.301	0.000
tusa	12	MET	-	5.440	6.000	1.000	7.000	1.000	7.000	1.174	0.777	-0.768	0.000
:sor_1	13	MET	0	4.235	4.000	1.000	7.000	1.000	7.000	1.469	-0.383	-0.042	0:000
sor_2	14	MET	0	3.076	3.000	1.000	7.000	1.000	7.000	1.651	-0.564	0.497	0.000
tsor_3	15	MET	0	3.988	4.000	1.000	7.000	1.000	7.000	1.478	-0.468	-0.061	0.000
tsor_4	16	MET	0	3.125	3.000	1.000	7.000	1.000	7.000	1.462	-0.463	0.197	0.000
tsor_5	17	MET	0	3.983	4.000	1.000	7.000	1.000	7.000	1.583	-0.696	-0.042	0.000
ssor_global	18	MET	0	4.988	5.000	1.000	7.000	1.000	7.000	1.289	-0.629	-0.142	0.000
attr_1	19	MET	0	4.991	5.000	1.000	7.000	1.000	7.000	1.458	-0.078	-0.567	0.000
attr_2	20	MET	0	2.945	2.000	1.000	7.000	1.000	7.000	2.098	-1.131	0.574	0.000
attr_3	21	MET	0	4.811	5.000	1.000	7.000	1.000	7.000	1.451	-0.568	-0.275	0.000
ittr_global	22	MET	0	5.587	6.000	2.000	7 000	000 0	2 000	1 214	-0.100	D.CEA	0000

9

Finally, at the very top of the window, you can select the target *Project* in the *Workspace* (e.g., *Corporate reputation*) and change the *File name* of the data set. In this example, we use the original name (i.e., *Corporate reputation data*) and proceed by clicking *Import*. SmartPLS will open a new tab (Exhibit A2.6), which provides information on the data set and its format (data view).

Below the menu bar, at the right of the screen appears a list with all variables, their number of missing values, the scale type, and basic descriptive statistics (e.g., mean, median, scale minimum, and scale maximum values, observed minimum, and observed maximum values, standard deviation, excess kurtosis, skewness, and Cramér-von Mises *p* values). At the top left of the list, you can see the number of indicators as well as sample size and the number of missing values. In case you need to adjust the data set settings (e.g., the missing value marker), click on *Setup* in the menu bar. Any changes that you make will dynamically update the values in the data view. For example, adjusting the minimum values will change the values in the *Scale min* column.

The window left of the data view allows you to navigate to the indicator correlation matric (*Indicator correlations*), *Data groups*, and *Raw data*. Note that in order to view the data groups, you need to first specify a grouping. This can be done by clicking on *Add group* or *Generate groups* in the menu bar at the top of the screen. However, as the specification of grouping variables is primarily relevant for more advanced analyses, we skip this aspect for now.

At this point, you can click on the orange arrow button with the label *Back* in the menu bar and return to the *Workspace* view. Note that you can always reopen the data view by double-clicking on the data set (i.e., *Corporate reputation data*) in the *Workspace* view.

Each project can have one or more path models and one or more data sets (i.e., .csv, .txt, .xlsx or .sav files). To create the first model, click on the information button labeled *Create model* and in the combo box, choose *PLS-SEM* under *Model type*, enter the *Model name* (e.g., *Simple model*), and click on *Save*. Note that you can also create a model by right-clicking on the project name and choosing *New PLS-SEM model* option from the dialog that opens.

SmartPLS now opens a new window, which shows the indicator list on the left and the *Modeling window* on the right where you can create a path model. As we initiated a new project (as opposed to working with a saved project), the modeling window is empty. You can start creating the path model shown in Exhibit A2.7. To do so, select *Latent variable* in the menu bar () and left-click in the *Modeling window*. A text box is shown, which suggests a name for the first latent variable (*Alpha*), which you should change into *COMP*. When pressing the *Enter* key on your keyboard, SmartPLS will place the construct labeled *COMP* into the *Modeling window*. You can rename the latent variable by right-clicking on the construct and clicking on *Rename* in the window that opens. Each time you left-click in the *Modeling window*, a new text box for entering the name of the new construct will appear. After press-



ing the *Enter* key, a new construct represented by a red circle will be added. To leave this mode, click on *Select* in the menu bar (①).

Once you have included all your constructs, you can left-click on any of the constructs to select, resize, or move it in the *Modeling window*. To connect the latent variables with each other (i.e., to draw path arrows), left-click on *Connect* in the menu bar (2). Next, left-click on an exogenous (independent) construct and move the cursor over the target endogenous (dependent) construct. Now left-click on the endogenous construct and a path relationship (directional arrow) will be inserted between the two constructs. Repeat the same process and connect all the constructs based on your structural theory. When you finish, it will look like Exhibit A2.7.

Next, you need to assign indicators to each of the constructs. On the left side of the screen, there is an Indicators window that shows all the indicators that are in your data set along with some basic descriptive statistics that you can request by left-clicking on an indicator. Start with the COMP construct by dragging the first competence indicator comp_1 from the Indicators window and dropping it on the construct (i.e., left-click the mouse and hold the button down, then move it until over a construct, then release). After assigning an indicator to a construct, it appears in the graphical Modeling window as a yellow rectangle attached to the construct. Assigning an indicator to a construct will also turn the color of the construct from red to blue. You can move the indicator around, but it will remain attached to the construct (unless you delete it). By right-clicking on the construct and choosing one of the options under Align (e.g., Align indicators to the top), you can align the indicator(s). You can also hide the indicators of a construct by selecting the corresponding option in the menu that appears when right-clicking on it. You can also access the align indicators option via the *Tools* on the right-hand side of the Modeling window. Continue until you have assigned all the indicators to the constructs as shown in Exhibit A2.8. Make sure to save the model by clicking on *Save* in the menu bar (

Right-clicking on selected construct(s) in the *Modeling window* opens a menu with several options. Apart from renaming the constructs, you can, for example, invert the measurement model from reflective to formative measurement, and vice versa (*Invert measurement model*), *Hide indicators* and *Show indicators* of the construct. Additionally, when double-clicking on construct, a different menu opens that allows you to select an indicator weighting scheme per construct (i.e., *Automatic, Mode A, Mode B, Equal Weights*, and *Pre-defined*) and prespecify a value for the (formative) construct reliability between 0 and 1 for formatively measured constructs. To add a note to your *Modeling window*, left-click on the *Comment* button () in the menu bar.

Clicking on the right mouse button while the cursor is placed over other elements also opens a menu with additional functions. As a further example, if you place the cursor in the *Workspace window* and right-click on the project name, you can create a new model (for example, *New PLS-SEM model, New*



REGRESSION model, New PROCESS model, and New CB-SEM model), create a new project (New project), or import a new data set (Import data file). Moreover, you can select the Copy resource, Paste resource, and Delete resource options for projects and models that appear in the Workspace window. For example, the Duplicate option is useful when you would like to modify a PLS path model but want to keep your initial model setup. Using the Copy resource option, you can copy and paste a PLS path model from one project to the other. The Import data file option allows you to add more data sets to an existing project (e.g., data from different years if available). You can also export a project by selecting the Export project option. Using this option, SmartPLS will export the entire project, including all models and data sets you may have included in it, in a .zip folder. You can also directly import this "ready-to-use" project by going to *Files* \rightarrow *Import project from backup file*. You can use this option to import the project that includes the PLS-SEM example on corporate reputation. The file name is Corporate Reputation.zip. This project is ready to download on your computer system in the download section at https://www.pls-sem.net. Download this file and save it on your computer system. Then, go to *Files* \rightarrow *Import project from backup file*. SmartPLS allows you to browse your computer and select the downloaded project Corporate Reputation.zip for import. Alternately, under Sample Projects (i.e., on the lower part of the main screen), you find the Corporate reputation-PLS-SEM book (primer) example in the list of ready-to-run SmartPLS project examples. After clicking on the link with the label Install next to the sample project's name, it will appear as a new project in the Workspace. After successful import, double-click on the model in this project, and the path model as shown in Exhibit A2.8 will appear in a new *Modeling window*.