Evaluation of the Structural Model

Contents

- 6.1 Assess Collinearity Issues
of the Structural Model – 117
- 6.2 Assess the Significance
and Relevance of the Structural
Model Relationships – 117
- 6.3 Assess the Model's Explanatory
Power – 118
- 6.4 Assess the Model's Predictive
Power – 119
- 6.5 Model Comparisons – 121
- 6.6 Case Study Illustration:
Structural Model
Evaluation – 123
- References – 136

Learning Objectives

After reading this chapter, you should understand:

1. The steps involved in structural model assessment
2. The concept of explanatory power and how to evaluate it
3. How to use $PLS_{predict}$ to assess a model's predictive power
4. The concept of model comparison and metrics for selecting the best model
5. How to assess structural models using SEMinR

6

Once you have confirmed that the measurement of constructs is reliable and valid, the next step addresses the assessment of the structural model results. ■ Figure 6.1 shows a systematic approach to the structural model assessment. In the first step, you need to examine the structural model for potential collinearity issues. The reason is that the estimation of path coefficients in the structural models is based on ordinary least squares (OLS) regressions of each endogenous construct on its corresponding predictor constructs. Just as in an OLS regression, the path coefficients might be biased if the estimation involves high levels of collinearity among predictor constructs. Once you have ensured that collinearity is not a problem, you will evaluate the significance and relevance of the structural model relationships (i.e., the path coefficients). Steps 3 and 4 of the procedure involve examining the model's explanatory power and predictive power. In addition, some research situations involve the computation and comparison of alternative models, which can emerge from different theories or contexts. PLS-SEM facilitates the comparison of alternative models using established criteria, which are well known from the regression literature. As model comparisons are not relevant for every PLS-SEM analysis, Step 5 should be considered optional.

■ **Fig. 6.1** Structural model assessment procedure. (Source: Hair, Hult, Ringle, & Sarstedt, 2022, Chap. 6; used with permission by Sage)

Excuse

Model fit indices enable judging how well a hypothesized model structure fits the empirical data and are an integral part of any CB-SEM analysis. However, the notion of model fit as known from CB-SEM is not transferrable to PLS-SEM as the method follows a different aim when estimating model parameters (i.e., with the aim of maximizing the explained variance instead of minimizing the divergence between covariance matrices) – see Hair, Sarstedt, and Ringle (2019). Nevertheless, research has brought forward several PLS-SEM-based model fit measures such as SRMR, RMS_{θ} , and the exact fit test (Henseler et al., 2014; Lohmöller, 1989, Chap. 2), which, however, have proven ineffective in detecting model misspecifications in settings commonly encountered in applied research. Instead, structural model assessment in PLS-SEM focuses on evaluating the model's explanatory and predictive power. For a detailed discussion of model fit in PLS-SEM, see Chap. 6 in Hair et al. (2022).

6.1 Assess Collinearity Issues of the Structural Model

Structural model coefficients for the relationships between constructs are derived from estimating a series of regression equations. As the point estimates and standard errors can be biased by strong correlations of each set of predictor constructs (Sarstedt & Mooi, 2019; Chap. 7), the structural model regressions must be examined for potential **collinearity** issues. This process is similar to assessing formative measurement models, but in this case, the construct scores of the predictor constructs in each regression in the structural model are used to calculate the **variance inflation factor (VIF)** values. VIF values above 5 are indicative of probable collinearity issues among predictor constructs, but collinearity can also occur at lower VIF values of 3–5 (Becker, Ringle, Sarstedt, & Völckner, 2015; Mason & Perreault, 1991). If collinearity is a problem, a frequently used option is to create higher-order constructs (Hair, Risher, Sarstedt, & Ringle, 2019; Hair, Sarstedt, Ringle, & Gudergan, 2018; Chap. 2; Sarstedt et al., 2019).

6.2 Assess the Significance and Relevance of the Structural Model Relationships

In the next step, the **significance of the path coefficients** and **relevance of the path coefficients** are evaluated. Analogous to the assessment of formative indicator weights (► Chap. 5), the significance assessment builds on bootstrapping standard errors as a basis for calculating *t*-values of path coefficients or alternatively confidence intervals (Streukens & Leroi-Werelds, 2016). A path coefficient is significant at the 5% level if the value zero does not fall into the 95% confidence interval. In general, the percentile method should be used to construct the confidence intervals (Aguirre-Urreta & Rönkkö, 2018). For a recap on bootstrapping, return to ► Chap. 5.

In terms of relevance, path coefficients are usually between -1 and $+1$, with coefficients closer to -1 representing strong negative relationships and those closer to $+1$ indicating strong positive relationships. Note that values below -1 and above $+1$ may technically occur, for instance, when collinearity is at very high levels. Path coefficients larger than ± 1 are not acceptable, and multicollinearity reduction methods must be implemented. As PLS-SEM processes standardized data, the path coefficients indicate the changes in an endogenous construct's values that are associated with standard deviation unit changes in a certain predictor construct, holding all other predictor constructs constant. For example, a path coefficient of 0.505 indicates that when the predictor construct increases by one standard deviation unit, the endogenous construct will increase by 0.505 standard deviation units.

The research context is important when determining whether the size of a path coefficient is meaningful. Thus, when examining structural model results, researchers should also interpret total effects, defined as the sum of the direct effect (if any) and all indirect effects linking one construct to another in the model. The examination of total effects between constructs, including all their indirect effects, provides a more comprehensive picture of the structural model relationships (Nitzl, Roldán, & Cepeda Carrión, 2016).

6.3 Assess the Model's Explanatory Power

The next step involves examining the **coefficient of determination (R^2)** of the endogenous construct(s). The R^2 represents the variance explained in each of the endogenous constructs and is a measure of the model's **explanatory power** (Shmueli & Koppius, 2011), also referred to as **in-sample predictive power** (Rigdon, 2012). The R^2 ranges from 0 to 1 , with higher values indicating a greater explanatory power. As a general guideline, R^2 values of 0.75 , 0.50 , and 0.25 can be considered substantial, moderate, and weak, respectively, in many social science disciplines (Hair, Ringle, & Sarstedt, 2011). But acceptable R^2 values are based on the research context, and in some disciplines, an R^2 value as low as 0.10 is considered satisfactory, as for example, in predicting stock returns (e.g., Raithel, Sarstedt, Scharf, & Schwaiger, 2012).

Researchers should also be aware that R^2 is a function of the number of predictor constructs – the greater the number of predictor constructs, the higher the R^2 . Therefore, the R^2 should always be interpreted relative to the context of the study, based on the R^2 values from related studies as well as models of similar complexity. R^2 values can also be too high when the model overfits the data. In case of model overfit, the (partial regression) model is too complex, which results in fitting the random noise inherent in the sample rather than reflecting the overall population. The same model would likely not fit on another sample drawn from the same population (Sharma, Sarstedt, Shmueli, Kim, & Thiele, 2019). When measuring a concept that is inherently predictable, such as physical processes, R^2 values of (up to) 0.90 might be plausible. However, similar R^2 value levels in a model that predicts human attitudes, perceptions, and intentions would likely

indicate model overfit (Hair, Risher, et al., 2019). A limitation of R^2 is that the metric will tend to increase as more explanatory variables are introduced to a model. The adjusted R^2 metric accounts for this by adjusting the R^2 value based upon the number of explanatory variables in relation to the data size and is seen as a more conservative estimate of R^2 (Theil, 1961). But because of the correction factor introduced to account for data and model size, the adjusted R^2 is not a precise indication of an endogenous construct's explained variance (Sarstedt & Mooi, 2019; Chap. 7).

Researchers can also assess how the removal of a selected predictor construct affects an endogenous construct's R^2 value. This metric is the f^2 effect size and is similar to the size of the path coefficients. More precisely, the rank order of the relevance of the predictor constructs in explaining a dependent construct in the structural model is often the same when comparing the size of the path coefficients and the f^2 effect sizes. In such situations, the f^2 effect size is typically only reported if requested by editors or reviewers. Otherwise (i.e., if the rank order of constructs' relevance in explaining a dependent construct in the structural model differs when comparing the size of the path coefficients and the f^2 effect sizes), the researcher may report the f^2 effect size to offer an alternative perspective on the results. In addition, some other research settings call for the reporting of effect sizes, such as in moderation analysis (Memon et al., 2019; see Chap. 7).

6.4 Assess the Model's Predictive Power

Many researchers interpret the R^2 statistic as a measure of their model's predictive power (Sarstedt & Danks, 2021; Shmueli & Koppius, 2011). This interpretation is not entirely correct, however, since the R^2 only indicates the model's in-sample explanatory power – it says nothing about the model's **predictive power** (Chin et al., 2020; Hair & Sarstedt, 2021), also referred to as **out-of-sample predictive power**, which indicates a model's ability to predict new or future observations. Addressing this concern, Shmueli, Ray, Estrada, and Chatla (2016) introduced PLS_{predict} , a procedure for out-of-sample prediction. Execution of PLS_{predict} involves estimating the model on a **training sample** and evaluating its predictive performance on a **holdout sample** (Shmueli et al., 2019). Note that the holdout sample is separated from the total sample before executing the initial analysis on the training sample data, so it includes data that were not used in the model estimation.

➤ Important

The R^2 is a measure of a model's explanatory power. It does not, however, indicate a model's out-of-sample predictive power.

PLS_{predict} executes **k -fold cross-validation**. A fold is a subgroup of the total sample, while k is the number of subgroups. That is, the total dataset is randomly split into k equally sized subsets of data. For example, a cross-validation based on $k = 5$ folds splits the sample into five equally sized data subsets (i.e., groups of data). PLS_{predict}

■ **Fig. 6.2** *k*-fold cross-validation procedure. (Source: authors' own figure)

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Holdout 1	Training	Training	Training	Training
Training	Holdout 2	Training	Training	Training
Training	Training	Holdout 3	Training	Training
Training	Training	Training	Holdout 4	Training
Training	Training	Training	Training	Holdout 5

6

then combines $k-1$ subsets (i.e., four groups of data) into a single training sample that is used to predict the remaining fifth subset. The fifth subset is the holdout sample for the first cross-validation run. This cross-validation process is then repeated k times (i.e., five times), with each of the five subsets used once as the holdout sample and the remaining cases being combined into the training sample. Thus, each case in every holdout sample has a predicted value estimated with the respective training sample in which that case was not used to estimate the model parameters. Leave-one-out cross-validation (LOOCV) is a subset of k -fold cross-validation where only one observation is included in the holdout sample.

■ Figure 6.2 visualizes the cross-validation process. Shmueli et al. (2019) recommend setting $k = 10$, but researchers need to make sure that the training sample for each fold meets minimum sample size guidelines (e.g., by following the inverse square root method, see also ► Chap. 1).

The generation of the k subsets of data is a random process and can therefore result in extreme partitions that potentially lead to abnormal solutions. To avoid such abnormal solutions, researchers should run PLS_{predict} multiple times. Shmueli et al. (2019) generally recommend running the procedure ten times. However, when the aim is to mimic how the PLS model will eventually be used to predict a new observation using a single model (estimated from the entire dataset), PLS_{predict} should be run only once (i.e., without repetitions).

To assess a model's predictive power, researchers can draw on several **prediction statistics** that quantify the amount of **prediction error** in the indicators of a particular endogenous construct. When analyzing the prediction errors, the focus should be on the model's key endogenous construct – and not on examining the prediction errors for the indicators of all endogenous constructs. The most popular metric to quantify the degree of prediction error is the **root-mean-square error (RMSE)**. This metric is the square root of the average of the squared differences between the predictions and the actual observations. Note that the RMSE squares the errors before averaging, so the statistic assigns a greater weight to larger errors, which makes it particularly useful when large errors are undesirable – as it is common in predictive analyses. Another popular metric is the **mean absolute error (MAE)**. This metric measures the average magnitude of the errors in a set of predictions without considering their direction (over- or underestimation). The MAE

thus is the average absolute difference between the predictions and the actual observations, with all the individual differences having equal weight.

In most instances, researchers should use the RMSE to examine a model's predictive power. But if the prediction error distribution is highly nonsymmetric, as evidenced in a long left or right tail in the distribution of prediction errors (Danks & Ray, 2018), the MAE is the more appropriate prediction statistic (Shmueli et al., 2019). These prediction statistics depend on the indicators' measurement scales, so the absolute size of their raw values does not have much meaning.

To interpret these metrics, researchers need to compare each indicator's RMSE (or MAE) values with a naïve **linear regression model (LM) benchmark**. The LM benchmark values are obtained by running a linear regression of each of the dependent construct's indicators on the indicators of the exogenous constructs in the PLS path model (Danks & Ray, 2018). In comparing the RMSE (or MAE) values with the LM values, the following guidelines apply (Shmueli et al., 2019):

1. If *all* indicators in the PLS-SEM analysis have lower RMSE (or MAE) values compared to the naïve LM benchmark, the model has high predictive power.
2. If the *majority* (or the same number) of indicators in the PLS-SEM analysis yields smaller prediction errors compared to the LM, this indicates a medium predictive power.
3. If a *minority* of the dependent construct's indicators produce lower PLS-SEM prediction errors compared to the naïve LM benchmark, this indicates the model has low predictive power.
4. If the PLS-SEM analysis (compared to the LM) yields lower prediction errors in terms of the RMSE (or the MAE) for *none* of the indicators, this indicates the model lacks predictive power.

An important decision when using $PLS_{predict}$ is how to generate the predictions when the PLS path model includes a mediator construct (mediation is discussed further in ► Chap. 7), which is both a predictor to the outcome and itself the outcome of an antecedent. SEMinR offers two alternatives to generate predictions in such a model setup (Shmueli et al., 2016). Researchers can choose to generate predictions using either the direct antecedents (DAs) or the earliest antecedents (EAs). In the DA approach, $PLS_{predict}$ would consider both the antecedent and the mediator as predictors of outcome constructs, whereas in the EA approach, the mediator would be excluded from the analysis. Danks (2021) presents simulation evidence that the DA approach generates predictions with the highest accuracy. We therefore recommend using this approach.

6.5 Model Comparisons

In a final, optional step, researchers may be interested in conducting **model comparisons**. Models are compared across different model configurations resulting from different theories or research contexts and are evaluated for predictive power.

Sharma et al. (2019, 2021) recently compared the efficacy of various metrics for model comparison tasks and found that Schwarz's (1978) **Bayesian information criterion (BIC)** and Geweke and Meese's (1981) **criterion (GM)** achieve a sound trade-off between model fit and predictive power in the estimation of PLS path models. These (information theoretic) model selection criteria facilitate the comparison of models in terms of model fit and predictive power without having to use a holdout sample, which is particularly useful for PLS-SEM analyses that often draw on small sample sizes. In applying these metrics, researchers should estimate each model separately and select the model that minimizes the value in BIC or GM for a certain target construct. That is, the model that produces the lowest value in BIC or GM is to be selected. While BIC and GM exhibit practically the same performance in model selection tasks, BIC is considerably easier to compute. Hence, our illustrations will focus on this criterion.

6

➤ Important

When comparing different models, researchers should select the one that minimizes the BIC value for a certain target construct.

One issue in the application of the BIC is that – in its simple form (i.e., raw values) – the criterion does not offer any insights regarding the relative weights of evidence in favor of models under consideration (Burnham & Anderson, 2004; Chap. 2.9). More precisely, while the differences in BIC and GM values are useful in ranking and selecting models, such differences can often be small in practice, leading to model selection uncertainty. To resolve this issue, researchers can use the BIC (and GM) values to compute **Akaike weights**. These weights determine a model's relative likelihood of being the data generation model, given the data and a set of competing models (Danks, Sharma, & Sarstedt, 2020) – see Wagenmakers and Farrell (2004) for a sample application. The higher the Akaike weights, the more likely that the selected model better represents the data generation model.

Excuse

A further development of the prediction-oriented model comparisons in PLS-SEM is the **cross-validated predictive ability test (CVPAT)** (Lienggaard et al., 2021). This approach offers a statistical test to decide whether an alternative model offers significantly higher out-of-sample predictive power than an established model. A statistical test is particularly advantageous if the differences in BIC values for deciding for one or the other model are relatively small. In addition, the test statistic of the CVPAT is suitable for prediction-oriented model comparison in the context of the development and validation of theories. As such, CVPAT offers researchers an important tool for selecting a model on which they can base, for example, strategic management and policy decisions. Future extensions of CVPAT will also support a test for the predictive power assessment of a single model (Hair, 2021).

■ **Table 6.1** Rules of thumb for structural model assessment (Source: authors' own table)

Criterion	Metrics and thresholds
Collinearity	Critical collinearity issues likely occur if $VIF \geq 5$ Collinearity issues are usually uncritical if $VIF = 3-5$ Collinearity is not a problematic issue if $VIF < 3$
Significance and relevance of the path coefficients	Apply bootstrapping to assess the significance of the path coefficients on the ground of t -values or confidence intervals Assess the magnitude of path coefficients Assess the f^2 values for each path and check that they follow the same rank order as the path coefficient magnitude
R ² value	R ² values of 0.75, 0.50, and 0.25 are considered substantial, moderate, and weak. However, R ² values have to be interpreted in the context of the model and its complexity. Excessive R ² values indicate that the model overfits the data
PLS _{predict}	Focus on one key target construct in the analysis Set $k = 10$, assuming each subgroup meets the minimum required sample size Use ten repetitions Compare the RMSE (or the MAE) values produced by PLS-SEM with those produced by the LM for each indicator. Check if the PLS-SEM analysis (compared to the LM) yields lower prediction errors in terms of RMSE (or MAE) for all (high predictive power), the majority or the same number (medium predictive power), the minority (low predictive power), or none of the indicators (no predictive power) Use the DA approach to generate predictions in mediation models
Model comparisons	Select the model that minimizes the value in BIC or GM compared to other models in the set Compute Akaike weights for additional evidence for a model's relative likelihood

■ Table 6.1 summarizes the metrics that need to be applied when evaluating the structural model.

6.6 Case Study Illustration: Structural Model Evaluation

We continue evaluating the extended corporate reputation model introduced in ► Chap. 5. In the prior chapters, we focused on the evaluation of the measurement models. We now turn our attention to the structural model, which describes the relationships between constructs.

```
# Load the SEMinR library
library(seminr)
# Load the data
corp_rep_data <- corp_rep_data
# Create measurement model
```

```

corp_rep_mm_ext <- constructs(
  composite("QUAL", multi_items("qual_", 1:8), weights = mode_B),
  composite("PERF", multi_items("perf_", 1:5), weights = mode_B),
  composite("CSOR", multi_items("csor_", 1:5), weights = mode_B),
  composite("ATTR", multi_items("attr_", 1:3), weights = mode_B),
  composite("COMP", multi_items("comp_", 1:3)),
  composite("LIKE", multi_items("like_", 1:3)),
  composite("CUSA", single_item("cusa")),
  composite("CUSL", multi_items("cusl_", 1:3))
)
# Create structural model
corp_rep_sm_ext <- relationships(
  paths(from = c("QUAL", "PERF", "CSOR", "ATTR"), to = c("COMP",
"LIKE")),
  paths(from = c("COMP", "LIKE"), to = c("CUSA", "CUSL")),
  paths(from = c("CUSA"), to = c("CUSL"))
)
# Estimate the model
corp_rep_pls_model_ext <- estimate_pls(
  data = corp_rep_data,
  measurement_model = corp_rep_mm_ext,
  structural_model = corp_rep_sm_ext,
  missing = mean_replacement,
  missing_value = "-99")
# Summarize the results of the model estimation
summary_corp_rep_ext <- summary(corp_rep_pls_model_ext)
# Bootstrap the model
boot_corp_rep_ext <- bootstrap_model(
  seminr_model = corp_rep_pls_model_ext,
  nboot = 1000,
  cores = parallel::detectCores(),
  seed = 123)
# Summarize the results of the bootstrap
summary_boot_corp_rep_ext <- summary(boot_corp_rep_ext,
                                     alpha = 0.05)

```

We follow the structural model assessment procedure (Step 1 in ■ Fig. 6.1) and begin with an evaluation of the collinearity of predictor constructs in relation to each endogenous construct. The corporate reputation model has four endogenous constructs (■ Fig. 6.3), namely, *COMP*, *LIKE*, *CUSA*, and *CUSL*. We examine the VIF values for the predictor constructs by inspecting the **vif_antecedents** element within the **summary_corp_rep_ext** object:

```

# Inspect the structural model collinearity VIF
summary_corp_rep_ext$vif_antecedents

```

As can be seen in ■ Fig. 6.4, all VIF values are clearly below the threshold of 5. However, *QUAL*'s VIF value (3.487) is above 3, suggesting the possibility of

■ **Fig. 6.3** The extended corporate reputation model. (Source: Hair et al., 2022, Chap. 5; used with permission by Sage)

```
Console Terminal x Jobs x
~/seminr/
> # Inspect the structural model collinearity VIF
> summary_corp_rep_ext$vif_antecedents
COMP :
  QUAL PERF  CSOR  ATTR
3.487 2.889 2.083 2.122

LIKE :
  QUAL PERF  CSOR  ATTR
3.487 2.889 2.083 2.122

CUSA :
  COMP  LIKE
1.686 1.686

CUSL :
  COMP  LIKE  CUSA
1.716 1.954 1.412
> |
```

■ **Fig. 6.4** VIF values for structural model. (Source: authors' screenshot from RStudio)

collinearity problems. Since the one exception is close to 3, we can conclude that collinearity among predictor constructs is likely not a critical issue in the structural model, and we can continue examining the result report.

Next, in the structural model assessment procedure (Step 2 in ■ Fig. 6.1), we need to evaluate the relevance and significance of the structural paths. The results of the bootstrapping of structural paths can be accessed by inspecting the `bootstrapped_paths` element nested in the `summary_boot_corp_ext` object (■ Fig. 6.5).

The screenshot shows an RStudio console window with the following content:

```
~/seminr/
> # Inspect the structural paths
> summary_boot_corp_rep_ext$bootstrapped_paths
```

		Original Est.	Bootstrap Mean	Bootstrap SD	T Stat.	2.5% CI	97.5% CI
QUAL	→ COMP	0.430	0.431	0.065	6.603	0.303	0.552
QUAL	→ LIKE	0.380	0.384	0.067	5.699	0.253	0.514
PERF	→ COMP	0.295	0.301	0.064	4.611	0.173	0.422
PERF	→ LIKE	0.117	0.123	0.073	1.613	-0.011	0.261
CSOR	→ COMP	0.059	0.059	0.054	1.084	-0.044	0.165
CSOR	→ LIKE	0.178	0.177	0.056	3.205	0.065	0.282
ATTR	→ COMP	0.086	0.084	0.055	1.565	-0.018	0.194
ATTR	→ LIKE	0.167	0.165	0.065	2.573	0.034	0.291
COMP	→ CUSA	0.146	0.147	0.071	2.047	0.007	0.281
COMP	→ CUSL	0.006	0.006	0.055	0.104	-0.104	0.115
LIKE	→ CUSA	0.436	0.435	0.062	7.069	0.312	0.555
LIKE	→ CUSL	0.344	0.343	0.056	6.175	0.231	0.449
CUSA	→ CUSL	0.505	0.505	0.042	12.074	0.420	0.586

■ Fig. 6.5 Path coefficient estimates, significance, and confidence intervals. (Source: authors' screenshot from RStudio)

```
# Inspect the structural paths
summary_boot_corp_rep_ext$bootstrapped_paths
# Inspect the total effects
summary_boot_corp_rep_ext$bootstrapped_total_paths
# Inspect the model RSquares
summary_corp_rep_ext$paths
# Inspect the effect sizes
summary_corp_rep_ext$fsquare
```

First, let's consider the original path coefficient estimates (column: **Original Est.**, ■ Fig. 6.5) for the exogenous driver constructs. For example, we find that *QUAL* has a strong positive impact on both *COMP* (**0.430**) and *LIKE* (**0.380**). A similar pattern of relationships emerges for *PERF*, but with lower effect sizes. On the contrary, *CSOR* exerts a much lower impact on these two constructs as evidenced in path coefficient estimates of **0.059** for *COMP* and **0.178** for *LIKE*. Similarly, *ATTR* has only a low impact on *COMP* and *LIKE*. Further analyzing the path coefficient estimates, we find that *LIKE* is the primary driver of both *CUSA* (**0.436**) and *CUSL* (**0.344**), as demonstrated by the larger path coefficients compared with those of *COMP*.

Let's now review the results for statistical significance. Assuming a 5% significance level (as specified with the parameter **alpha = 0.05** in the **bootstrap_model()** function), the *t*-values (T Stat. column, ■ Fig. 6.5) estimated from the bootstrapping should exceed the value of 1.960. We find that several relationships are significant, including five of the exogenous driver construct relationships (*QUAL* → *COMP*, *t* = **6.603**; *QUAL* → *LIKE*, *t* = **5.699**; *PERF* → *COMP*, *t* = **4.611**; *CSOR* → *LIKE*, *t* = **3.205**; *ATTR* → *LIKE*, *t* = **2.573**). At the same time, however, three of the exogenous driver relationships are not statistically significant (*PERF* → *LIKE*,

$t = 1.613$; $CSOR \rightarrow COMP$, $t = 1.084$; $ATTR \rightarrow COMP$, $t = 1.565$). Reviewing the statistical significance of the path coefficients for the endogenous construct relationships, we see that four of the five paths are statistically significant. The one path that is not significant is from $COMP$ to $CUSL$ ($t = 0.104$). These results suggest that companies should concentrate their marketing efforts on enhancing their likeability (by strengthening customers' quality perceptions) rather than their perceived competence to maximize customer loyalty. This is not surprising, considering that customers rated mobile network operators. Since their services (provision of network capabilities) are intangible, affective judgments play a much more important role than cognitive judgments for establishing customer loyalty.

➤ Note that, although a seed is set, the random results generated might differ across software and hardware combinations. Hence, your results will likely look slightly different from those in ■ Fig. 6.5.

It is also important to consider the total effects to gain an idea of the impact of the four exogenous driver constructs on the outcome constructs $CUSA$ and $CUSL$. To evaluate the total effects, we need to inspect the **bootstrapped_total_paths** element of **summary_boot_corp_rep_ext** (■ Fig. 6.6).

Of the four driver constructs, $QUAL$ has the strongest total effect on $CUSL$ (0.248), followed by $CSOR$ (0.105), $ATTR$ (0.101), and $PERF$ (0.089). Therefore,

		Original Est.	Bootstrap Mean	Bootstrap SD	T Stat.	2.5% CI	97.5% CI
QUAL	-> COMP	0.430	0.431	0.065	6.603	0.303	0.552
QUAL	-> LIKE	0.380	0.384	0.067	5.699	0.253	0.514
QUAL	-> CUSA	0.228	0.230	0.039	5.923	0.154	0.310
QUAL	-> CUSL	0.248	0.251	0.044	5.679	0.165	0.337
PERF	-> COMP	0.295	0.301	0.064	4.611	0.173	0.422
PERF	-> LIKE	0.117	0.123	0.073	1.613	-0.011	0.261
PERF	-> CUSA	0.094	0.098	0.040	2.373	0.024	0.179
PERF	-> CUSL	0.089	0.094	0.045	1.968	0.009	0.180
CSOR	-> COMP	0.059	0.059	0.054	1.084	-0.044	0.165
CSOR	-> LIKE	0.178	0.177	0.056	3.205	0.065	0.282
CSOR	-> CUSA	0.086	0.086	0.028	3.133	0.031	0.138
CSOR	-> CUSL	0.105	0.105	0.033	3.166	0.038	0.172
ATTR	-> COMP	0.086	0.084	0.055	1.565	-0.018	0.194
ATTR	-> LIKE	0.167	0.165	0.065	2.573	0.034	0.291
ATTR	-> CUSA	0.085	0.084	0.031	2.731	0.026	0.147
ATTR	-> CUSL	0.101	0.100	0.038	2.652	0.028	0.178
COMP	-> CUSA	0.146	0.147	0.071	2.047	0.007	0.281
COMP	-> CUSL	0.079	0.081	0.069	1.155	-0.052	0.213
LIKE	-> CUSA	0.436	0.435	0.062	7.069	0.312	0.555
LIKE	-> CUSL	0.564	0.563	0.061	9.219	0.444	0.676
CUSA	-> CUSL	0.505	0.505	0.042	12.074	0.420	0.586

■ Fig. 6.6 Total effect estimates, significance, and confidence intervals. (Source: authors' screenshot from RStudio)

it is advisable for companies to focus on marketing activities that positively influence the customers' perception of the quality of their products and services (*QUAL*). As can be seen, all of *QUAL*'s total effects are significant at a 5% level.

In Step 3 of the structural model assessment procedure (■ Fig. 6.1), we need to consider the model's explanatory power by analyzing the R^2 of the endogenous constructs and the f^2 effect size of the predictor constructs. To start with, we need to examine the R^2 values of the endogenous constructs. The R^2 values of *COMP* (0.631), *CUSL* (0.562), and *LIKE* (0.558) can be considered moderate, whereas the R^2 value of *CUSA* (0.292) is weak (■ Fig. 6.7). The weak R^2 value of *CUSA* may be the result of this construct being measured as a single item. We recommend customer satisfaction always be measured as a multi-item construct.

■ Figure 6.8 shows the f^2 values for all combinations of endogenous constructs (represented by the columns) and corresponding exogenous (i.e., predictor) constructs (represented by the rows). For example, *LIKE* has a medium effect size of 0.159 on *CUSA* and of 0.138 on *CUSL*. On the contrary, *COMP* has no effect

6

■ Fig. 6.7 Path coefficient estimates, R^2 , and adjusted R^2 values. (Source: authors' screenshot from RStudio)

■ Fig. 6.8 f^2 effect sizes. (Source: authors' screenshot from RStudio)

■ **Table 6.2** A list of arguments for the `predict_pls()` function

Argument	Value
<i>model</i>	The PLS model, which contains the structural and measurement models used to generate predictions
<i>technique</i>	The <code>predict_DA</code> option for direct antecedent (DA) approach or <code>predict_EA</code> for the earliest antecedent (EA) approach (<code>predict_DA</code> is set as default)
<i>noFolds</i>	Number of folds to employ in the <i>k</i> -fold process, <code>NULL</code> is default whereby leave-one-out cross-validation (LOOCV) is performed
<i>reps</i>	Number of replications to run (<code>NULL</code> is default whereby the <i>k</i> -fold cross-validation is performed once)
<i>cores</i>	The number of cores to use for parallel processing

Source: authors' own table

on *CUSA* (0.018) or *CUSL* (0.000). The rank order of effect sizes is identical to the rank order on the grounds of the path coefficients.

Step 4 in the structural model assessment procedure (■ Fig. 6.1) is the evaluation of the model's predictive power. To do so, we first have to generate the predictions using the `predict_pls()` function. ■ Table 6.2 lists this function's arguments.

We run the PLS_{predict} procedure with *k* = 10 folds and ten repetitions and thus set *noFolds* = 10, and *reps* = 10. In addition, we use the `predict_DA` approach. Finally, we summarize the PLS_{predict} model and assign the output to the `sum_predict_corp_rep_ext` object:

```
# Generate the model predictions
predict_corp_rep_ext <- predict_pls(
  model = corp_rep_pls_model_ext,
  technique = predict_DA,
  noFolds = 10,
  reps = 10)
# Summarize the prediction results
sum_predict_corp_rep_ext <- summary(predict_corp_rep_ext)
```

The distributions of the prediction errors need to be assessed to decide the best metric for evaluating predictive power. If the prediction error is highly skewed, the MAE is a more appropriate metric than the RMSE. In order to assess the distribution of predictive error, we use the `plot()` function on the object `sum_predict_corp_rep_ext` and specify the *indicator* argument to the indicators of interest. We should focus on the key outcome construct *CUSL* and evaluate the indicators *cusl_1*, *cusl_2*, and *cusl_3*. First, we set the number of plots to display in the output to three plots arranged horizontally using the `par(mfrow=c(1, 3))`

command. Remember to set `par(mfrow=c(1,1))` after outputting the plots; otherwise, all future plots will be arranged horizontally in a sequence of three:

```
# Analyze the distribution of prediction error
par(mfrow=c(1,3))
plot(sum_predict_corp_rep_ext,
     indicator = "cusl_1")
plot(sum_predict_corp_rep_ext,
     indicator = "cusl_2")
plot(sum_predict_corp_rep_ext,
     indicator = "cusl_3")
par(mfrow=c(1,1))
```

6

The results in ■ Fig. 6.9 show that while all three plots have a left tail and are mildly skewed to the right, the prediction error distributions are rather symmetric. We should therefore use the RMSE for our assessment of prediction errors.

We can investigate the RMSE and MAE values by calling the `sum_predict_corp_rep_ext` object.

```
# Compute the prediction statistics
sum_predict_corp_rep_ext
```

Analyzing the *CUSL* construct's indicators (■ Fig. 6.10), we find that the PLS path model has lower out-of-sample predictive error (RMSE) compared to the naïve LM model benchmark for all three indicators (sections: **PLS out-of-sample metrics** and **LM out-of-sample metrics**): *cusl_1* (PLS, 1.192; LM, 1.228), *cusl_2* (PLS, 1.239; LM, 1.312), and *cusl_3* (PLS, 1.312; LM, 1.380). Accordingly, we conclude that the model has a high predictive power.

In Step 5 of the structural model assessment procedure (■ Fig. 6.1), we will perform model comparisons. First, we set up three theoretically justifiable competing models (Model 1, Model 2, and Model 3, shown in ■ Fig. 6.11). Specifically, we compare the original model that serves as the basis for our prior analyses (Model 1), with two more complex versions, in which the four driver constructs also relate to *CUSA* (Model 2) and *CUSL* (Model 3). As the models share the same measurement models, we need to specify them only once. Since each model has a unique structural model, we must specify three structural models according to ■ Fig. 6.11. To begin, we assign the outputs to `structural_model1`, `structural_model2`, and `structural_model3`. We can then estimate three separate PLS path models and summarize the results.

■ **Fig. 6.9** Distribution of prediction error for indicators *cusl_1*, *cusl_2*, and *cusl_3*. (Source: authors' screenshot from RStudio)

```
# Estimate alternative models
# Create measurement model
measurement_model <- constructs(
  composite("QUAL", multi_items("qual_", 1:8), weights = mode_B),
  composite("PERF", multi_items("perf_", 1:5), weights = mode_B),
  composite("CSOR", multi_items("csor_", 1:5), weights = mode_B),
  composite("ATTR", multi_items("attr_", 1:3), weights = mode_B),
  composite("COMP", multi_items("comp_", 1:3)),
  composite("LIKE", multi_items("like_", 1:3)),
  composite("CUSA", single_item("cusa")),
  composite("CUSL", multi_items("cusl_", 1:3))
)
```

```

# Create structural models
# Model 1
structural_model1 <- relationships(
  paths(from = c("QUAL", "PERF", "CSOR", "ATTR"), to = c("COMP",
    "LIKE")),
  paths(from = c("COMP", "LIKE"), to = c("CUSA", "CUSL")),
  paths(from = "CUSA", to = c("CUSL"))
)
# Model 2
structural_model2 <- relationships(
  paths(from = c("QUAL", "PERF", "CSOR", "ATTR"), to = c("COMP",
    "LIKE", "CUSA")),
  paths(from = c("COMP", "LIKE"), to = c("CUSA", "CUSL")),
  paths(from = "CUSA", to = c("CUSL"))
)
# Model 3
structural_model3 <- relationships(
  paths(from = c("QUAL", "PERF", "CSOR", "ATTR"),
    to = c("COMP", "LIKE", "CUSA", "CUSL")),
  paths(from = c("COMP", "LIKE"), to = c("CUSA", "CUSL")),
  paths(from = "CUSA", to = c("CUSL"))
)

# Estimate and summarize the models
pls_model1 <- estimate_pls(
  data = corp_rep_data,
  measurement_model = measurement_model,
  structural_model = structural_model1,
  missing_value = "-99"
)
sum_model1 <- summary(pls_model1)

pls_model2 <- estimate_pls(
  data = corp_rep_data,
  measurement_model = measurement_model,
  structural_model = structural_model2,
  missing_value = "-99"
)
sum_model2 <- summary(pls_model2)

pls_model3 <- estimate_pls(
  data = corp_rep_data,
  measurement_model = measurement_model,
  structural_model = structural_model3,
  missing_value = "-99"
)
sum_model3 <- summary(pls_model3)

```

```

Console Terminal Find in Files Jobs
~/Dropbox/GITrepos/seminr/
> # Generate the model predictions
> predict_corp_rep_ext <- predict_pls(
+   model = corp_rep_pls_model_ext,
+   technique = predict_DA,
+   noFolds = 10,
+   reps = 10)
> # Summarize the prediction results
> sum_predict_corp_rep_ext <- summary(predict_corp_rep_ext)
> # Analyze the distribution of prediction error
> par(mfrow=c(1,3))
> plot(sum_predict_corp_rep_ext, indicator = "cusl_1")
> plot(sum_predict_corp_rep_ext, indicator = "cusl_2")
> plot(sum_predict_corp_rep_ext, indicator = "cusl_3")
> par(mfrow=c(1,1))
> # Inspect the results of PLSpredict
> sum_predict_corp_rep_ext

PLS in-sample metrics:
      comp_1 comp_2 comp_3 like_1 like_2 like_3 cusa cusl_1 cusl_2 cusl_3
RMSE  1.022  1.081  1.112  1.101  1.452  1.478  0.985  1.181  1.225  1.300
MAE   0.784  0.868  0.881  0.837  1.127  1.146  0.769  0.874  0.909  0.948

PLS out-of-sample metrics:
      comp_1 comp_2 comp_3 like_1 like_2 like_3 cusa cusl_1 cusl_2 cusl_3
RMSE  1.046  1.101  1.134  1.136  1.489  1.512  1.000  1.192  1.239  1.312
MAE   0.799  0.882  0.895  0.861  1.155  1.175  0.778  0.883  0.919  0.959

LM in-sample metrics:
      comp_1 comp_2 comp_3 like_1 like_2 like_3 cusa cusl_1 cusl_2 cusl_3
RMSE  0.938  1.011  1.011  0.968  1.315  1.346  0.770  1.086  1.161  1.222
MAE   0.717  0.799  0.788  0.744  1.020  1.050  0.609  0.794  0.873  0.911

LM out-of-sample metrics:
      comp_1 comp_2 comp_3 like_1 like_2 like_3 cusa cusl_1 cusl_2 cusl_3
RMSE  1.064  1.142  1.134  1.121  1.520  1.555  0.886  1.228  1.312  1.380
MAE   0.805  0.898  0.877  0.849  1.154  1.196  0.690  0.889  0.977  1.029
>

```

■ **Fig. 6.10** Prediction metrics for outcome construct items. (Source: authors' screenshot from RStudio)

We focus our analysis on the *CUSA* construct as the immediate consequence of the two dimensions of corporate reputation (*LIKE* and *COMP*). In order to compare the models, we must first inspect each model for the estimated BIC value for the outcome construct of interest (i.e., *CUSA*). The matrix of the model's information criteria can be accessed by inspecting the `it_criteria` element in the `sum_model1` object, `sum_model1$it_criteria`. This matrix reports the BIC value for each outcome construct along with the Akaike information criterion (AIC), which is known to favor too complex models (Sharma et al., 2019, 2021). We can subset this matrix to return only the BIC row of the *CUSA* column by entering `sum_model1$it_criteria["BIC", "CUSA"]`. To compare the BIC for the

■ **Fig. 6.11** Three competing models for predictive model comparison. (Source: Hair et al., 2022, Chap. 6; used with permission by Sage)

three models, we need to assign the BIC for *CUSA* for each model to a vector. We then name the vector using the `names()` function and inspect the `itcriteria_vector`. In a final step, we request the BIC Akaike weights for the three models under consideration using the `compute_it_criteria_weights()` function.

```
# Inspect the IT Criteria matrix of Model1
sum_model1$it_criteria
# Subset the matrix to only return the BIC row and CUSL column
sum_model1$it_criteria["BIC", "CUSA"]
# Collect the vector of BIC values for CUSL
itcriteria_vector <- c(sum_model1$it_criteria["BIC", "CUSA"],
                      sum_model2$it_criteria["BIC", "CUSA"],
                      sum_model3$it_criteria["BIC", "CUSA"])
# Assign the model names to IT Criteria vector
names(itcriteria_vector) <- c("Model1", "Model2", "Model3")
# Inspect the IT Criteria vector for competing models
itcriteria_vector
# Calculate the model BIC Akaike weights
compute_itcriteria_weights(itcriteria_vector)
```

We can now compare the BIC values (■ Fig. 6.12) of Model 1 (-102.206), Model 2 (-93.965), and Model 3 (-97.401). The results suggest that Model 1 is superior in terms of model fit. To learn about the models' relative likelihoods, we can consult the BIC-based Akaike weights for Model 1 (0.904), Model 2 (0.015), and Model 3 (0.082). It is clear that Model 1 has a very strong weighting, so we conclude the model comparison indicates Model 1 is the superior model.

```

Console Terminal x Jobs x
~/seminr/
> # Inspect the IT Criteria matrix of Model1
> sum_model1$it_criteria
      COMP      LIKE      CUSA      CUSL
AIC -333.825 -271.581 -113.728 -276.964
BIC -314.622 -252.378 -102.206 -261.602
> # Subset the matrix to only return the BIC row and CUSL column
> sum_model1$it_criteria["BIC", "CUSA"]
[1] -102.2062
> # Collect the vector of BIC values for CUSL
> itcriteria_vector <- c(sum_model1$it_criteria["BIC", "CUSA"],
+                        sum_model2$it_criteria["BIC", "CUSA"],
+                        sum_model3$it_criteria["BIC", "CUSA"])
> # Assign the model names to IT Criteria vector
> names(itcriteria_vector) <- c("Model1", "Model2", "Model3")
> # Inspect the IT Criteria vector for competing models.
> itcriteria_vector
      Model1      Model2      Model3
-102.20623  -93.96473  -97.40109
> # Calculate the model BIC Akaike Weights
> compute_itcriteria_weights(itcriteria_vector)
      Model1      Model2      Model3
0.90357310 0.01466713 0.08175977
>

```

■ **Fig. 6.12** BIC values and BIC Akaike weights for *CUSA*. (Source: authors' screenshot from RStudio)

Summary

The structural model assessment in PLS-SEM starts with the evaluation of potential collinearity among predictor constructs in structural model regressions, followed by the evaluation of the path coefficients' significance and relevance and concluding with the analysis of the model's explanatory and predictive power. After ensuring the model estimates are not affected by high levels of collinearity by examining VIF values, we need to test the path coefficients' significance by applying the bootstrapping routine and examining *t*-values or bootstrapping confidence intervals. To assess a model's explanatory power, researchers rely on the coefficient of determination (R^2). Predictive power assessment builds on PLS_{predict}, a holdout sample-based procedure that applies *k*-fold cross-validation to estimate the model parameters. Some research situations call for the comparison of alternative models. To compare different model configurations and select the best model, the BIC criterion should be used. The model, which yields the smallest BIC value, is considered the best model in the set. BIC-based Akaike weights offer further evidence for the relative likelihood of a model compared to alternative models in the set.

Exercise

We continue evaluating the influencer model introduced in the exercise section of ► Chap. 3 (► Fig. 3.10; ► Tables 3.9 and 3.10) and subsequently evaluated in the follow-up chapters. To start the exercise, load the influencer data, reproduce the influencer model in the SEMinR syntax, and estimate the model. We have evaluated both reflectively and formatively measured constructs in ► Chaps. 4 and 5, so we can now turn our attention to the structural model evaluation as follows:

1. Do any predictor constructs suffer from collinearity issues?
2. Are all structural paths significant and relevant? Which paths are of low or weak relevance?
3. Now focus on the key target construct *PI* as follows:
 - (a) Does the model have satisfactory explanatory power in terms of this construct?
 - (b) Does the model have satisfactory predictive power in terms of this construct?
 - (c) Construct a theoretically justified competing model and conduct a model comparison in order to detect if the original influencer model is supported by the BIC statistic. Can you generate a model with a higher BIC-based Akaike weight than the original influencer model?

References

- Aguirre-Urreta, M. I., & Rönkkö, M. (2018). Statistical inference with PLSc using bootstrap confidence intervals. *MIS Quarterly*, 42(3), 1001–1020.
- Becker, J.-M., Ringle, C. M., Sarstedt, M., & Völckner, F. (2015). How collinearity affects mixture regression results. *Marketing Letters*, 26(4), 643–659.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Research Methods*, 33(2), 261–304.
- Chin, W. W., Cheah, J.-H., Liu, Y., Ting, H., Lim, X.-J., & Cham, T. H. (2020). Demystifying the role of causal-predictive modeling using partial least squares structural equation modeling in information systems research. *Industrial Management & Data Systems*, 120(12), 2161–2209.
- Danks, N. P., & Ray, S. (2018). Predictions from partial least squares models. In F. Ali, S. M. Rasoolimanesh, & C. Cobanoglu (Eds.), *Applying partial least squares in tourism and hospitality research* (pp. 35–52). Bingley: Emerald.
- Danks, N. P., Sharma, P. N., & Sarstedt, M. (2020). Model selection uncertainty and multimodel inference in partial least squares structural equation modeling (PLS-SEM). *Journal of Business Research*, 113, 13–24.
- Danks, N.P. (2021). The piggy in the middle: The role of mediators in PLS-SEM-based prediction. *The Data Base for Advances in Information Systems*, forthcoming.
- Geweke, J., & Meese, R. (1981). Estimating regression models of finite but unknown order. *International Economic Review*, 22(1), 55–70.
- Hair, J. F. (2021). Next generation prediction metrics for composite-based PLS-SEM. *Industrial Management & Data Systems*, 121(1), 5–11.
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, 19(2), 139–151.
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24.

References

- Hair, J. F., & Sarstedt, M. (2021). Explanation plus prediction – The logical focus of project management research. *Project Management Journal*, 52(4), 319–322.
- Hair, J. F., Sarstedt, M., & Ringle, C. M. (2019). Rethinking some of the rethinking of partial least squares. *European Journal of Marketing*, 53(4), 566–584.
- Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2018). *Advanced issues in partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.
- Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., et al. (2014). Common beliefs and reality about partial least squares: Comments on Rönkkö & Evermann (2013). *Organizational Research Methods*, 17(1), 182–209.
- Liengaard, B., Sharma, P. N., Hult, G. T. M., Jensen, M. B., Sarstedt, M., Hair, J. F., & Ringle, C. M. (2021). Prediction: Coveted, yet forsaken? Introducing a cross-validated predictive ability test in partial least squares path modeling. *Decision Sciences*, 52(2), 362–392.
- Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Heidelberg: Physica.
- Mason, C. H., & Perreault, W. D. (1991). Collinearity, power, and interpretation of multiple regression analysis. *Journal of Marketing Research*, 28(3), 268–280.
- Memon, M. A., Cheah, J. H., Ramayah, T., Ting, H., Chuah, F., & Cham, T. H. (2019). Moderation analysis: Issues and guidelines. *Journal of Applied Structural Equation Modeling*, 3(1), i–xi.
- Nitzl, C., Roldán, J. L., & Cepeda Carrión, G. (2016). Mediation analysis in partial least squares path modeling. *Industrial Management & Data Systems*, 119(9), 1849–1864.
- Raithel, S., Sarstedt, M., Scharf, S., & Schwaiger, M. (2012). On the value relevance of customer satisfaction. Multiple drivers and multiple markets. *Journal of the Academy of Marketing Science*, 40(4), 509–525.
- Rigdon, E. E. (2012). Rethinking partial least squares path modeling: In praise of simple methods. *Long Range Planning*, 45(5–6), 341–358.
- Sarstedt, M. & Danks, N. P. (2021). Prediction in HRM research—A gap between rhetoric and reality. *Human Resource Management Journal*, forthcoming.
- Sarstedt, M., Hair, J. F., Cheah, J.-H., Becker, J.-M., & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order models. *Australasian Marketing Journal*, 27(3), 197–211.
- Sarstedt, M., & Mooi, E. (2019). *A concise guide to market research. The process, data, and methods using IBM SPSS statistics* (3rd ed.). Berlin: Springer.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464.
- Sharma, P. N., Sarstedt, M., Shmueli, G., Kim, K. H., & Thiele, K. O. (2019). PLS-based model selection: The role of alternative explanations in Information Systems research. *Journal of the Association for Information Systems*, 40(4), 346–397.
- Sharma, P. N., Shmueli, G., Sarstedt, M., Danks, N., & Ray, S. (2021). Prediction-oriented model selection in partial least squares path modeling. *Decision Sciences*, 52(3), 567–607.
- Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *MIS Quarterly*, 35(3), 553–572.
- Shmueli, G., Ray, S., Estrada, J. M. V., & Chatla, S. B. (2016). The elephant in the room: Predictive performance of PLS models. *Journal of Business Research*, 69(10), 4552–4564.
- Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J. H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing*, 53(11), 2322–2347.
- Streukens, S., & Leroi-Werelds, S. (2016). Bootstrapping and PLS-SEM: A step-by-step guide to get more out of your bootstrap results. *European Management Journal*, 34(6), 618–632.
- Theil, H. (1961). *Economic Forecasts and Policy*. Amsterdam: North-Holland.
- Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, 11(1), 192–196.

Suggested Reading

- Chin, W. W., Cheah, J.-H., Liu, Y., Ting, H., Lim, X.-J., & Cham, T. H. (2020). Demystifying the role of causal-predictive modeling using partial least squares structural equation modeling in information systems research. *Industrial Management & Data Systems*, 120(12), 2161–2209.
- Hair, J. F., Hult, T., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Thousand Oaks, CA: Sage.
- Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J. H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing*, 53(11), 2322–2347.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.