JMP Academic Case Study 058

Modeling Online Shopping Perceptions

Confirmatory Factor Analysis (CFA),
Structural Equation Modeling (SEM), Model Comparison

Produced by

Muralidhara A
muralidhara.a@jmp.com

Ross Metusalem
ross.metusalem@jmp.com

JMP STATISTICAL DISCOVERY

Modeling Online Shopping Behavior

Confirmatory Factor Analysis (CFA), Structural Equation Modeling (SEM), Measurement and Structural Regression Models, Model Comparison

Key Ideas

This case study demonstrates the use of confirmatory factor analysis (CFA) and structural equation modeling (SEM) as applied to survey responses from online shoppers. The case begins by using CFA to assess a measurement model and then proceeds to build a structural regression model to test a theory regarding the factors that influence shoppers' intent to purchase from an online seller.

Background

Online shopping has influenced the word of marketing significantly. Various online retailers across the globe sell a multitude of products and services, and consumers often prefer the convenience of shopping from home. While it gives an opportunity to shop 24/7, online shopping also exposes customers to increased risk. They can fall victim to fraud through fake websites set up to steal personal information or through the theft of their information from a seller's database. Even without fraud, the relatively lower overhead of setting up an online store gives inexperienced or incompetent sellers easy access to the marketplace, presenting shoppers with the risk of incorrect, incomplete, or lost orders. Therefore, to promote sales, online sellers have a strong incentive to encourage customers to view them as trustworthy.

The Task

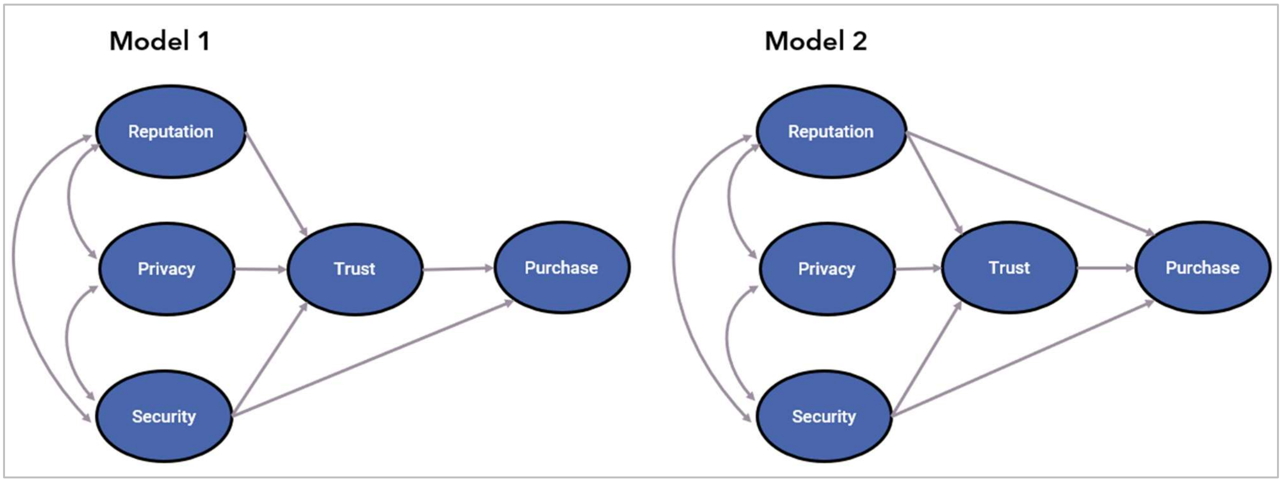
Anna, a marketing research analyst, is studying how different aspects of shoppers' trust in an online seller influence their intent to purchase a product from that seller. Through prior knowledge and research literature, she has two competing models that she wants to test. In the first model, she theorizes that shoppers' intent to purchase (Purchase) is driven in part by trust in the seller (Trust), which itself is driven by shoppers' perceptions of their personal privacy during the shopping experience (Privacy), the reputation of the seller (Reputation), and the perceived security of the online shopping experience (Security). In addition, Security is proposed to directly influence Purchase, because there may be elements of perceived security unrelated to trust in the seller (e.g., the shopper is using public Wi-Fi) that influence a shopper's intent to purchase. However, an alternative model proposes that Reputation also has a direct effect on Purchase because, theoretically, a positive perception of a seller's reputation may increase intent to purchase without necessarily causing increased trust in that seller.

Here, Purchase, Trust, Security, Reputation, and Privacy are considered latent variables or constructs, meaning that they cannot be observed directly and must be inferred from objective measurements.

Anna's competing theories are presented in more detail by the path diagrams in Exhibit 1. Both make testable claims regarding the influence of Privacy, Reputation, Security, and Trust on Purchase. In Model 1, Privacy, Reputation, and Security are all proposed to influence Trust, which in turn influences Purchase. In addition, Security is proposed to influence Purchase directly. Finally, Privacy, Reputation, and Security are all proposed to covary with one another. Model 2 includes all the effects of Model 1 while adding a direct effect of Reputation on Purchase.

Anna has collected survey data with the goal of testing this theory. Respondents rated a series of statements regarding a recent online shopping experience. For each of the five latent variables in her theory, Anna asked respondents to rate between three and five individual statements on a 0-5 scale. For example, one of the Privacy items stated, "I was not asked to provide any unnecessary personal information." A response of 5 indicated strong agreement.

Exhibit 1 Structures of the competing theoretical models.



Anna performs the following tasks using structural equation modeling (SEM):

- Conduct confirmatory factor analysis (CFA) to validate the measurement model.
- Build competing structural equation models to test the theory.
- Compare the models to identify which better fits the data.

The Data

Each of the latent variables was measured by a set of observed or indicator variables. Anna collected data from 843 respondents, which is provided in **online-purchase.jmp**. All indicator variables are measured using a 0-5 scale.

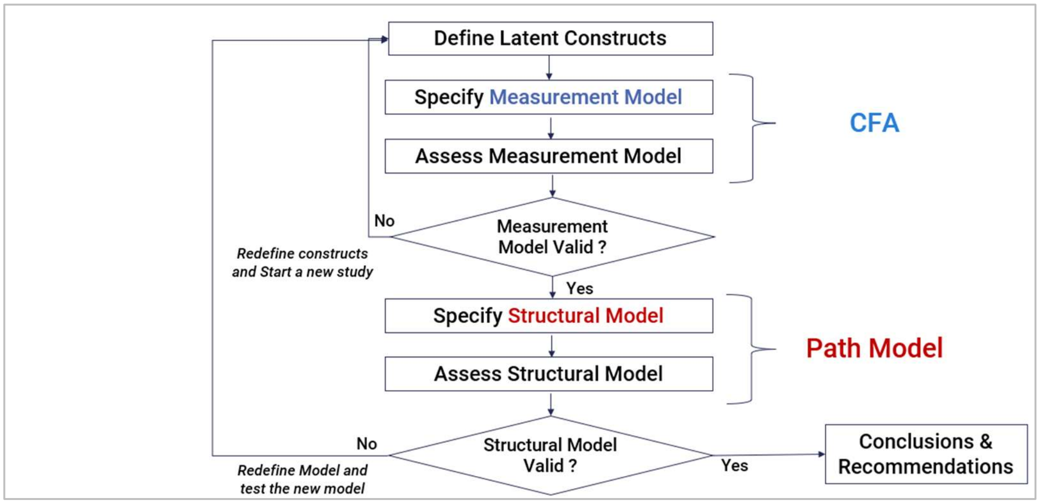
Variable	Description
ID	Respondent ID
Privacy_1	Indicator variables for Privacy
Privacy_2	
Privacy_3	
Privacy_4	
Reput_1	Indicator variables for Reputation
Reput_2	
Reput_3	
Reput_4	
Security_1	Indicator variables for Security
Security_2	
Security_3	
Security_4	
Security_5	
Trust_1	Indicator variables for Trust
Trust_2	
Trust_3	
Trust_4	
PI_1	Indicator variables for Purchase
PI_2	
PI_3	

Analysis

Fitting Structural Equation Models

Summary of the SEM process, which involves estimation of multiple models is shown below.

Exhibit 2 SEM process steps



Measurement models represent the theories that specify the observed variables for each latent variable. Confirmatory factor analysis (CFA) is a technique used to estimate a measurement model. It involves the confirmation of the number of latent variables and the loadings of observed variables on them. Reliability and validity of the constructs are assessed as part CFA.

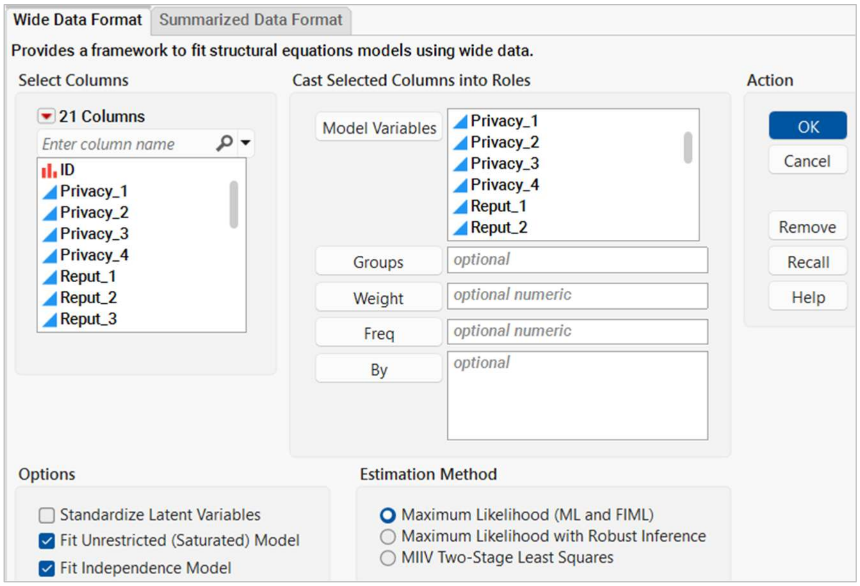
Structural regression models represent theories that specify how the latent variables are related to each other. The models are assessed for their ability to account for the correlational structure of the observed variables. A high degree of fit is taken as evidence that the theory accounts for the data well.

There are two approaches to build SEM models: covariance-based (CB-SEM) and component-based (PLS-SEM). JMP leverages the covariance-based method, where the goal is to fit a model that explains the observed covariation between variables. Competing models, representing alternative theories about the data, are often specified and compared statistically to determine which describes the data best.

Measurement Model Estimation (Confirmatory Factor Analysis)

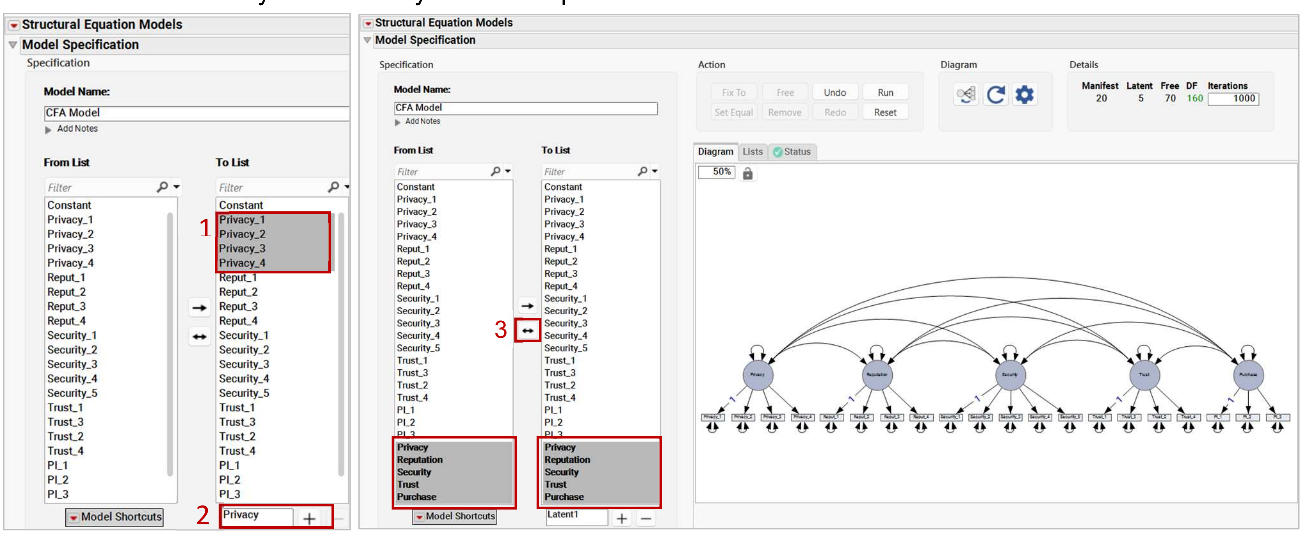
CFA enables you to test alternative measurement models and is the initial step prior to fitting structural regression models. In this example, we assess measurement validity and reliability by performing confirmatory factor analysis for the five latent variables: Privacy, Security, Reputation, Trust, and Purchase. Exhibits 3 and 4 show how to enter the survey data into the Structural Equation Models platform and specify the confirmatory factor model.

Exhibit 3 Structural Equation Modeling dialog



(Analyze > Multivariate Methods > Structural Equation Models. Select all the indicator variables starting from Privacy_1 to PI_3 and move them to Model Variables > Click OK.)

Exhibit 4 Confirmatory Factor Analysis model specification



(To specify the latent variables, specify each construct using their corresponding indicator variables. Select all the indicator variables related to Privacy (Privacy_1, Privacy_2, Privacy_3 and Privacy_4) in the To List (red box 1), label the latent variable as Privacy, and click the plus sign at the bottom (red box 2). This adds a latent variable called Privacy and populates it into both To List and From List. Follow similar steps to add all other latent variables: Reputation, Security, Trust, and Purchase.

After adding all the constructs, select all the constructs in the From List and To List and click the bidirectional arrow (red box 3). This adds factor covariances for all constructs. This is the complete CFA model. One can also use Model Shortcuts, select Cross-Sectional Classics > Confirmatory Factor Analysis. This will provide another window to specify the confirmatory factor model.)

Exhibit 5 Lists and status panels

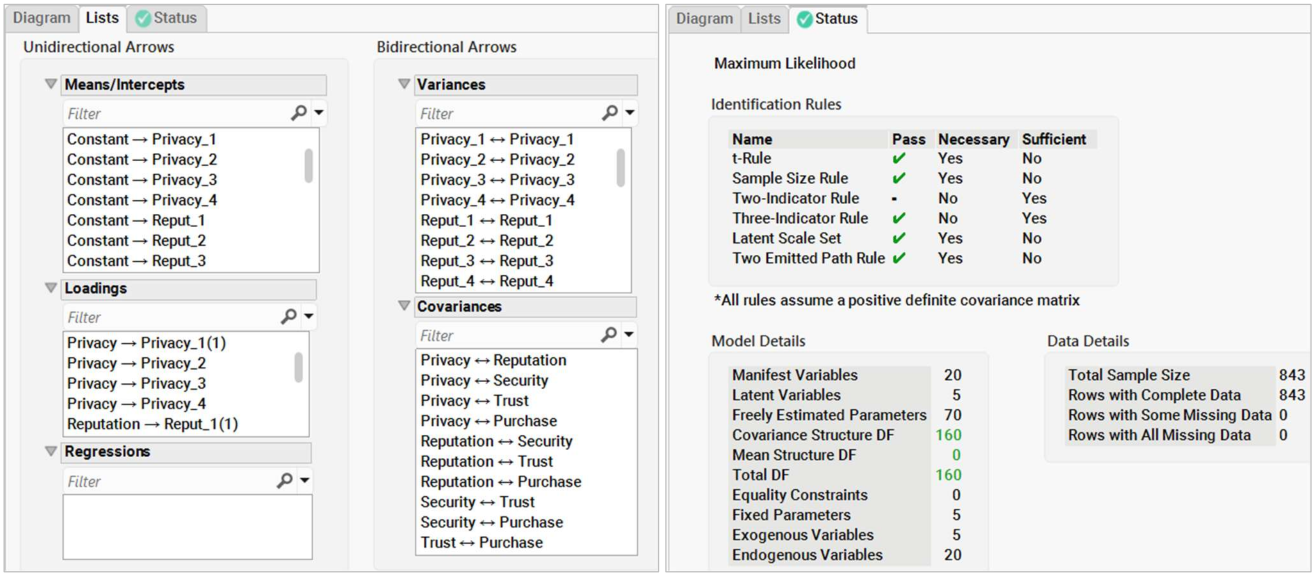


Exhibit 5 shows the information about the model that is currently specified.

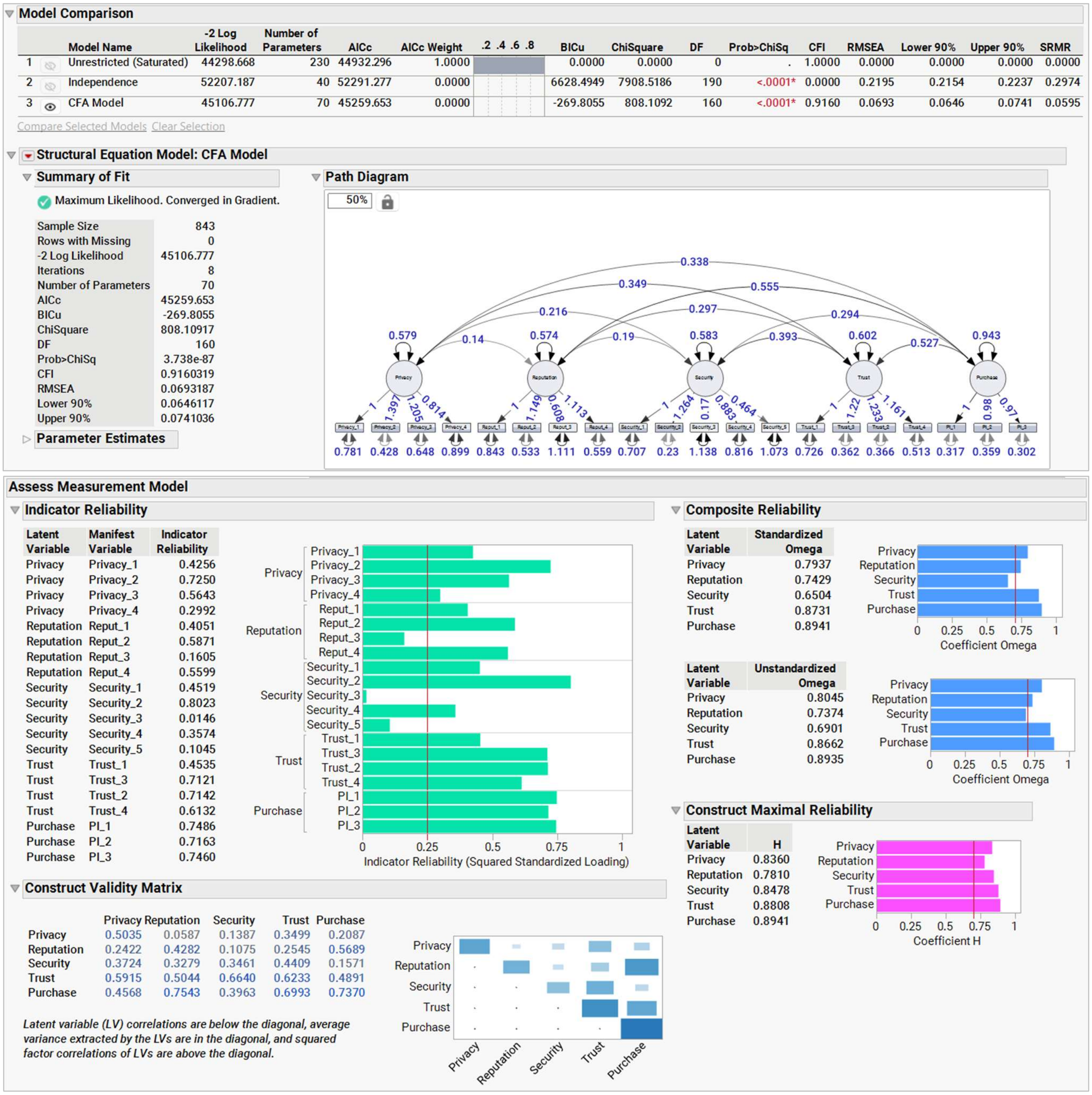
- The Diagram tab contains the model diagram. In the diagram, latent variables are represented by circles, and indicator variables are represented by rectangles. Unidirectional arrows represent loadings and regressions. Bidirectional arrows represent variances and covariances. Please note that the path diagram is interactive, so you can drag items to rearrange them or customize any aspect of the diagram by using the Customize Diagram option in the context menu.
- The Lists tab contains the lists for each type of parameter in the model. Means and intercepts for the model appear here. By right clicking on the canvas of the diagram and looking under Show, you can find Show Means/Intercepts option to display them in the diagram, too.
- The Status tab contains checks for model identifiability, which means that unique solutions for every model parameter can be derived based on the covariance matrix of the input data. The Status tab includes three panels of information: Identification Rules, Model Details, and Data Details. A list of identification rules (necessary and sufficient) and their overall pass status will be reflected by an icon as shown below:

All applicable identification rules pass

At least one necessary identification rule does not pass and needs correction

Model Details provides descriptive values of the model, while Data Details provides information about the input data.

Exhibit 6 Partial output from Confirmatory Factor Analysis



Model Comparison

	Model Name	-2 Log Likelihood	Number of Parameters	AICc	AICc Weight	BICu	ChiSquare	DF	Prob>ChiSq	CFI	RMSEA	Lower 90%	Upper 90%	SRMR
1	Unrestricted (Saturated)	44298.668	230	44932.296	1.0000	0.0000	0.0000	0	.	1.0000	0.0000	0.0000	0.0000	0.0000
2	Independence	52207.187	40	52291.277	0.0000	6628.4949	7908.5186	190	<.0001*	0.0000	0.2195	0.2154	0.2237	0.2974
3	CFA Model	45106.777	70	45259.653	0.0000	-269.8055	808.1092	160	<.0001*	0.9160	0.0693	0.0646	0.0741	0.0595

Compare Selected Models Clear Selection

Structural Equation Model: CFA Model

Summary of Fit

Maximum Likelihood. Converged in Gradient.

Sample Size	843
Rows with Missing	0
-2 Log Likelihood	45106.777
Iterations	8
Number of Parameters	70
AICc	45259.653
BICu	-269.8055
ChiSquare	808.10917
DF	160
Prob>ChiSq	3.738e-87
CFI	0.9160319
RMSEA	0.0693187
Lower 90%	0.0646117
Upper 90%	0.0741036

Parameter Estimates

Path Diagram

Assess Measurement Model

Indicator Reliability

Latent Variable	Manifest Variable	Indicator Reliability
Privacy	Privacy_1	0.4256
Privacy	Privacy_2	0.7250
Privacy	Privacy_3	0.5643
Privacy	Privacy_4	0.2992
Reputation	Reput_1	0.4051
Reputation	Reput_2	0.5871
Reputation	Reput_3	0.1605
Reputation	Reput_4	0.5599
Security	Security_1	0.4519
Security	Security_2	0.8023
Security	Security_3	0.0146
Security	Security_4	0.3574
Security	Security_5	0.1045
Trust	Trust_1	0.4535
Trust	Trust_3	0.7121
Trust	Trust_2	0.7142
Trust	Trust_4	0.6132
Purchase	PI_1	0.7486
Purchase	PI_2	0.7163
Purchase	PI_3	0.7460

Indicator Reliability (Squared Standardized Loading)

Composite Reliability

Latent Variable	Standardized Omega
Privacy	0.7937
Reputation	0.7429
Security	0.6504
Trust	0.8731
Purchase	0.8941

Latent Variable	Unstandardized Omega
Privacy	0.8045
Reputation	0.7374
Security	0.6901
Trust	0.8662
Purchase	0.8935

Construct Maximal Reliability

Latent Variable	H
Privacy	0.8360
Reputation	0.7810
Security	0.8478
Trust	0.8808
Purchase	0.8941

Construct Validity Matrix

	Privacy	Reputation	Security	Trust	Purchase
Privacy	0.5035	0.0587	0.1387	0.3499	0.2087
Reputation	0.2422	0.4282	0.1075	0.2545	0.5689
Security	0.3724	0.3279	0.3461	0.4409	0.1571
Trust	0.5915	0.5044	0.6640	0.6233	0.4891
Purchase	0.4568	0.7543	0.3963	0.6993	0.7370

Latent variable (LV) correlations are below the diagonal, average variance extracted by the LVs are in the diagonal, and squared factor correlations of LVs are above the diagonal.

(*Click the Run button in the Action section above the model diagram. Click the red triangle next to Structural Equation Model: CFA Model and choose Assess Measurement Model.)*

Reliability and Validity

Exhibit 6 shows the Assess Measurement Model output. The indicator reliability plot shows the squared standardized loadings of the latent variables along with a suggested minimum threshold for acceptable reliability (0.25). Low values for a variable indicate that the variable does not do a good job of capturing variability in the corresponding latent variable. The indicator variables Reput_3 (0.16), Security_3 (0.01) and Security_5 (0.10) are below the 0.25 threshold, suggesting they do a poor job of capturing variability in the corresponding latent variables. The Security_3 item has such a low value that it could be removed

with negligible loss to the reliability of the construct, whereas the Security_5 and Reput_3 could be revised to improve their reliability.
The composite reliability (CR) plot shows the coefficient omega (McDonald's omega), which is preferred over Cronbach's alpha because it assumes neither unidimensionality nor equal factor loadings. Omega represents the proportion of variance of the latent variable(s) in the observed composite score(s). The values range from 0 to 1, and an acceptable value is 0.7 and above. Composite reliability for Security (0.67) could be improved by focusing on improving the indicator reliability for Security_3 and Security_5.

The construct maximal reliability report shows coefficient H, which represents the proportion of latent variable variance represented by the indicators. The values range from 0 to 1, and it is recommended that they be 0.70 or greater.

The construct validity matrix provides a way to determine whether latent variables are measuring what they are intended to measure. It has three sets of values:

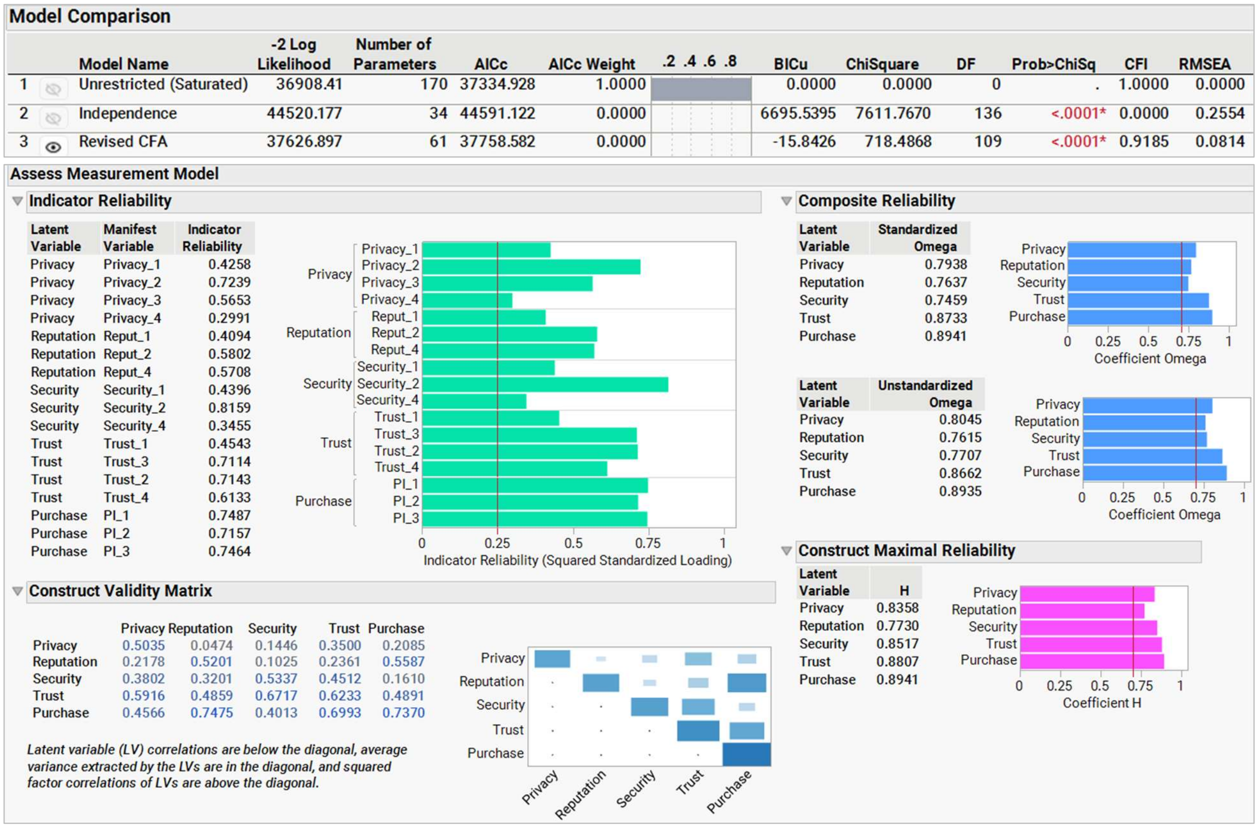
- The lower triangular entries contain the latent variable correlations, which help to check the strength of correlation between latent constructs.
- The upper triangular entries are the squared latent variable correlations.
- The diagonal entries contain the average amount of variance extracted (also called average variance extracted, or AVE) by each latent variable, which is equivalent to the average of the indicator reliabilities of each latent variable.

Convergent validity can be established by high factor loadings or by a higher value of AVE. Security and reputation constructs have AVE is less than 0.5, which is below the acceptable limit.

Discriminant validity exists when no two constructs are highly correlated (correlation greater than 0.85). If two constructs are highly correlated, it is recommended that the constructs be combined unless there is a strong theoretical reason not to do so. In this case, the correlation between the constructs is less than 0.85. An alternative test of discriminant validity, called the Fornell-Larcker criterion, is based on the logic that a construct should explain its observed variables better than it explains any other constructs. Discriminant validity is achieved if the AVE is greater than the square of its correlations (diagonal entry for each latent construct should be higher than the entries above and to the right of it). Visualization of the construct validity matrix helps to compare the diagonal entries to the upper triangular entries. Note that the diagonal values for Privacy, Trust, and Purchase are larger than the values above and to the right of them. However, this is not the case for Security and Reputation, which is further evidence that the items related to Security and Reputation can be improved to better measure these latent variables.

One can conclude that the reliability and validity can be improved by removing or revising some of the items related to Security and Reputation. Anna has decided to omit the indicator variables Reput_3, Security_3 and Security_5. Exhibit 7 shows the revised CFA output after omitting these variables. The revised measurement model assures the reliability and validity of the constructs.

Exhibit 7 Partial output from the revised CFA model



Model Comparison

	Model Name	-2 Log Likelihood	Number of Parameters	AICc	AICc Weight	BICu	ChiSquare	DF	Prob>ChiSq	CFI	RMSEA
1	Unrestricted (Saturated)	36908.41	170	37334.928	1.0000	0.0000	0.0000	0	.	1.0000	0.0000
2	Independence	44520.177	34	44591.122	0.0000	6695.5395	7611.7670	136	<.0001*	0.0000	0.2554
3	Revised CFA	37626.897	61	37758.582	0.0000	-15.8426	718.4868	109	<.0001*	0.9185	0.0814

Assess Measurement Model

Indicator Reliability

Latent Variable	Manifest Variable	Indicator Reliability
Privacy	Privacy_1	0.4258
Privacy	Privacy_2	0.7239
Privacy	Privacy_3	0.5653
Privacy	Privacy_4	0.2991
Reputation	Reput_1	0.4094
Reputation	Reput_2	0.5802
Reputation	Reput_4	0.5708
Security	Security_1	0.4396
Security	Security_2	0.8159
Security	Security_4	0.3455
Trust	Trust_1	0.4543
Trust	Trust_3	0.7114
Trust	Trust_2	0.7143
Trust	Trust_4	0.6133
Purchase	PI_1	0.7487
Purchase	PI_2	0.7157
Purchase	PI_3	0.7464

Indicator Reliability (Squared Standardized Loading)

Composite Reliability

Latent Variable	Standardized Omega
Privacy	0.7938
Reputation	0.7637
Security	0.7459
Trust	0.8733
Purchase	0.8941

Latent Variable	Unstandardized Omega
Privacy	0.8045
Reputation	0.7615
Security	0.7707
Trust	0.8662
Purchase	0.8935

Construct Maximal Reliability

Latent Variable	H
Privacy	0.8358
Reputation	0.7730
Security	0.8517
Trust	0.8807
Purchase	0.8941

Construct Validity Matrix

	Privacy	Reputation	Security	Trust	Purchase
Privacy	0.5035	0.0474	0.1446	0.3500	0.2085
Reputation	0.2178	0.5201	0.1025	0.2361	0.5587
Security	0.3802	0.3201	0.5337	0.4512	0.1610
Trust	0.5916	0.4859	0.6717	0.6233	0.4891
Purchase	0.4566	0.7475	0.4013	0.6993	0.7370

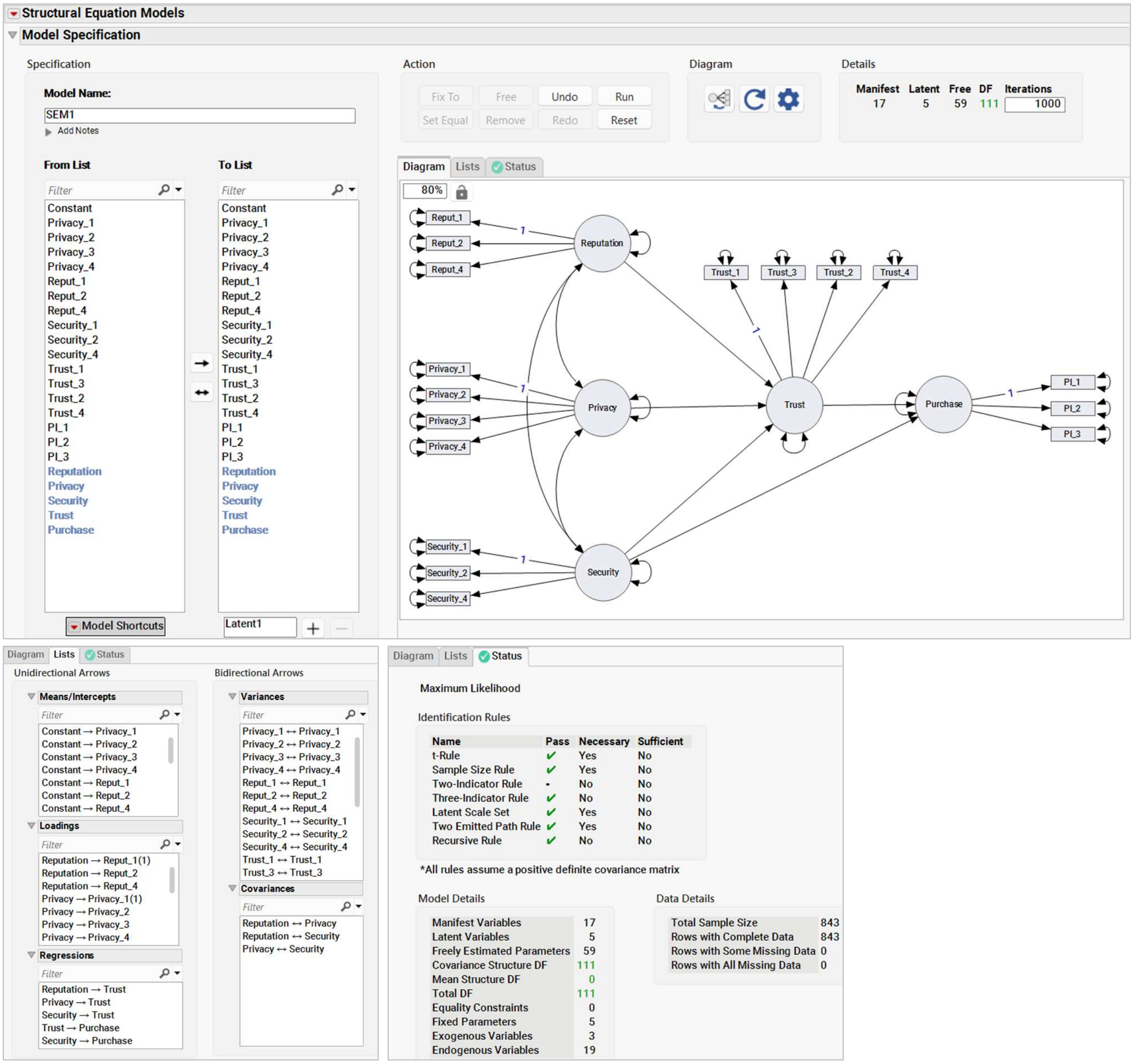
Latent variable (LV) correlations are below the diagonal, average variance extracted by the LVs are in the diagonal, and squared factor correlations of LVs are above the diagonal.

(Repeat the steps from Exhibits 4-6, but exclude Reput_3, Security_3, and Security_5 from the model.)

Fitting Competing Structural Equation Models

The structural regression model represents the theory that specifies how the latent and observed variables are related to each other. The model is created using prior knowledge and then assessed according to how well it fits the correlational structure among the observed variables. Exhibit 8 shows the specification in the SEM platform of the structural regression model corresponding to Model 1 in Exhibit 1.

Exhibit 8 Model specifications for SEM1 (Model 1)

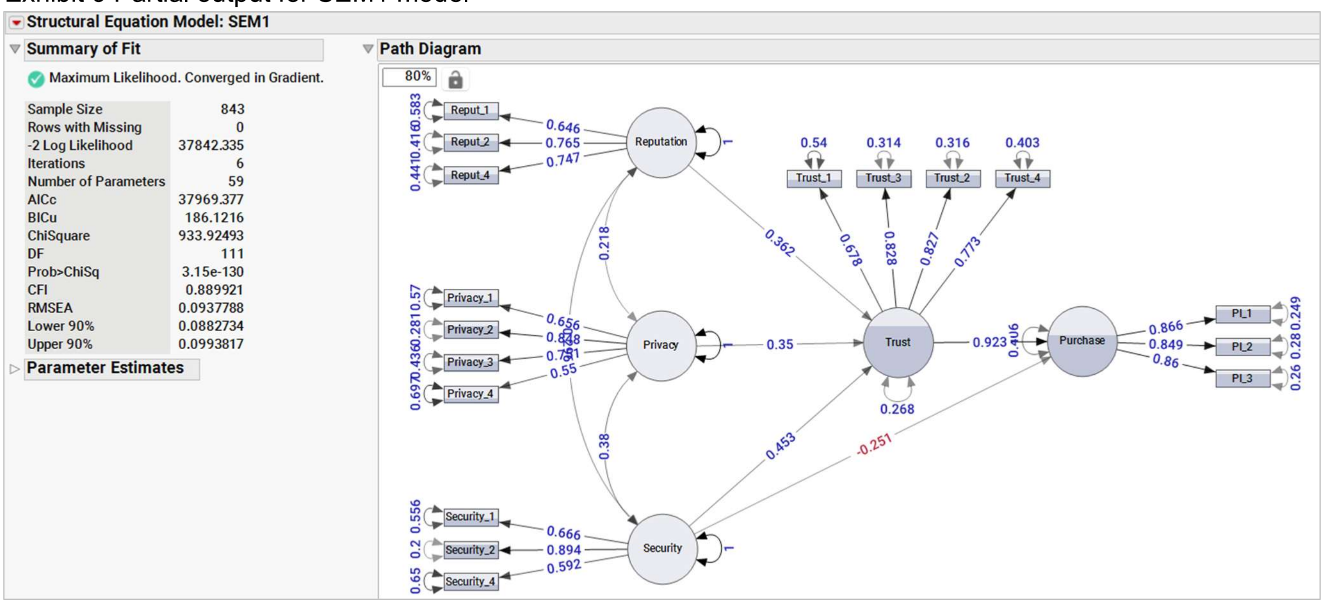


(Step 1: Analyze > Multivariate Methods > Structural Equation Models. Select all the indicator variables starting from Privacy_1 to PI_3 (except Reputation_3, Security_3 and Security_5) and move them to model variables > Click OK. This will launch the model specification dialog.
Step 2: Under Model Name, type SEM1.
Step 3: Create the latent constructs. Using the To List under Model Specification, select the variables that are presumed to load onto a latent variable, enter the name for the latent variable in the box below the To List, and click the plus button. For example, select Privacy_1, Privacy_2, Privacy_3 and Privacy_4, type Privacy in the box, and click the plus button to add the latent construct to the diagram. Repeat this process until all the latent constructs for the model have been specified.
Step 4: Add covariances. Select Privacy, Reputation and Security in the From List, then select Privacy, Reputation and Security in the To List, and finally click the bidirectional arrow button.
Step 5: Add Regressions. Select Privacy, Reputation and Security in the From List, then select Trust in the To List, and click the unidirectional arrow. Select Trust in the From List and Purchase in the To List, and click the unidirectional arrow. Select Security in the From List and Purchase in the To List, and click the unidirectional arrow button.

Step 6: Check the Status tab to ensure the specified model passes all identification rules.
Step 7: Click Run.)
Note that the latent variable specification from the CFA can be copied and pasted into the SEM Model Specification panel using the red triangle options of Copy Model Specification and Paste Model Specification options, respectively. By doing so, the steps to specify the structural regression model are minimized, as you can simply replace the covariances between latent variables with the hypothesized regressions. In other words, all the point-and-click effort to specify the latent variables can be eliminated once an acceptable CFA is fit.

Also note that the SEM platform always includes a mean structure, so all the indicator variables are listed in the Means/Intercepts list as outcomes of the Constant term. Moreover, all latent variables are automatically identified by setting the loading of their first indicator to 1 (default) or their variance to 1 if the Standardized Latent constructs option was selected in the launch window. Exhibit 9 shows the partial output of the SEM1 model.

Exhibit 9 Partial output for SEM1 model



(Click the red triangle next to Structural Equation Model: SEM1 and select Path Diagram Settings > Layout > Top to Bottom. Click the gray disclosure icon next to Parameter Estimates. Closing the Parameter Estimates report enables you to see the entire path diagram. If you require standardized parameters estimates in the path diagram, right click in the diagram and select Show > Show Estimates > Standardized.)

Summary of Fit for SEM1 Model

As shown in Exhibit 9, the summary of fit table provides information about the model fit, including the convergence status and the estimation method. Because the chi-square value depends on the sample size, some poorly fitted models can still produce a significant chi-square value. The comparative fit index (CFI) and root mean square error of approximation (RMSEA) provide additional guidance for determining model fit. These indices are bounded between 0 and 1. CFI values greater than 0.90 and RMSEA values less than 0.10 are preferred. Here, the CFI of 0.89 and RMSEA of 0.09 indicate a reasonably good fit. Note that additional measures of fit are available through the Fit Indices option under the red triangle.

Now that we have built Model 1, follow similar steps listed in Exhibit 8 to build SEM2 for Model 2 from Exhibit 1. Exhibit 10 shows the partial output of the SEM2 model after incorporating the direct effect of Reputation on Purchase.

Exhibit 10 Partial output for SEM2 model

Structural Equation Model: SEM2

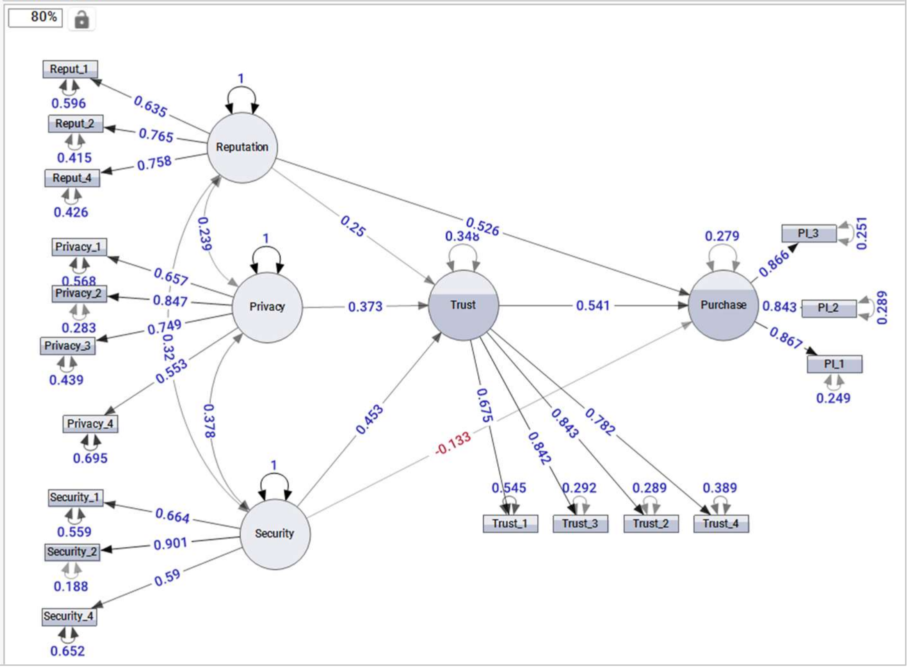
Summary of Fit

Maximum Likelihood. Converged in Gradient.

Sample Size	843
Rows with Missing	0
-2 Log Likelihood	37637.491
Iterations	6
Number of Parameters	60
AICc	37766.852
BICu	-11.98491
ChiSquare	729.08146
DF	110
Prob>ChiSq	5.276e-92
CFI	0.9171882
RMSEA	0.0817078
Lower 90%	0.0761264
Upper 90%	0.0873958

Parameter Estimates

Path Diagram



(Return to the Model Specification section. Change the model name to SEM2. Select Reputation in the From List and Purchase in the To List and click the unidirectional arrow. Run the model.)

For SEM2, the CFI of 0.92 and RMSEA of 0.08 indicate a good fit and are slightly better than the values for SEM1. The next step is to perform a direct comparison of the two models.

Comparing Competing Models (Model Comparison)

The model comparison report in Exhibit 11 contains a table of all the models that have been fit. The metrics to the right of each model's name enable comparison of models based on various criteria. By default, two default models (unrestricted and independence) are also shown. Select the SEM1 and SEM2 models and click Compare Selected Models. The two models are compared statistically, and a chi-square difference test report is generated. Note that this comparison should only be done with nested models. In this case, SEM1 is nested in SEM2, so we can proceed.

Exhibit 11 Model comparison report

Model Comparison

	Model Name	-2 Log Likelihood	Number of Parameters	AICc	AICc Weight	.2 .4 .6 .8	BICu	ChiSquare	DF	Prob>ChiSq	CFI	RMSEA
1	Unrestricted (Saturated)	36908.41	170	37334.928	1.0000		0.0000	0.0000	0	.	1.0000	0.0000
2	Independence	44520.177	34	44591.122	0.0000		6695.5395	7611.7670	136	<.0001*	0.0000	0.2554
3	SEM1	37842.335	59	37969.377	0.0000		186.1216	933.9249	111	<.0001*	0.8899	0.0938
4	SEM2	37637.491	60	37766.852	0.0000		-11.9849	729.0815	110	<.0001*	0.9172	0.0817

Compare Selected Models Clear Selection

Chi-Square Difference Test

Model nested...	...in model	ΔChiSquare	ΔDF	Prob>ChiSq	ΔCFI	ΔRMSEA
SEM1	SEM2	204.8435	1	<.0001*	-0.027	0.0121

Difference tests are meaningful only for nested models

(Select SEM1 and SEM2 in the Model Comparison section, then click Compare Selected Models (highlighted by red box).)

The results show the differences in CFI and RMSEA for the two models, along with a chi-square test and p-value. SEM1 has a lower CFI and higher RMSEA, and the p-value is statistically significant at the 0.05 level. These results suggest that SEM2 fits the data better than SEM1. We choose to prefer SEM2 and move on to interpreting that model.

Interpreting Parameter Estimates for SEM2

Exhibit 12 shows the parameter estimates for SEM2. All six of the regression parameters are statistically significant. Therefore, we conclude that the true regression effects specified in the model are nonzero. The parameter estimates show positive effects of Privacy, Reputation, and Security on Trust as expected. They also suggest a positive effect of Trust and Reputation on Purchase and a small negative effect of Security on Purchase. This negative effect was unanticipated, as Anna had hypothesized that increased perception of security should increase purchase intent. This surprising result deserves scrutiny.

Exhibit 12 Parameter estimates for SEM2

Parameter Estimates

Means/Intercepts	Estimate	Std Error	Wald Z	Prob>|Z|
Constant → Privacy_1	3.2657177	0.0401657	81.306095	<.0001*
Constant → Privacy_2	3.0960854	0.0429874	72.023172	<.0001*
Constant → Privacy_3	3.1447212	0.0420127	74.851653	<.0001*
Constant → Privacy_4	3.4211151	0.0390109	87.696278	<.0001*
Constant → Reput_1	3.4033215	0.0410046	82.998464	<.0001*
Constant → Reput_2	2.8896797	0.0391412	73.826992	<.0001*
Constant → Reput_4	3.0652432	0.0388285	78.943132	<.0001*
Constant → Security_1	2.9513642	0.0391101	75.463063	<.0001*
Constant → Security_2	3.5065243	0.0371108	94.48797	<.0001*
Constant → Security_4	3.5077106	0.0388156	90.368578	<.0001*
Constant → Trust_1	2.6417556	0.0396958	66.550067	<.0001*
Constant → Trust_3	2.5432977	0.0386424	65.816172	<.0001*
Constant → Trust_2	3.2941874	0.0390026	84.460752	<.0001*
Constant → Trust_4	2.7912218	0.0396505	70.395675	<.0001*
Constant → PI_1	2.5717675	0.038665	66.514102	<.0001*
Constant → PI_2	3.0604982	0.0387288	79.023931	<.0001*
Constant → PI_3	2.5551601	0.0375531	68.041185	<.0001*

Loadings	Estimate	Std Error	Wald Z	Prob>|Z|
Reputation → Reput_1	1	0	.	.
Reputation → Reput_2	1.1486499	0.070481	16.297287	<.0001*
Reputation → Reput_4	1.1293556	0.067998	16.60865	<.0001*
Privacy → Privacy_1	1	0	.	.
Privacy → Privacy_2	1.3794339	0.0791624	17.425366	<.0001*
Privacy → Privacy_3	1.192235	0.0707474	16.851994	<.0001*
Privacy → Privacy_4	0.8166223	0.0574297	14.219515	<.0001*
Security → Security_1	1	0	.	.
Security → Security_2	1.2874218	0.0759916	16.94163	<.0001*
Security → Security_4	0.8812307	0.0577866	15.249739	<.0001*
Trust → Trust_1	1	0	.	.
Trust → Trust_3	1.2142215	0.0576532	21.060769	<.0001*
Trust → Trust_2	1.2279899	0.0581771	21.107801	<.0001*
Trust → Trust_4	1.1575764	0.0580722	19.933391	<.0001*
Purchase → PI_1	1	0	.	.
Purchase → PI_2	0.9746678	0.0318577	30.59444	<.0001*
Purchase → PI_3	0.9701517	0.03015	32.177507	<.0001*

Regressions	Estimate	Std Error	Wald Z	Prob>|Z|
Reputation → Trust	0.2568974	0.0360763	7.1209355	<.0001*
Reputation → Purchase	0.6766962	0.0555146	12.189509	<.0001*
Privacy → Trust	0.3786197	0.0383771	9.8657735	<.0001*
Security → Trust	0.4670238	0.044375	10.524472	<.0001*
Security → Purchase	-0.171121	0.0542631	-3.153549	0.0016*
Trust → Purchase	0.6763818	0.0620815	10.895058	<.0001*

Variances	Estimate	Std Error	Wald Z	Prob>|Z|
Privacy_1 ↔ Privacy_1	0.7726511	0.0461715	16.73439	<.0001*
Privacy_2 ↔ Privacy_2	0.4401621	0.0438914	10.028434	<.0001*
Privacy_3 ↔ Privacy_3	0.6530818	0.042544	15.35074	<.0001*
Privacy_4 ↔ Privacy_4	0.8912371	0.0492872	18.082516	<.0001*
Reput_1 ↔ Reput_1	0.8452103	0.0491676	17.190387	<.0001*
Reput_2 ↔ Reput_2	0.5365573	0.0395814	13.555777	<.0001*
Reput_4 ↔ Reput_4	0.5411508	0.0392661	13.781611	<.0001*
Security_1 ↔ Security_1	0.7206195	0.044381	16.237103	<.0001*
Security_2 ↔ Security_2	0.2181793	0.0405904	5.3751434	<.0001*
Security_4 ↔ Security_4	0.8283717	0.0476026	17.40181	<.0001*
Trust_1 ↔ Trust_1	0.7236799	0.0390618	18.526536	<.0001*
Trust_3 ↔ Trust_3	0.3672994	0.0244604	15.016069	<.0001*
Trust_2 ↔ Trust_2	0.3705395	0.0247856	14.949793	<.0001*
Trust_4 ↔ Trust_4	0.5150695	0.0302839	17.008039	<.0001*
PI_1 ↔ PI_1	0.313963	0.0222833	14.089614	<.0001*
PI_2 ↔ PI_2	0.3654594	0.0241435	15.136945	<.0001*
PI_3 ↔ PI_3	0.2981723	0.0211505	14.097631	<.0001*
Reputation ↔ Reputation	0.5721929	0.0615965	9.2893731	<.0001*
Privacy ↔ Privacy	0.587348	0.0607227	9.6726319	<.0001*
Security ↔ Security	0.56883	0.0587395	9.6839454	<.0001*
Trust ↔ Trust	0.2105785	0.0226714	9.2882693	<.0001*
Purchase ↔ Purchase	0.2641074	0.0265887	9.9330682	<.0001*

Covariances	Estimate	Std Error	Wald Z	Prob>|Z|
Reputation ↔ Privacy	0.1387963	0.0263735	5.2627218	<.0001*
Reputation ↔ Security	0.182337	0.0277546	6.5696177	<.0001*
Privacy ↔ Security	0.2186352	0.0283733	7.7056651	<.0001*

Summary

Statistical Insights

To summarize, this case called upon Anna to use JMP Pro to do the following:

- Perform confirmatory factor analysis (CFA) to confirm measurement validity and reliability.
- Build competing structural regression models.
- Compare the models using a chi-square difference test.
- Interpret the parameter estimates of the best fitting model.

Managerial Implications

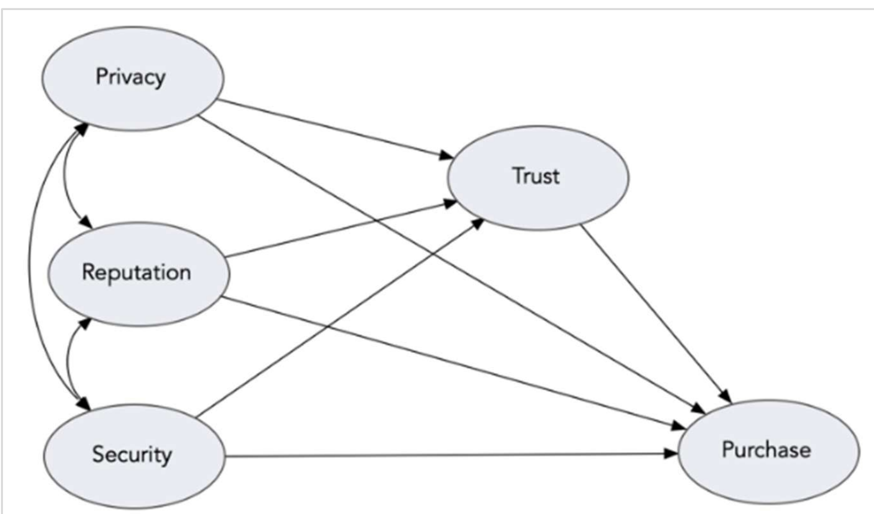
Some of the survey items were not reliable and led to construct validity issues. Anna removed the relevant indicator variables to achieve the desired level of measurement reliability and validity. She then fit two theoretically motivated structural regression models, and determined that the model in which Reputation direction influences Purchase was preferable to the alternative. She concluded that the seller's reputation not only builds customer's trust, but it also drives purchase intentions directly.

JMP Features and Hints

This case study uses the Structural Equation Model platform in JMP Pro. This platform was used to fit both CFA and structural regression models. Model shortcuts and the ability to copy-paste model specifications were used to reduce the number of clicks to specify models.

Exercise

Anna has a third theory in which all three constructs of Privacy, Reputation, and Security influence both Trust and Purchase, as shown below.



- Using the same data, build a structural regression model to test the theory.
- Is the overall model statistically significant? What are the values of CFI and RMSEA, and do they indicate a good model fit?
- Are there any regressions in the model that are not statistically significant? If so, which one(s)?
- Perform a nested model comparison between this model and Model 2 (SEM2). Do you obtain evidence that this model fits the data better than Model 2? How do you know?