CS-BIGS 2(2): 99-108 http://www.bentley.edu/csbigs/vol2-2/wilson.pdf

Combining Web Mining Techniques and Structural Equations Modeling for Measuring E-commerce Perceptions

Dax Jacobson, Ekin Pehlivan, Sonia Vilvovsky and Wilson Wong
Bentley University, USA

The goal of this study is to illustrate a novel approach to a common research question. Much research has been conducted to measure perceived value in e-commerce and several studies have utilized structural equation modeling (SEM). However, the vast majority of these studies rely on data collected from survey instruments. This study, which is appropriate for intermediate students of statistics, demonstrates what can be done using a combination of web mining and SEM techniques to analyze data captured directly from an online store, in this case data on LCD (Liquid Crystal Display) television sets. Both the data collection technique and the SEM methodology provide new opportunities for measuring perception in e-commerce contexts which can also be replicated and built upon by future researchers.

1. Introduction

Structural equation modeling (SEM) is a research approach used in many academic disciplines, including information systems (Gefen, et al., 2000) and marketing (Steenkamp & Baumgartner, 2000). This method is often used to examine the perceptions of customers and has been used in e-commerce contexts to measure perceived value (Chen & Dubinsky, 2003) and ratings of web-sites (Kwon, et al., 2002). However, SEMs are not used as frequently as they could be due to data availability issues and the large number of observations needed (Steenkamp & Baumgartner, 2000). The data availability and sample size issues arise because of the reliance on survey data, which can be difficult to obtain. Surveys take time, are expensive, and require a decent response rate. In addition, problems may arise from self-reporting.

In this paper, we argue for the importance of SEM as a research method that should be used more often. We illustrate a web mining data collection technique which does not require a survey instrument and is thus able to circumvent many of the problems associated with survey data.

For example, we avoid the problems with self-reporting data by collecting data directly from a single e-commerce site. We then build a simple model based on these data that contributes to the research on customer perception from another angle.

The purpose of our paper, though, is not to build or test theory but to explore an alternative approach to survey data collection when creating SEM models that students can use for learning SEM techniques. To ensure clarity in our reporting we follow Chin's (1998) recommendations. In the next section we describe the data collection techniques, the population from which the data sample was obtained and the distribution of the data (to determine the adequacy of the statistical estimation procedure). This is followed by a description of the methods and the conceptual model (to determine the appropriateness of the statistical models analyzed). We then discuss the results and a comparison model and finally conclude with a discussion of the implications of this study for students of SEM techniques as well as implications for future research.

2. Data

The web mining techniques employed in this study consisted of Perl programs that were written to download and extract data from a website. The software WGET 1.10.2 is a freely available software utility from http://www.gnu.org/software/wget/ (the Free Software Foundation) and is a non-interactive network program that retrieves files from the World Wide Web. The Perl programming language was a natural choice for extracting the downloaded data for this study. Perl, short for Practical Extraction and Report Language, is a programming language that is "optimized for scanning arbitrary text files, extracting information from those text files, and printing reports based on that information" (Ashton 2001). Using the network utility, WGET 1.10.2, a Perl program was written to download LCD and plasma television data from the CNET website by first retrieving the popular TV web pages beginning at http://shopper.cnet.com/4032-6475_9.html (Figure 1) and then the TV specifications web page for each of the 649 televisions listed. The television specifications web page for the Samsung LN-T3253H is given in Figure 2. An example of Perl processing code that extracts a single field, the diagonal size, from a downloaded television webpage is given in Appendix A.

For each television model's specifications web page, the program extracted data fields that included price information, user rating, make and model, the list of online stores, vendor ratings, vendor prices including shipping and tax, and technical specifications (see Table 1). In addition to the data that were taken from the web pages, additional calculated fields were created (Table 2) to analyze television properties that could not be represented by any of the single downloaded fields.

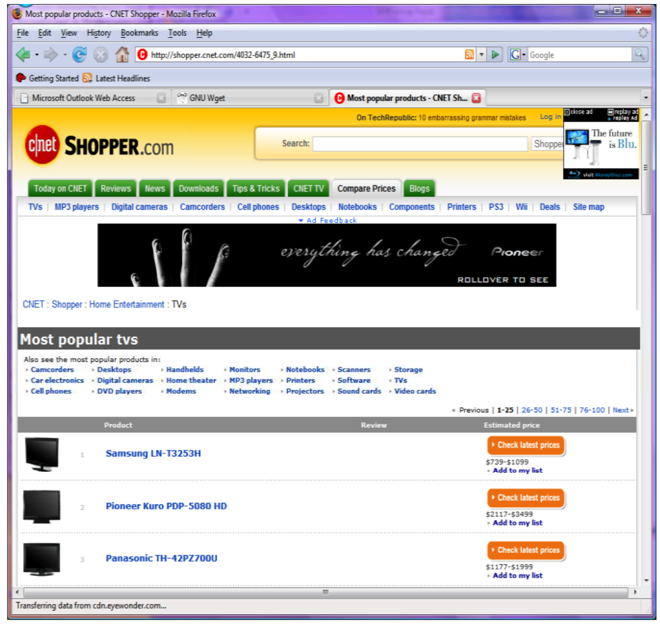

Figure 1. Screen Shot of CNET Shopper Website

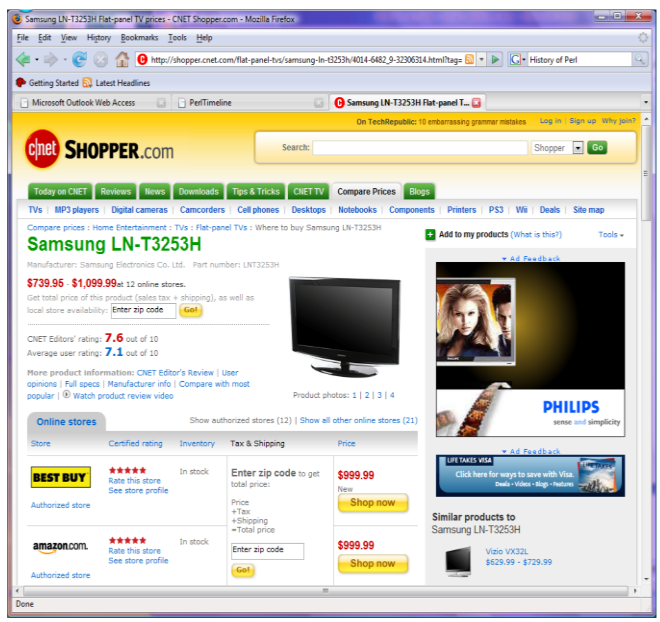

Figure 2. Screen Shot of CNET Information for Samsung LN-T3253H

Table 1. Observed Variables in the Data Set

Variable	Description	Minimum	Maximum	Mean	Std. Deviation
Low Price	The lowest price for this set among all vendors, in US $	199.95	6337	1188.22	825.8
High Price	The highest price for this set among all vendors, in US $	214.99	7999.98	1824.87	1242.61
Base Price	The price offered by this vendor, in US $	199.95	7999.98	1477.68	1050.44
Full Price	Base price + taxes + shipping, in US $	199.95	8479.98	1525.55	1079.53
Diagonal Size	Of Screen, in inch	15	65	37.77	9.64
Width	TV set width, in inch	14.6	61.9	37.32	9.07
Depth	TV set depth, in inch	2.9	19.4	6.81	3.49
Height	TV set height, in inch	12.4	38.4	25.13	5.53
Weight	TV set weight, in pounds	8	143	50.86	24.89
Speakers	Total speakers power, in watt	2	50	20.75	7.1
ResHor	Horizontal Resolution, in pixels/inch	640	1920	1584.42	293.35
ResVert	Vertical Resolution, in pixels/inch	480	1080	895.96	158.16
NumDigT	# of digital tuner types	0	3	1.57	0.68
IO Total	# of all input/output ports	4	23	13.26	3.87
NumSlots	# of slots	0	5	0.03	0.23
Stores	Total number of stores carrying this model	1	30	13.71	8.07

On November 3, 2007 at 1:33pm, the Perl program, over the course of 3 hours, extracted all of the television data from the CNET web pages and produced pipe delimited records which use vertical bars to separate fields within a record. These pipe-delimited records were then imported into Excel and SPSS for analysis. Each observation includes data for a specific television set offered for sale by a specific vendor. From the initial set of 3350 observations, 585 observations that had more than four missing values in different fields were dropped. Two outlier categories, state-of-the-art televisions and portable televisions, were identified and removed from our data set. Television sets which incorporate state-of-the-art technology command an extremely high premium and would therefore skew our results. This category includes 14 observations of very large (diagonals greater than 65”) television sets and 6 observations of sets that are priced higher than $8,000. Portable televisions are another outlier category because their mobility makes their use significantly different from most household televisions. This category includes four observations of very small (diagonals less than 15”) television sets.

Table 2. Computed Variables in the Data Set

Variable	Description	Minimum	Maximum	Mean	Std. Deviation
HighDelta	The difference of the highest and base prices, in US $	0	3000	347.18	368.6
LowDelta	The difference of base and lowest prices, in US $	0	3000	289.46	349.64
LowRatio	The ratio of base price to low price	1	2.4	1.24	0.23
HighRatio	The ratio of high price to base price	1	2.4	1.26	0.21
PriceRange	The ratio of highest price to lowest price	1	3	1.54	0.3
Overhead	Tax + Shipping	0	480	47.86	65.76
OverRatio	The ratio of overhead to the base price	0	0.26	0.03	0.04

After removing the outliers and observations with too many missing values, 2,841 observations were left. Missing values for size and weight were replaced with the means of the four closest points and high price missing values were replaced by the base price (see Table 1). The resulting set of 2,841 observations contained 2,122 observations for LCD television sets and 719 observations for plasma television sets.

Further analysis showed that technology is an important differentiator for customers’ perception of the product. We continue our analysis based on the LCD televisions sets data only, for a total of 2,122 observations. Each observation includes the name of the manufacturing company and the name of online vendor. It also includes the model and series number of the television set. Table 1 summarizes continuous variables in the data set of 2,122 observations for LCD TVs from which we began the analysis.

3. Structural Equations Modeling

SEM is a technique used for specifying and estimating models of linear relationships among variables which may include both measured variables (MVs) and latent variables (LVs). LVs are hypothetical constructs that cannot be directly measured; they are typically represented by multiple MVs that serve as indicators of the construct. Hence we can say that a structural equation model is a hypothesized pattern of directional and non-directional linear relationships among a set of MVs and LVs. Directional relationships imply some sort of directional influence of one variable on another whereas non-directional relationships are correlated and imply no directed influence (MacCallum and Austin, 2000).

Although the application of structural equation modeling for comprehensive investigations of both measurement and theoretical issues is generally acknowledged (e.g., Anderson & Gerbing, 1988; Bagozzi, 1984; Bagozzi & Yi, 1988; Dillon, 1986; Steenkamp & van Trijp, 1991) some authors have commented critically on the technique's value for empirical research. These criticisms range from outright denial of the method's usefulness because of the presumed implausibility of underlying assumptions (e.g., Freedman, 1987) to concerns about the way in which SEM has been applied in practice (e.g., Breckler, 1990; Biddle & Marlin, 1987; Cliff, 1983; Fornell, 1983; Martin, 1987). However this statistical technique is invaluable for the purposes of this paper in our attempts to uncover latent constructs and their relationships.

In this study, structural equation modeling is used to understand customer perceptions of LCD TVs. Instead of using survey data to understand customer perceptions, we employed web mining techniques to access available information on the CNET website. There are several advantages to SEM in comparison to methods accomplishing similar objectives. Chin (1998) states that SEM provides the researcher with the flexibility to: (a) model relationships among multiple predictor and criterion variables, (b) construct unobservable latent variables (hereafter LVs), (c) model errors in measurements for observed variables, and (d) statistically test a priori substantive/theoretical and measurement assumptions against empirical data (i.e., confirmatory analysis).

We began building the model by performing a factor analysis. Factor analysis summarizes the interrelationships among the variables in a concise but accurate manner by representing as much of the variability of the original variables in a smaller number of derived variables, or factors, in order to achieve a comprehensible solution (Gorsuch, 1983). A factor loading is a measure of the degree of generalizability found between each variable and each factor. The farther the factor loading is from zero, the easier it is to generalize from that factor to that variable. Comparing loadings of the same variable on several factors provides information concerning how appropriate it is to generalize to that variable from each factor (Gorsuch, 1983). Because factor analysis does not result in a unique solution, equivalent solutions can be obtained by rotating the factor matrices. Statistical packages such as SPSS can perform several types of factor rotations so that most variables are strongly loaded on a single factor. In building the initial SEM model, we wanted each factor to have at least three strongly loaded variables since SEM typically requires at least three variables per latent construct.

Table 3. Rotated Component Matrix

	Component			
	1	2	3	4
Vendor Full Price	0.533	0.754	0.032	0.073
Stores	0.012	-0.062	0.798	0.084
Speakers	0.818	0.013	0.061	0.022
ResVert	0.51	0.565	0.189	0.047
NumPCs	0.007	-0.126	-0.082	0.771
NumDigT	-0.215	0.201	0.288	0.613
IO Total	0.316	0.07	-0.424	0.615
Width	0.82	0.467	0.125	0.028
Depth	0.603	0.056	-0.292	-0.085
Height	0.878	0.347	0.063	0.018
Weight	0.835	0.363	0.075	0.02
HighDelta	0.476	0.05	0.651	-0.013
LowDelta	0.174	0.904	0.026	-0.025
PriceRangeRatio	-0.075	0.322	0.675	-0.211

Fourteen variables from the data set were selected for factoring. Some variables were omitted due to a high correlation with other variables. For example, the screen diagonal size is highly correlated with the width of the television set, and vertical resolution is correlated with horizontal resolution. Table 3 represents the SPSS output of the extracted principal components rotated with the Varimax rotation algorithm. Price and Vertical resolution variables were divided by 100, to bring their ranges closer to that of the other variables in the model.

The initial structural equation model was drafted using the AMOS statistical tool (AMOS 7.0.0 build 1140, .NET Framework Version). Four LVs of the initial model represent the extracted principle components:

- Component 1, "Perceived size": width, height, weight, speaker power, also price, resolution and depth.
- Component 2, "Perceived monetary value": Price, Low Price Delta, also resolution and width
- Component 3, "Perceived quality of service": number of online stores, prices range, high price delta
- Component 4, "Perceived technological sophistication": PC interfaces, No. of digital tuners, IO total

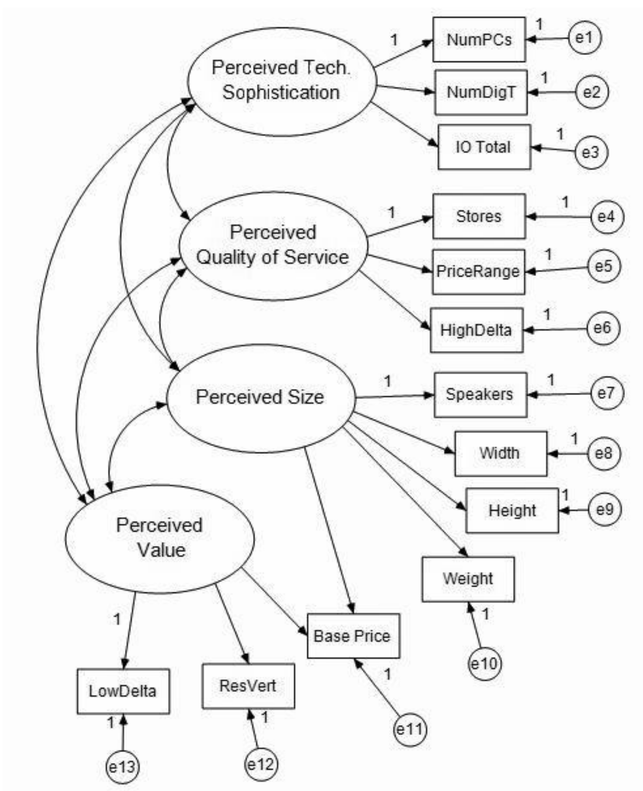

Figure 4. The Initial Measurement Model Based on Factor Analysis

The initial measurement model is shown in figure 4. In addition to insufficient fit indices, this model is unacceptable due to a negative estimated variance for the error term of the Total Number of I/O. As a result, exploratory research was conducted by refining the model through the use of fit indices, modification indices and other output provided by the modeling tool. A modification index "estimates the amount by which the overall model chi-square statistic would decrease if a particular fixed-to-zero path were freely estimated" (Kline, 2005). Modification indices with greater values indicate a better predicted improvement in overall model fit if those paths are added to the model (ibid).

We began the process of creating the model by first removing all covariances between the latent constructs. The first model rendered a Chi-square of 6427.3 with 64 degrees of freedom. Reviewing the modification indices revealed that the index for covariance of Perceived Size and Perceived Value had a large value of 620.124 so we added this covariance back to the model and tested the model again. This resulted in an improved Chi-square of 5219.4, with 63 degrees of freedom; however, the model was not admissible due to the negative variance of the error term for Price. In addition, the modification indices revealed two large values – one for a covariance of the error terms Delta Low and Delta High (differences between the base price and low or high price respectively), and Delta Low and Price Range (the ratio of high price to low price). These three variables are highly related and their presence in the model in this form was likely redundant.

We determined that different combinations of variables describing high, low and base prices could provide better fit to the model. As an example, price ratios may better reflect customers' perceptions than absolute differences, and the model could benefit from introduction of the overhead, e.g. the shipping and handling costs that are not included in the base price. We proceeded to replace the absolute price differences with price ratios (Low Ratio is the ratio of base price to low price, and High Ratio is the ratio of high price to base price), and the Price Range ratio – by Overhead Ratio (the ratio of overhead to base price). Price was then loaded on two latent constructs – as these constructs are now covariate in the model, we could load Price on only one of them. In the next step we replaced absolute differences in Price with the newly created Price Ratios and loaded Price on only the Perceived Value construct. Testing this model resulted in a Chi-square of 2455.8, with 64 degrees of freedom; however, the model did not converge – the limit of 49 iterations was reached, and no estimations were produced.

This was a discouraging result. However, even in a situation where the model does not converge, the modeling tool provides valuable output which can be used in the modeling process. In our case, the error of Num of Online Stores had a negative variance. The highest modification index for this error term suggested covariance with the error term for Low Ratio. However, when we introduced this covariance to the model, the model failed to converge. As the modification indices suggested a covariance between the error terms of Num of Online Stores and Low Ratio, we decided to add a covariance between the two constructs themselves. This resulted in a Chi Square of 2363.7 (with 63 degrees of freedom), with a persistent negative variance, this time for the High Ratio's error term.

At this point, we took a step back and examined the modification indices again. We noticed that the Perceived Technological Sophistication construct and the three variables loading on it did not show high modification indices. We had decided initially to use four principal components in our model, knowing that the Perceived Technological Sophistication component was the weakest component of the group. Based on the modification indices we determined that the model might be better off with only three principal components. To accomplish this we could either remove the Perceived Technological Sophistication construct entirely or merge it with the Perceived Value construct. We decided on the latter for two reasons: first, the perception of technological sophistication may be viewed as a part of the total perception of value; and second, when we added the covariance between the Perceived Technological Sophistication and the Perceived Value constructs, there was no change in the model's Chi square.

We observed a high modification index for the error term of the Num of PC interfaces, and several high modification indices for Width and Height. As there is some research evidence that suggests that the actual shape of the TV set influences perceptions of size more than height, we replaced the Height variable with SizeRatio. The Num of PC interfaces variable was also removed from the model. Testing this model resulted in an improved Chi square of 1983.0 (df=52); however, the Width and High Ratio's error terms had negative variances. Reviewing the modification indices suggested creating covariances between both the Width and Size Ratio, and the High and Low Ratios (the latter could be problematic as the respective latent constructs also covariate). Based on the modification indices we proceeded to replace the covariance between Perceived Quality of Service and Perceived Value with a covariance between Perceived Quality of Service and Perceived Size. Finally, we also removed the Number of Digital Tuners from the model based on the modification indices. This final measurement model, displayed in Figure 5, has a Chi square of 1349.9 (df=40).

The next step in the process is to build a structural model from the measurement model. We proceeded by hypothesizing (based on the prior research as well as experience) the possible relationships between the latent constructs and came up with a suggested influence of perceived value on perceived size and perceived quality of service. The structural model is presented in Figure 6, and has a Chi square 1327.4 (df=40).

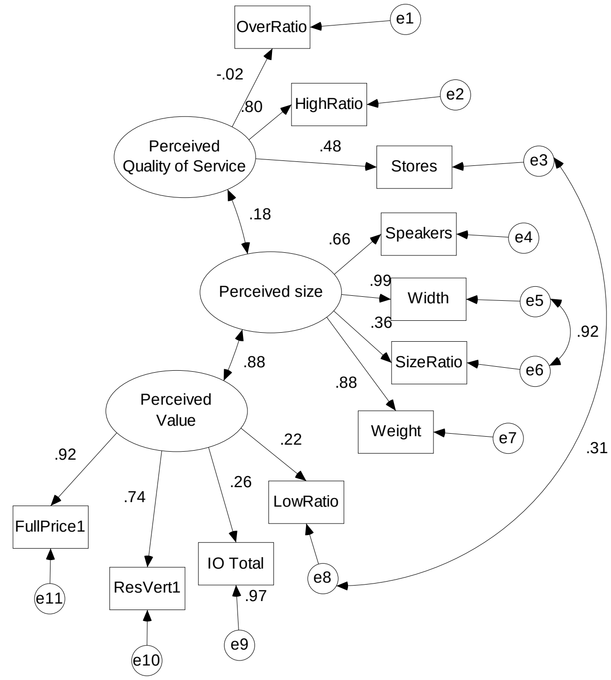

Figure 5. The Final Measurement Model

4. Model & Analysis

The final model is displayed in Figure 6 with standardized coefficients included. Table 4 lists the variables used in the final model and Table 5 is a summary of the model fit measurements. The values reveal that our model provides an acceptable fit to the data. The Goodness of Fit (GFI) and the Adjusted GFI (AGFI), which are the statistics measuring the fit (adjusted for degrees of freedom-for AGFI) of the combined measurement and structural model to the data, are near the limit (for acceptable values and detailed description see Gefen, et al., 2000). However, keeping in mind the nature of the data it is fair to say that the values (GFI=.898, AGFI=.832) are satisfactory. The BIC criterion, which helps select parsimonious models with a good fit to the data (Deichmann, et al., 2006) is the only model fit criterion that does not fit, however it is important to maintain the models conceptual integrity while improving the model fit. This tradeoff is one that has to be taken into account in the explanation of the model.

We recall that the measurement model is a sub-model in structural equation modeling that specifies the indicators for each construct and assesses the reliability of each construct for estimating the relationships (Gefen, et al., 2000).

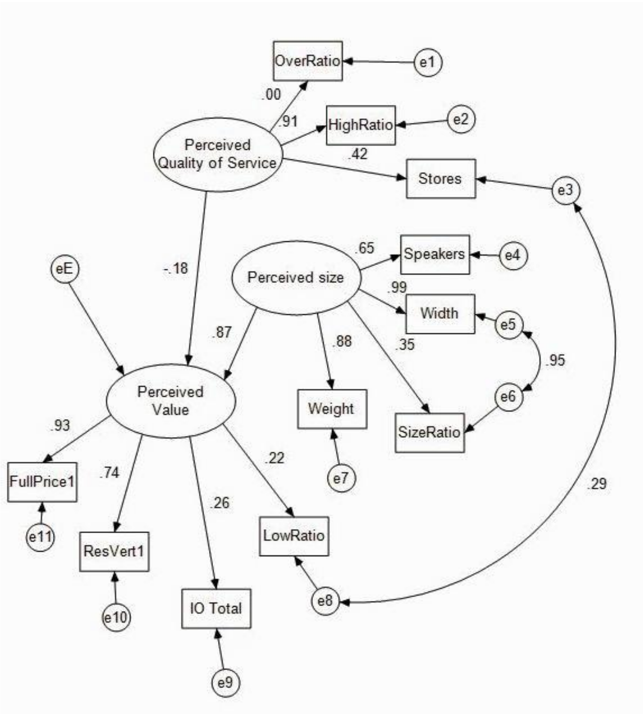

Figure 6. The Structural Model

The measurement model (see Figure 5) revealed that the variables represented by the number of stores, the ratio of the highest price to the base price for the specific TV model and the overhead defined as the shipping/handling fee and taxes load on to the latent construct Perceived Quality of Service. Higher ratios of overhead and highest price to base price may indicate a store's competitive strategy to concentrate on customer service rather than on low price. Since a higher Perceived Quality of Service partially derives from higher prices, the scale of this latent construct moves in the opposite direction from Perceived Value and Perceived Size. Perceived Size includes speakers, width, size ratio and weight. The height and depth were highly correlated with the width and weight, and were removed. The resolution, full price and the ratio of the lowest price to the specific TV and the I/O ports load on to the Perceived Value in the measurement model.

Table 5. Summary of Model Fit Values for Final Model

	Chosen model		
Chi-Square	1327.4		
Df	40	**Suggested values**	
GFI	0.898	closer to 1	
AGFI	0.832	closer to 1	
RMSEA	0.123	**Saturated model**	**Independence.**
BIC	1526	505.568	9942.892

Table 4. Summary of Variables in the Final Model

Variable	Description	Minimum	Maximum	Mean	Std. Deviation
	Observed variables				
Stores	The number of online stores offering this TV set	1	30	13.71	8.07
Speakers		2	50	20.75	7.099
Width		14.6	61.9	37.322	9.0657
Weight		8	143	50.86	24.89
	Computed variables				
Full Price1	Full price divided by 100	2	84.8	15.2555	10.795
ResVert1	Vertical resolution divided by 100	4.8	10.8	8.9596	1.5816
IO Total	Total number of input/output ports	4	23	13.26	3.872
LowRatio	The ratio of base price to low price	1	2.4	1.2396	0.22583
HighRatio	The ratio of high price to base price	1	2.4	1.2554	0.21295
OverRatio	The ratio of tax and shipping fees to the base price	0	0.26	0.0338	0.03951
SizeRatio	The ratio of width to height	0.96	2.22	1.481	1.35

Table 6 displays a portion of AMOS output with standardized regression weights.

Using these regression weights, we can represent the standardized structural equation model with the following set of regression equations, noting that all variables are assumed to be standardized in the equations below:

$$P.Vaul = -.186 * P.QualityofService + .870 * P.Size + E$$

$$OverheadRatio = .001 * P.QualityofService + e1$$

$$HighRatio = .911 * P.QualityofService + e2$$

$$OnlineStores = .421 * P.QualityofService + e3$$

$$TotalSpeakersPower = .650 * P.Size + e4$$

$$Width = .991 * P.Size + e5$$

$$SizeRatio = .352 * P.Size + e6$$

$$Width = .880 * P.Size + e7$$

$$LowRatio = .22 * P.Value + e8$$

$$IOTotal = .26 * P.Value + e9$$

$$Resolution = .74 * P.Value + e10$$

$$FullPrice = .93 * P.Value + e11$$

Table 6. Part of AMOS Output for Final Model

Standardized Regression Weights: (Group number 1 - Default model)			
			Estimate
P. Value	<---	P.Quality of Service	-0.185
P. Value	<---	PSize	0.87
TotalSpeakerPowerWatt	<---	PSize	0.65
OnlineStroes	<---	P.Quality of Service	0.421
HighRatio	<---	P.Quality of Service	0.911
Price1	<---	P.Value	0.926
IOTotal	<---	P.Value	0.259
LowRatio	<---	P.Value	0.218
SizeRatio	<---	Psize	0.352
OverRatio	<---	P.Quality of Service	0.001
ResVert100	<---	P.Value	0.742
Weight	<---	Psize	0.88
Width	<---	Psize	0.991

The structural model is used to illustrate a set of one or more dependent relationships linking the model constructs. The structural model presented in Figure 6 suggests that Perceived Quality of Service and Perceived Size of the TV influence the Perceived Value separately. The influence is significant for both latent constructs on Perceived Value. Thus, the interpretation is that the information presented on a website for each TV is perceived in three categories and two of these perceptions influence the third one, the overall perceived value of the specific TV.

For comparison to the structural model in Figure 6, we include Figure 7, a model with better fit measurements (see Table 7) but much less relevance to theory. The model correlates error term 4 (e4), the # of online stores, with error term 20 (e20), the # of I/O ports. This one change may significantly improve model fit but the coefficients of the paths relating the constructs to one another remain almost the same. Moreover, correlating these two errors has no basis in theory, does not make any intuitive sense, and the model loses relevance. Caution must be taken when trying to improve any model – fit is important but not to the point that links to theory are lost.

Table 7. Summary of Model Fit Values for Comparison Model

	Chosen model	Suggested values	
Chi-Square	1219.3		
Df	39		
GFI	0.904	closer to 1	
AGFI	0.838	closer to 1	
RMSEA	0.119	Saturated model	Independence.
BIC	1426	505.568	10648.052

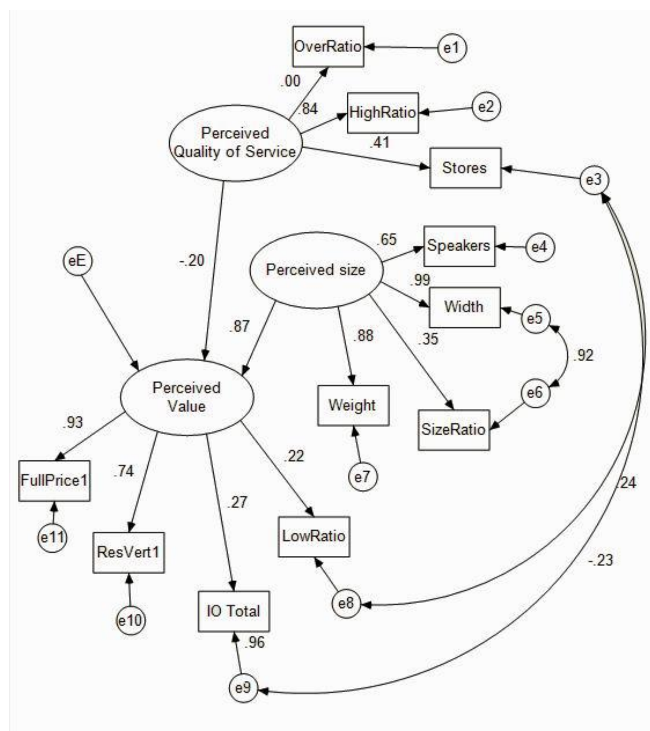

Figure 7. A Comparison Model

5. Conclusion

We have demonstrated the ability to combine web mining techniques with structural equation modeling in studying consumer perceptions. Although the web mining techniques require some programming, this offsets the requirements for survey data collection when creating SEMs. Moreover, the method presented in the paper allows for data to be collected relatively easily and frequently, allowing for frequent testing. The model we have presented is admittedly a simple one but it provides a foundation for future researchers to build upon and refine.

Future research could test other product categories as well as compare data taken directly from an e-commerce website with data taken from surveys and determine the implications for theory. At a minimum, the method and model presented in the paper provide new opportunities for SEM students as well as give researchers an additional source of data for testing.

Appendix A - Example of Perl Code

```
# datum field we are looking to extract
$spec = "Diagonal Size";

# call the subroutine getSpecs()
$diagonalSize = &getSpecs($currentwebpage, '<td>' . $spec . '</td>');

# Remove the text " in" from the diagonal size
($diagonalSize, $trash) = split(/ in/, $diagonalSize);

# $diagonalSize now contains the number representing the diagonal inches
# of the television

sub getSpecs()
{
  local $result = $_[0]; # first parameter is the downloaded webpage
  local $spec = $_[1];   # second parameter is the beginning of the text to be extracted
  local $trash;

  # remove all text before $spec
  ($trash,$result) = split(/${spec}/,$result);

  # In this example, the </tr> tag signifies the end of the text to be extracted
  ($result,$trash) = split(/<\/tr>/,$result);

  # Remove the <td></td> tags and extra spaces around the specification data
  ($trash, $result) = split(/<td>\s*/, $result);
  ($result, $trash) = split(/\s*<\/td>/, $result);

  return $result;
```

REFERENCES

Anderson, J.C. & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin,* 103: 411-423.

Ashton, E. (2001). "The Timeline of Perl and its Culture v3.0_0505." Retrieved December 23, 2007, from *http://history.perl.org/PerlTimeline.html*.

Bagozzi, R.P. (1984). A prospectus for theory construction in marketing. *Journal of Marketing* 48: 11-29.

Bagozzi, R.P. & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science,* 16: 74-94.

Biddle, B. J. & Marlin, M.M. (1987). Causality, confirmation, credulity, and structural equation modeling. *Child Development,* 58: 4-17.

Breckler, S.J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin,* 107: 260-273.

Chen, Z., & Dubinsky, A. J. (2003). A Conceptual Model of Perceived Customer Value in E-Commerce: A preliminary Investigation. *Psychology & Marketing*, 20(4): 323-347.

Chin, W. W. (1998). Issues and Opinion on Structural Equation Modeling. *Management Information Systems Quarterly*, 22(1): vii-xvi.

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research* 18: 115-126.

Deichmann, J., Eshgi, A., Haughton, D., Teebagy, N., & Topi, H. (2006). Determinants of Customer Churn Behavior: The Case of the Local Telephone Service. *Marketing Management Journal*, 16(2): 179-187.

Dillon, W.R. (1986). Building consumer behavior models with LISREL: Issues in applications. In D. Brinberg & R.J. Lutz (eds.), *Perspectives on Methodology in Consumer Research*, 107-154, New York: Springer.

Fornell, C. (1983). Issues in the application of covariance structure analysis. *Journal of Consumer Research*, 9: 443-448.

Freedman, D.A. (1987). As others see us: A case study in path analysis. *Journal of Educational Statistics* 12: 101-128.

Gefen, D., Straub, D. W., & Boudreau, M. (2000). Structural Equation Modeling and Regression: Guidelines for Research Practice. *Communications of the Association for Information Systems*, 4(7): 1-79.

Gorsuch, R.L. (1983). *Factor Analysis.* Hiilsdale, NJ: Lawrence Erlbaum.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling.* New York: The Guilford Press.

Kwon, O. B., Kim, C., & Lee, E. J. (2002). Impact of website information design factors on consumer ratings of web-based auction sites. *Behaviour & Information Technology*, 21(6): 387-402.

MacCallum R. C. & Austin J. T. (2000). Applications of Structural Equation Modeling in Psychological Research. *Annual Review of Psychology*, 51:201–226

Martin, J.A. (1987). Structural equation modeling: A guide for the perplexed. *Child Development,* 58: 33-37.

Steenkamp, J. E., & Baumgartner, H. (2000). On the Use of Structural Equation Models for Marketing Modeling. *International Journal of Research in Marketing*, 17(2-3): 195-202.

Steenkamp, J.B. & van Trijp, H. (1991). The use of LISREL in validating marketing constructs. *International Journal of Research in Marketing,* 8: 283-299.

Correspondence: wwong@bentley.edu