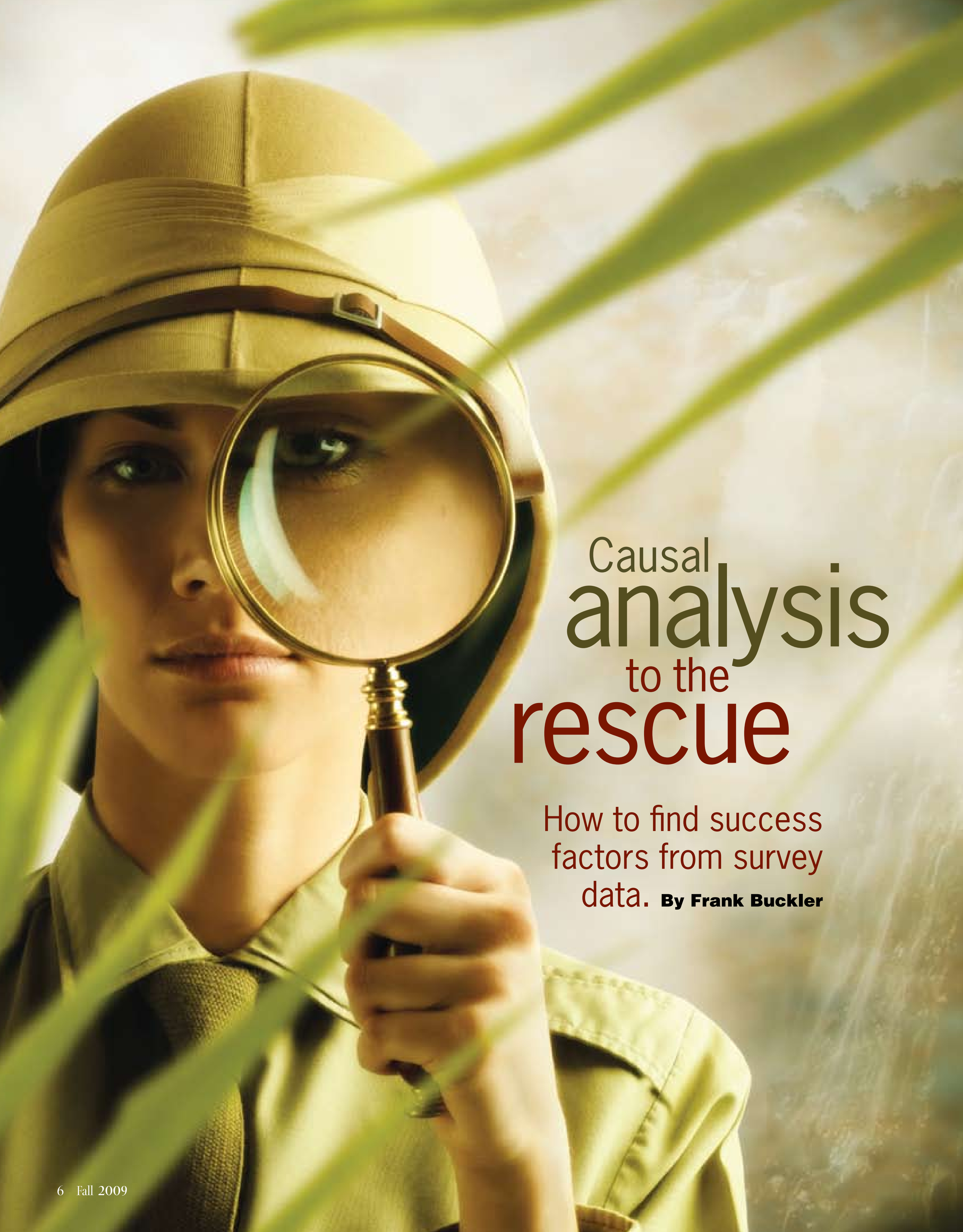


Causal analysis to the rescue

How to find success factors from survey data. **By Frank Buckler**

Surveys are conducted for many different reasons. While some are related to internal politics, most studies truly aim to find the key driver of success. Companies want to use their findings to boost the efficiency and effectiveness of marketing and sales. This turns out to be a tough call because many influencing factors are at work. In order to pull the right success triggers as a consequence of a study, it's necessary to learn more details about an effect, beyond just "X is a main driver." Even more important, it is crucial that the findings turn out to be effective and that they pay back every cent.

How would it be if you drew amazingly detailed cause-effect insights from survey data that were not possible until now? This article explains why existing analysis methods are often of limited use and how companies can profit from the opportunities offered by a new methodology.

Descriptive methods are normally used to infer success factors. Multivariate techniques, such as regression or structural equation modeling, are rarely deployed. Why are existing analysis methods of limited help in practice? A few reasons come to mind.

Spurious correlation. Most of the time results of surveys are presented as follows: "Successful enterprises invest on average 20 percent more into research and development." From this, one may conclude that investing more into R&D will foster success. In fact, this conclusion is probably wrong. The reason is spurious correlations. Perhaps this is a result of the innovative company culture that spreads its positive influence in every department and as a side effect makes the top management more willing to spend extra on R&D. Or maybe the R&D department itself creates more costs then effects, but is overcompensated by the productivity of the employees' working culture.

Executive Summary

Despite its popularity, descriptive analysis is barely suitable for concluding effective measures because it is prone to spurious correlations. Causal (path) analysis methods are needed to handle the risk, but today's methods are barely able to cope with the requirements and complexity of business applications. This article introduces a practical method that explores unknown success factors and is able to uncover complex but invaluable details. Case studies illustrate the value to market research.

Because of spurious correlations, nobody should draw causal conclusions, whenever possible, out of descriptive or bivariate analysis. You always have to take all relevant data within one analysis into account. Even though 95 percent of all analysis and conclusions drawn in day-to-day business are descriptive or bivariate, this remains a momentous mistake.

Spurious correlations can only be avoided using multivariate analysis as multiple regression. When data of all relevant terms are present, this method calculates the direct effect of one term on the other.

Latent variables and interconnectedness. Companies try hard to control things like customer satisfaction, loyalty and attitude toward a brand. To reliably model such terms, you have to measure them in several unique ways because these terms cannot be observed directly. The results are combined into indices or latent variables. After this step, researchers can analyze the causes of these terms. Standard multivariate analysis methods are not designed for this two-step process. Furthermore, causes often influence each other and are not independent of one other. This contradicts the working assumption of multivariate methods such as multiple regression. To address these issues, latent variable path modeling methods were developed. Until today there has been little use of these methods in business applications.

Unknown relations and properties. In practice it is rarely known which variable influences which and which are unrelated. But the standard path modeling methods (e.g., structural equation modeling) require exactly this, which is often a knockout criteria for business applications. That's why more exploratory approaches are needed in practice.

The Tetrad project developed processes and methods for explorative causal path modeling. Still another severe issue remained. In most cases relations are nonlinear, and variables are moderating the effects of others (so-called interactions). In fact, an analysis of four arbitrarily chosen data sets published in two leading marketing journals revealed that every model contained unexpected nonlinearities, interactions or paths. We know less then we think.

Applying advanced approaches of SEM, which are able to consider nonlinearities or interaction, unfortunately does not help. The reason is that the kind of nonlinearity and interaction is unknown beforehand. For confirmatory approaches like SEM you have to specify in advance exactly how the relations will look. In order to model unknown nonlinearities and interactions you must delve into the world of data mining methods.

Which Methodology to Use?

Universal structural modeling (USM) is a new causal analysis methodology based on artificial neural networks first introduced in 2001. Like all latent variable path analysis, it handles measurement and structural models in a combined way. USM is most similar to the partial least squares analysis (PLS), but instead of linear regression uses a data mining method. The USM process is as follows:

Step 1: Model specification. With USM one has to define which manifest variables (items) will build a latent variable. Next, the user has to define if the measurement model is of the reflexive or formative kind. Last, one has to define the direction of paths because this could only be inferred from data when longitudinal data are given. Of course the user is free to eliminate even more paths to influence the results with

> How would it be **if you drew** amazingly **detailed** cause-effect **insights** from survey data that were **not possible** until now?

his theoretical knowledge. However, in contrast to confirmatory methods, the user does not define which paths are in the model, but only which are not.

Step 2: Measurement model. Here factor loadings of measurement models are determined. The starting values are calculated using principle component analysis. In principle it is even possible to use alternative compression methods for special applications (e.g., ordinal PCA, nonlinear PCA or multidimensional PCA) in order to address, for instance, scale usage heterogeneity.

Step 3: Structural model. Separate regression models are calculated for every dependent latent variable. USM uses artificial neural networks as a regression method. Then Step 2 is repeated (using predictions of regression models as feedback information) as long as a stop criterion is not reached.

Within USM all Bayesian universal function approximators (in particular Bayesian multilayer perception and Gaussian processes) according to the MacKay framework (1992) are recommended. The main reason is because of their proven effectiveness in eliminating irrelevant variables (automated relevance detection) as well as parameters (automated soft pruning). It prevents not only the overfitting phenomenon of highly parameterized models but also leads to effective selection of true paths.

Instead of using the author's software, USM can be self-programmed by ambitious readers using statistical programming environments and freely available toolboxes. When doing this it's important to make sure that activation in hidden units is sigmoid and linear for output units intended for metric variables. For real life data sets I recommend using just one hidden layer because it is less likely to get stuck in local optimas. As a number of hidden units, I recommend approximately 10. Increase this number when high order nonlinearities are expected and large data sets are given. Of course other data mining techniques may be used as CART (although limited in kind of nonlinearities and interactions), MARS or nonparametric regression. Nevertheless I recommend Bayesian universal function approximators for the aforementioned reasons.

Step 4: Post processing. In the final step, measures for quantifying overall importance of effects, interaction effects, measurement validity (Chrombachs Alpha) and fit values as R2 and GoF are determined. The overall importance measure was developed for the USM environment because linear path coefficients are meaningless for nonlinear relations. This is how it is calculated: By replacing the cases of a given variable by its mean, you can experience a difference in explained variance using the trained neural network. This drop in variance explanation in relation to the overall variance represents the importance of the variable for explaining overall variance.

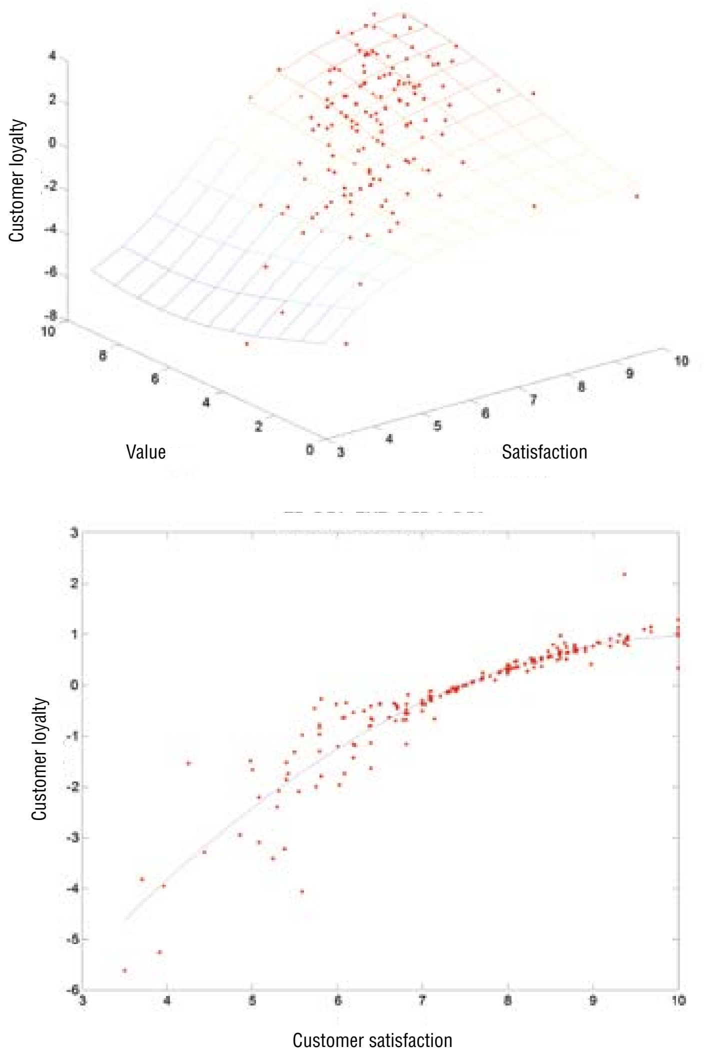
In addition, a bootstrapping statistic is computed for all measures in order to retrieve a significance measure. Nonlinear relations and moderating effects (two-way interactions) are visually extracted using a method introduced by Plate (1998). Using these techniques the "black box" of neural networks can be discovered.

When Should USM Be Used?

USM is an exploratory, flexible method for modeling causal paths using latent variables. Whenever you have good theoretically backed knowledge about the causal relations behind data, you should use confirmatory methods (SEM, PLS or econometric models). In this case they are expected to always produce models with higher validity. It is the key learning of Bayesian statistics that the more true a priori knowledge you incorporate in a method, the better your results get.

But be careful because the method will not tell you if your assumption is true. You have to know it beforehand. If in doubt, it's better to use—at least as a complement—exploratory methods. Looking at dozens of confirmatory studies, my experience is that in practice it is a quite rare event to have good a priori true knowledge about paths and shapes of relations (i.e., nonlinearities and interactions). Exhibit 1 may guide your choice of latent variable path methods.

Exhibit 1 Effects of satisfaction on loyalty



If you do not have latent variables, other data mining techniques (e.g., CART or MARS) can give you instant insights and might be a quick alternative to USM. I recommend applying USM in these circumstances because it is a methodical package that is predesigned to take care of overfitting, input selection issues, differentiating between paths and interactions, significance measures and so forth.

Some words on causality: According to Granger's definition of causality, one can infer causality from data when (1) all relevant variables are available (closed world assumption) and (2) a time difference between cause and effect is measured. If we do not have time serial data, one has to derive the causal direction from a priori knowledge. Assuming a closed world and truly defined path directions, the variable that incorporates the most information that helps explain the effect of another can be seen as the cause. The term "information" is broad because the definition allows nonlinear transformation of variances.

Your USM analysis may get biased or even wrong if you failed to incorporate important variables that influence the model. Furthermore, it may fail if you assume wrong path directions. In fact, this is true for every kind of statistical method. However, it is a sensible objection to doubt any statistical inference because of our ignorance of omitted variables.

Another limitation may come from the fact that multistep procedures such as USM do not always end with a globally optimal solution. To minimize the risk, it's a good idea to increase the number of iteration loops and run the same analysis several times.

The size of your data set may also limit the ability of USM to reliably explore new properties. At least the bootstrapping procedure should point to not reliable findings. The more cases (given a fixed number of variables), the more complex relation can be in principle revealed. I recommend as a rule of thumb to always have more than 200 cases. Other rules for estimating the needed number of cases (like VC dimension) are not practical because they depend on huge unrealistic data sets. And, because you don't know the complexity of the relations at hand, you also do not know how many data cases are needed.

Case Study: Customer Satisfaction

A national cellular network corporation asked its consultancy to find the specific driver of customer loyalty in order to craft proper measures. For the analysis we came back to a customer satisfaction survey that had already been conducted.

The conceptual framework behind the survey proposed that customer satisfaction results from customers' expectations and quality and value perceptions and in turn influences perceptions of loyalty and complaint handling. We used a sample of 250 customers, whose data were collected through personal CATI telephone interviews.

All direct additive effects (i.e., paths) found significant by formerly applied confirmatory methods also are significant in the USM estimation, and vice versa. The only exception is the path from perceived quality to complaint handling, which was not included in confirmatory models. The goodness-of-fit measure, which combines structural model and measurement model accuracy, is more than 10 percent higher for USM compared to PLS. Why?

USM reveals some hidden nonlinear and interactive model relations. With regard to nonlinearity, satisfaction affects loyalty in a nonlinear way by following a degressive growth function, a finding consistent with previous research. The existence of an interaction effect of perceived value and satisfaction on customer loyalty, which is relatively strong and significant. See Exhibit 3, right.

The three-dimensional interaction surface graphic illustrates that a saturation level exists for the effect of satisfaction on loyalty; after a critical level of loyalty, an increase in satisfaction does not transform into higher loyalty rates. As we can derive from the interaction graphic, this saturation level is lower when customer value is small and higher when that value is high. In other words, high loyalty can be achieved only when the customer is both highly satisfied with a product and assigns a high value to it.

The application of USM revealed the following:

Medium satisfaction is sufficient to keep customers loyal.

Unlike textbook theories, value for money has a direct moderating influence on loyalty.

High perceived value for money increases loyalty, if a customer is satisfied.

Consequently, increasing perceived value by establishing a particular level of satisfaction is needed.

With this, we helped our client to craft a very efficient strategy that focuses on eliminating "satisfaction killers" instead of "getting perfect." The complaint management process was optimized. Skipping some discounts for unsatisfied customers saved the company $1.5 million. Furthermore, communication was directed to emphasize the delivered value of service even more. This example shows the practical value of exploring interaction effects.

Case Study: Brand Image Analysis

A regional utility provider asked its consultant to figure out: "How does a green image improve customer loyalty?" We took the customer satisfaction monitor survey every year. It contained 700 respondents and image items including an image component "perceived as environmentally friendly." In addition, the loyalty status was measured by means of the individual willingness to change provider. With this data we analyzed the effects of the image components on customer loyalty with the software Neusrel, which is an implementation of USM.

Supprizingly we found an inverted-U-shaped effect. The actual level of perceived environmental friendliness resulted already in the maximum of customer loyalty. We learned from NEUSRELs plot, that an increase of the environmental friendliness to maximum would lead to a potential loss of 4% of customers. But a decrease of the perception of the same item would results in the same loss.

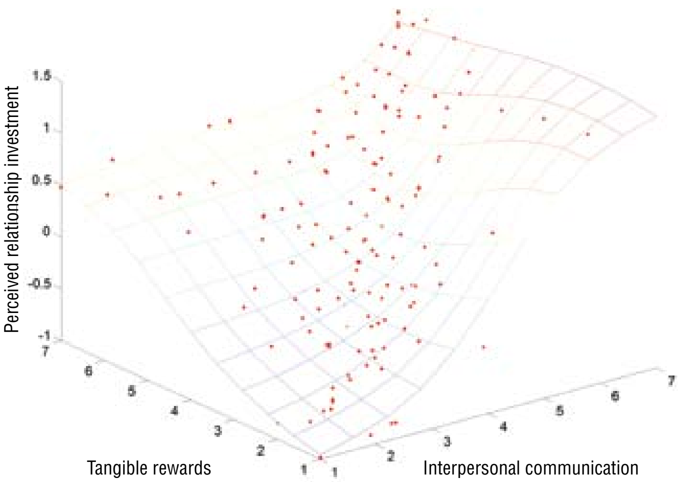
We offered our customer the following explanation of the findings: A moderately green image is perfect, and your image is today already close to perfection. By focusing on green initiatives you might win some strongly ecologically oriented customers. All other customers get the impression that you would waste money instead of lowering prices. This finding was also confirmed within focus groups that where conducted within the project.

The customer was surprised by these eye-opening findings. The evidence as well the sound rationale convinced him. The enterprise boiled down the wrong-leaded image campaign and saved $2 million. The best of it: The strategy worked and the number of customers did not decrease. This example shows how USM is able to explore hidden nonlinearities that were not expected in advance but make perfect sense.

Case Study: Marketing Mix Optimization

A woman's clothing retail store chain wanted to boost profits and asked for our help. We interviewed ad hoc 250

Exhibit 2 Drivers of perceived relationship investment



customers in some shops. We asked about attitude and perception that might lead to higher follow-up purchases and deployed USM. When analyzing the survey data we found that perceived relationship investment is a main prerequisite for repetitive purchases. The resulting question would be: How can we increase perceived relationship investment effectively?

With our analysis we showed that excellent interpersonal communication with customers is doing all the work. Expensive "tangible rewards" (especially free gifts such as shoe polish or rebates) are only an alternative but a less effective tool. Interestingly you can see in Exhibit 2 that an increase in tangible rewards does not have any effect when interpersonal communication is excellent (6 or 7). Similarly, an increase in interpersonal communication does result in much less effect when tangible rewards are high compared to the effect when they are low.

As a consequence you can skip the least effective instrument. We suggested exactly this. In addition, personnel selection and leadership processes were reviewed and refined. With these measures we cut 1.5 percent of overall costs, which boosted the profit by almost 30 percent.

This example shows the relevance of interaction effects. Classical methods are only able to model the separate effect of the causes. They miss the fact that the effects of the causes do not—in this case—add up. Therefore they overestimate the effects and fail to see that either one cause is neglectable. Again, with customized econometric models it is possible to model the same relation—but you have to know in advance which causes interact in which functional form.

Why USM Works

Universal structural modeling is a new causal analysis using artificial neural networks that offers the following advantages:

1. Causality: USM uncovers direct causal paths. As a result, by pulling the right success triggers the taken measures result in intended effects.
2. Exploration: USM needs less a priori knowledge about relations at hand.
3. Nonlinearity: USM explores nonlinear relationships (even unknown).
4. Interactions: USM finds, shows and quantifies interactions between causes.
5. Universality: USM makes use of arbitrary distributed variables, especially nominal scaled variables such as gender, profession, brand name, etc.
6. Quantification: USM quantifies every important property—no matter if for path strength, linear path coefficient, interaction strength or significance figures.
7. Simplicity: USM is very easy to use, with no need for detailed option settings.

Ease of use is actually the main fact that convinces many market research companies to apply USM to their data. The user just defines in an Excel sheet which items form a latent variable and which paths should not be considered. The rest is up to the software. Mistakes and insecurity in setting certain options are therefore rare. There is only one thing the software does not claim to solve: the interpretation of its findings. Sound expert knowledge about the topic at hand remains indispensable.

We have to conclude that, despite of its popularity, descriptive analysis is often not suitable for concluding effective measures. Due to spurious correlations, the risk of wrong conclusions is often high when, for instance, just means between groups are compared. Causal (path) analysis methods are needed to handle the risk. A practical method has to be explorative and able to uncover nonlinearities and interactions at the same time.

Universal structural modeling is the first methodical package that can cope with business requirements without compromising the quality of results. The ease of use of the available software puts the method in wide application. Numerous success stories show the significant value USM delivers. In nearly every sizable corporation the deployment of USM can help save millions in costs and create millions in additional profits. ●

Dr. Frank Buckler is founder of NEUSREL Causal Analytics and develops solutions based on neural networks for socio-economic applications for more than 15 years. He employed his methods as a project manager for leading specialized consultancies as Simon, Kucher & Partners. His academic background is electrical engineering and marketing science with a focus on quantitative methods. He may be reached at www.neusrel.com