Chapter 19
Marketing Analytics

S. Arunachalam and Amalesh Sharma

1 Introduction

It is very hard to ignore the potential of analytics in bringing robust insights to the boardroom in order to make effective firm, customer, and product/brand level decisions. Advance analytics tools, available data, and allied concepts have enormous potential to help design effective business and marketing strategies. In such a context, understanding the tools and their various implications in various different contexts is essential for any manager. Indeed, the robust use of the analytics tools has helped firms increase performance in terms of sales, revenues, profits, customer satisfaction, and competition. For details of how marketing analytics can help firms increase its performance, please refer to Kumar and Sharma (2017).

Marketing analytics can help firms realize the true potential of data and explore meaningful insights. Marketing analytics can be defined as a "high technology-enabled and marketing science model-supported approach to harness the true values of the customer, market, and firm level data to enhance the effect of marketing strategies" (Kumar and Sharma 2017; Lilien 2011).

Basically, marketing analytics is the creation and use of data to measure and optimize marketing decisions. Marketing analytics comprises tools and processes

---

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-319-68837-4_19) contains supplementary material, which is available to authorized users.

S. Arunachalam (✉)
Indian School of Business, Hyderabad, Telangana, India
e-mail: S_Arunachalam@isb.edu

A. Sharma
Texas A&M University, College Station, TX, USA

B. Pochiraju, S. Seshadri (eds.), *Essentials of Business Analytics*, International Series in Operations Research & Management Science 264,
https://doi.org/10.1007/978-3-319-68837-4_19

that can inform and evaluate marketing decisions right from product/service development to sales (Farris et al. 2010; Venkatesan et al. 2014). According to CMO survey, spending on analytics will increase from 4.6% to 22% in the next 3 years.[1] This shows the increasing importance of analytics in the field of marketing. Top marketers no longer rely on just intuition or past experience to make decisions. They want to make decisions based on data. But in the same survey, "Lack of processes or tools to measure success through analytics" and "Lack of people who can link to marketing practice" have been cited as the top two factors that prevent marketers from using advanced marketing analytic tools in the real world.

This chapter is a step toward closing these two gaps. We present tools that both inform and measure the success of marketing activities and strategies. We also hope that this will help current and potential marketers get a good grasp of marketing analytics and how it can be used in practice (Lilien et al. 2013; Kumar et al. 2015). Since analytics tools are vast in numbers and since it is not feasible to explain the A–Z of every tool here, we select a few commonly used tools and attempt to give a comprehensive understanding of these tools. Interested readers can look at the references in this chapter to get more advanced and in-depth understanding of the discussed tools.

The processes and tools discussed in this chapter will help in various aspects of marketing such as target marketing and segmentation, price and promotion, customer valuation, resource allocation, response analysis, demand assessment, and new product development. These can be applied at the following levels:

- *Firm*: At this level, tools are applied to the firm as a whole. Instead of focusing on a particular product or brand, these can be used to decide and evaluate firm strategies. For example, data envelopment analysis (DEA) can be used for all the units (i.e., finance, marketing, HR, operation, etc.) within a firm to find the most efficient units and allocate resources accordingly.
- *Brand/product*: At the brand/product level, tools are applied to decide and evaluate strategies for a particular brand/product. For example, conjoint analysis can be conducted to find the product features preferred by customers or response analysis can be conducted to find how a particular brand advertisement will be received by the market.
- *Customer*: Tools applied at customer level provide insights that help in segmenting and targeting customers. For example, customer lifetime value is a forward-looking customer metric that helps assess the value provided by customers to the firm (Fig. 19.1).

Before we move further, let us look at what constitutes marketing analytics. Though it is an ever-expanding field, for our purpose, we can segment marketing analytics into the following processes and tools:

[1] https://cmosurvey.org/2017/02/cmo-survey-marketers-to-spend-on-analytics-use-remains-elusive/ (accessed on Jul 6, 2018).

Fig. 19.1 Levels of marketing analytics application

1. Multivariate statistical analysis: It deals with the analysis of more than one outcome variable. Cluster analysis, factor analysis, perceptual maps, conjoint analysis, discriminant analysis, and MANOVA are a part of multivariate statistical analysis. These can help in target marketing and segmentation, optimizing product features, etc., among other applications.
2. Choice analytics: Choice modeling provides insights on how customers make decisions. Understanding customer decision-making process is critical as it can help to design and optimize various marketing mix strategies such as pricing and advertising. Largely, Logistic, Probit, and Tobit models are covered in this section.
3. Regression models: Regression modeling establishes relationships between dependent variables and independent variables. It can be used to understand outcomes such as sales and profitability, on the basis of different factors such as price and promotion. Univariate analysis, multivariate analysis, nonlinear analysis, and moderation and mediation analysis are covered in this section.
4. Time-series analytics: Models stated till now mainly deal with cross-sectional data (however, choice and regression models can be used for panel data as well). This section consists of auto-regressive models and vector auto-regressive models for time-series analysis. These can be used for forecasting sales, market share, etc.
5. Nonparametric tools: Non parametric tools are used when the data belongs to no particular distribution. Data envelopment analysis (DEA) and stochastic frontier analysis (SFA) are discussed in this section and can be used for benchmarking, resource allocation, and assessing efficiency.
6. Survival analysis: Survival analysis is used to determine the duration of time until an event such as purchase, attrition, and conversion happens. Baseline hazard model, proportional hazard model, and analysis with time varying covariates are covered in this section.

7. Salesforce/sales analytics: This section covers analytics for sales, which includes forecasting potential sales, forecasting market share, and causal analysis. It comprises various methods such as chain ratio method, Delphi method, and product life cycle analysis.
8. Innovation analytics: Innovation analytics deals specifically with new products. New product analysis differs from existing product analysis as you may have little or no historical data either for product design or sales forecasting. Bass model, ASSESSOR model, conjoint analysis can be used for innovation analytics.
9. Conjoint analysis: This section covers one of the most widely used quantitative methods in marketing research. Conjoint (trade-off) analysis is a statistical technique to measure customer preferences for various attributes of a product or service. This is used in various stages of a new product design, segmenting customers and pricing.
10. Customer analytics: In this section, we probe customer metrics such as customer lifetime value, customer referral value, and RFM (recency, frequency, and monetary value) analysis. These can be used for segmenting customers and determining value provided by customers.

This chapter provides a comprehensive understanding of the marketing analytics field. We explain a part of it through basic and advanced regression, data envelopment analysis, stochastic frontier analysis, innovation analytics, and customer analytics.

Each section in this chapter has marketing related problems and solutions to reinforce concepts and enhance understanding. We describe how various tools can be implemented for a particular objective. We also discuss limitations of the tools under consideration. You should familiarize yourself with all the processes and tools so that you can choose appropriate ones for your analysis.

2 Methods of Marketing Analytics

2.1 Regression

Interaction Effect

This section covers the steps for conducting interactions between continuous variables in multiple regression (MR). We provide an intuitive understanding of this concept and share a step-by-step method for statistically conducting and probing an interaction effect. Chapter 7 on linear regression analysis contains detailed explanation of the theory underlying MR.

Interaction effects, also called moderator effects, are the combined effects of multiple independent variables on the dependent variable. They represent situations where the effect of a variable on a dependent variable is contingent on another variable (Baron and Kenny 1986). Let us consider a hypothetical example where a manager is trying to understand the effect of advertisement and price on sales of a product. It can be argued that price and advertisement may have interaction effect on sales, that is, when both advertisement and price change it may have a different effect on product sales as compared to when either advertisement or price changes independently. To capture such effects, we require the following MR equation (error is omitted for ease of presentation in all equations):

$$\mathrm{Y} = i_0 + a\,\mathrm{X} + b\,\mathrm{Z} + c\,\mathrm{XZ} + \mathrm{Error} \tag{19.1}$$

where

Y = continuous dependent variable (sales in the above example)

X, Z = continuous independent variables (advertisement and price in the above example)

XZ = new variable computed as the product of X and Z (product of advertisement and price in the above example)

i_0 = intercept

a, b, and c = slopes

An important aspect to remember in specifying an equation with interaction terms is that the lower order terms should always be present in that equation. That is, it is incorrect to test for interaction by omitting X and Z and just having the XZ term alone. Here we assume that the independent variables, X and Z, are mean-centered. To mean-center a variable the value of a variable is subtracted by the mean of the variable in the sample (mean-centered variable = variable – mean [variable]). The product term XZ is computed after mean-centering both X and Z. The dependent variable, Y, is generally not mean-centered. Mean-centering helps in ease of interpretation of effects and also removes nonessential multicollinearity. More details about benefits of mean-centering (it is very useful to know) can be found from the book Aiken and West (1991) and from other sources mentioned in the references section (Cohen et al. 2003).

How does the presence of the interaction term, XZ, help in addressing the managerial question above? Why is it needed and why cannot interaction effects be uncovered using the simple MR equation: Y = i_0 + a X + b Z that does not contain the interaction term, XZ? To understand this let us start with the simple MR equation that does not have the XZ term. Here the slopes "a" and "b" are the effect of X and Z on Y, respectively. The regression of Y on X (i.e., slope "a") is constant across legit (explained below) values of Z and vice versa for regression of Y on Z (i.e., slope "b"). This means that the regression of Y on X is independent of Z because for any value of Z, value of the slope would be "a." Hence, this does not answer the managerial question asked above as it does not capture the variation in the effect of X according to Z and vice versa. However, this is not the case if we consider Eq. (19.1) that has the interaction term XZ. To understand why and how let us rewrite Eq. (19.1) in this way:

$$\mathrm{Y} = (\boldsymbol{a} + \boldsymbol{c}\,\mathbf{Z})\ \mathrm{X} + (\boldsymbol{i_0} + \boldsymbol{b}\,\mathbf{Z}) \tag{19.2}$$

This expanded view of Eq. (19.1) helps answer the two questions we raised above. It is seen from Eq. (19.2) that the slope of the regression of Y on X is now the term (***a*** + ***c*** **Z**), unlike just "a" from the simple equation. This term **(*a* + *c* Z)** is called simple slope in the moderation literature as the effect of Y on X is now conditional on the value of Z! Therefore, the effect is now dependent on Z rather than being independent as noted above. To go back to our managerial example this would mean that effect of advertisement (X) on sales (Y) is dependent on price (Z). This is precisely what we would like to achieve or test. This also means that for every value of Z, regression of Y on X would have a different line, which is called a simple regression line. Equation (19.2) could also be rewritten by pulling out Z and having a simple slope that is (b + c X), which would help understand how the effect of price (Z) on sales (Y) is dependent on advertisement (X).

Estimation: Estimating Eq. (19.2) using any statistical software like *STATA (or R)*, is straightforward. The command *regress (lm)* can be used to estimate the equation after mean-centering the independent variables and creating the XZ term. Interpretation and probing of interaction effects need more understanding (Preacher et al. 2006). Let us assume that the slope "c" of XZ term is statistically significant (at $p < 0.05$). What does this mean? How to interpret this significant effect?

There are two very important takeaways when the slope "c" is significant. First, it means that simple slopes for any pair of simple regression lines obtained using two different legit values of Z are statistically different. Second, for any specific value of Z the researcher has to test whether the simple regression line is significantly (i.e., statistical significance) different from zero. These two points are better understood and observed by plotting graphs of the interaction. We recommend that the researcher should always plot graphs to thoroughly understand the interaction effects. Before we probe these effects through graphs let us understand the meaning of "legit values" of Z. Let us recall that we have mean-centered the variables X and Z and hence the mean of these two variables would be zero and the range would be from negative to positive values around zero. It is always a good practice to choose two values of Z—one low value and another high value as one standard deviation below and above the mean (which is zero now) respectively, to plot the simple regression lines. The reason is that these two values would be within the range of values of Z in our sample. This is important because we should not choose a value that is not within the sample range. The researcher is free to choose any two values that are within the range of Z. If we choose an arbitrary value and if that value is outside of the range of Z, then we are testing the effects for out-of-sample data, which is incorrect. Therefore, researchers should pay attention in choosing legit values of Z.

Let us assume we choose two legit values of Z, namely, Z-low (Z_L) and Z-high (Z_H). So the simple slopes from Eq. (19.2) that determine the simple regression lines are ($a + c\ Z_L$) and ($a + c\ Z_H$). Therefore, we can plot two regression lines by substituting the two values of Z in Eq. (19.2). These two lines would then represent the effect of advertisement (X) on sales (Y) when the price is low (Z_L) and when the price is high (Z_H). Deriving from Eq. (19.2) the simple regression equations are:

$$\begin{aligned}&\text{Sales for lower price}: \mathrm{Y_L} = (\boldsymbol{a} + \boldsymbol{c}\mathbf{Z_L})\,\mathrm{X} + (i_0 + b\mathrm{Z_L})\\&\text{Sales for higher price}: \mathrm{Y_H} = (\boldsymbol{a} + \boldsymbol{c}\mathbf{Z_H})\,\mathrm{X} + (i_0 + b\mathrm{Z_H})\end{aligned} \tag{19.3}$$

Let us assume we have estimated Eq. (19.2) from a sample dataset that contained values for Y, X, and Z and that we have this prediction equation (intercept i_0 is omitted for pedagogical reason; X and Z are mean-centered):

$$\widehat{\mathrm{Y}} = 0.12\,\mathrm{X} + 0.09\mathrm{Z} + 0.46\mathrm{XZ} \tag{19.4}$$

The parameters "a," "b," and "c" are statistically significant ($p < 0.05$). Please note it is not necessary for "a" and "b" to be significant to probe the interaction effect. It is however important to have "c," the estimate for the interaction term XZ, to be statistically significant to probe the interaction effect (refer above for the two important takeaways). Let us also assume that one standard deviations of X and Z are 0.5 and 0.5. The following steps help in plotting the graphs.

We calculate the predicted value of Y (sales) for the two legit values of X and Z. So we take low and high values of X and Z and plug it in Eq. (19.4). As one standard deviation of mean-centered X and Z are 0.5 and 0.5, legit low and high values would be −0.5 and 0.5, respectively. When these values are used to compute the predicted sales (Y) for low advertisement ($\mathrm{X_L}$) and low price ($\mathrm{Z_L}$), Eq. (19.4) becomes:

$$\widehat{\mathrm{Y}} = 0.12 * (-0.5) + 0.09 * \ (-0.5) + 0.46 * \ (-0.5 * -0.5)$$

This leads to predicted sales of 0.01 at low advertisement and low price. Similarly, predicted values for all other combinations of X and Z can be obtained. We recommend using an excel sheet to compute this 2 * 2 table as shown below (Table 19.1).

Once we have the above table, it becomes easy to plot the two regression lines corresponding to low and high values of price (Z). The two rows in the table are the two regression lines. The graph can be plotted as a line graph in excel (Fig. 19.2).

(Excel approach: To plot, click the "Insert" option and select line graph and choose the two data points for low and high to get the two regression lines as shown below.)

Interpretation of the graph: The line titled "High Price" shows the effect of advertisement on sales when the price is also increased to a high value. The line shows that the slope is positive; that is, when price is increased, as advertisement increases, the effect on sales is positive. The line titled Low Price shows the effect

Table 19.1 Predicted sales values

Predicted sales value in each cell			
		X (Advertisement)	
		Low	*High*
Z (Price)	*Low*	0.01	−0.1
	High	−0.13	0.22

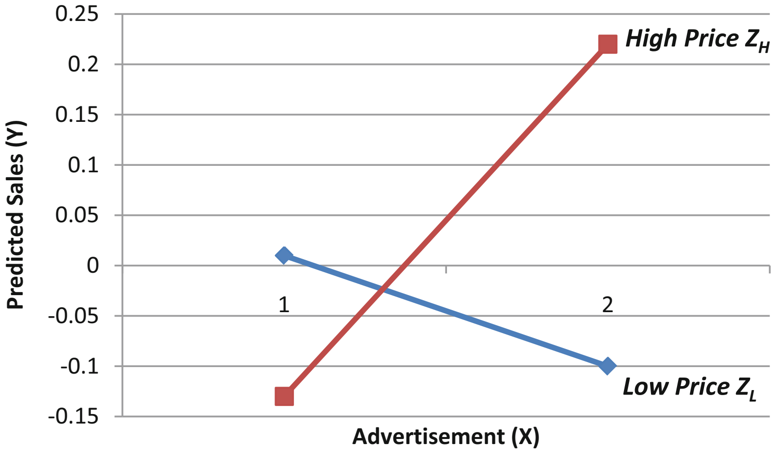

Fig. 19.2 Regression lines corresponding to price (Z)

of advertisement on sales when price is decreased to a low value. Interestingly, the slope of the line is negative; that is, when price is decreased to a low value, as advertisement increases, the effect on sales is negative. These two plots are important takeaways for a manager as it shows that price and advertisement are complements to each other (needless to say, this effect is for the sample dataset we have). Therefore, a manager should not increase advertisement and decrease price as this causes lower sales (the regression line for Z_L is negative).

This way we can test and understand interaction effects thoroughly. Effects in marketing studies are rarely just main effects. Most often, variables interact to show complex effects on the outcome variable of interest. We have shown one way to interpret this effect for interaction of two continuous variables. Similar steps can be undertaken for interaction between a categorical variable and a continuous variable as well. This strategy can also be extended to three-way interaction effects, that is, effects where three variables interact to produce differential effect on the outcome. These are advanced concepts but can be directly extended using the steps narrated above. We recommend interested readers to learn more about these advanced concepts by following the materials in the references.

Curvilinear Relationships

Many a time, in marketing studies, relationships between independent variables (X) and dependent variables (Y) are complex instead of linear. Most often, they represent a curvilinear relationship and studies tend to hypothesize a U-shaped or an inverted U-shaped effect. Rather than thinking about curvilinear relationships as just U-shaped (or inverted U), it would help to appreciate curvilinear relationships according to the effect of X on Y. The effect of X on Y could be (1) increasing at an increasing rate, (2) increasing at a decreasing rate, or (3) decreasing at a decreasing rate, thus altering the slope away from a simple linear trend.

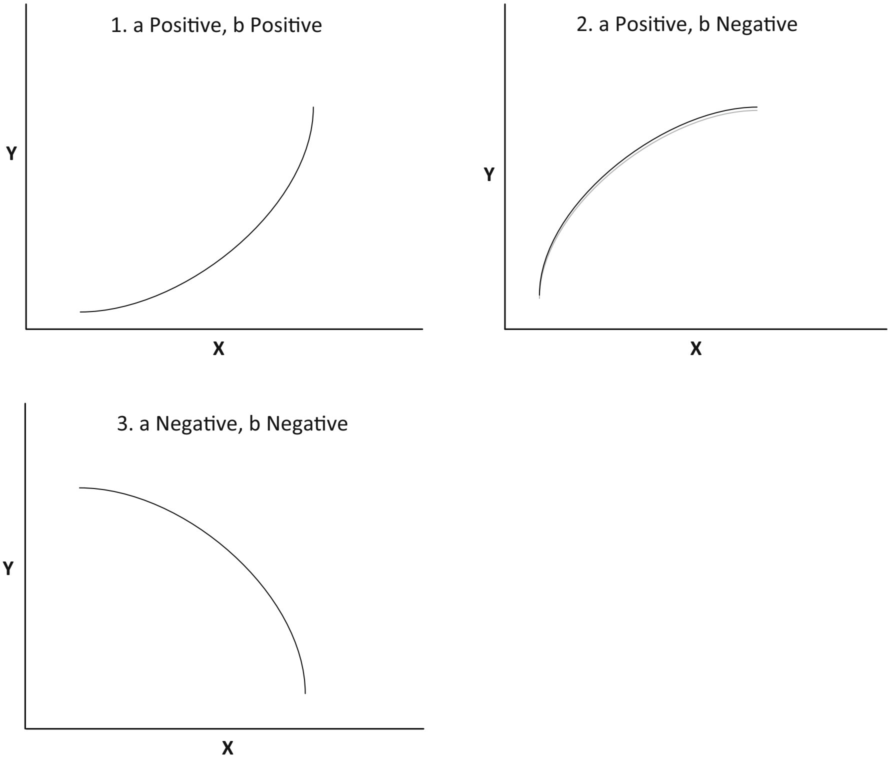

Fig. 19.3 Curvilinear relationship between variables

The three relationships mentioned above can be better understood as a regression equation. For testing any hypothesized curvilinear relationship between X and Y, the following equation helps:

$$\mathrm{Y} = i_0 + a\,\mathrm{X} + b\,\mathrm{X}^2 \tag{19.5}$$

(Note: X is mean-centered)

As noted in the section on interaction effect, lower order term (i.e., X) should be present when a higher order term, formed as product of X*X, is present. The magnitude and the sign of slopes "a" and "b" of Eq. (19.5) contain a wealth of information that helps the researcher in understanding what could be the shape of this specification. Refer Fig. 19.1 to understand the expected shapes depending on the sign (Fig. 19.3).

Going back to the example of the effect of advertisement and price on sales (Y); let us now consider (for pedagogical reason) only the effect of advertisement (X) on sales (Y). A manager could argue that as advertisement increases, product sales tend to increase. Also, beyond a point the effect of advertisement on sales might not increase at the same rate. The manager is actually talking about a curvilinear relationship between X and Y, wherein the effect of X on Y increases, but beyond a point it increases at a decreasing rate. This leads to a prediction line as depicted in Graph (19.3, 2) above, wherein slope "a" is positive and "b" is negative.

As we learnt in the interaction section, we could try to compute the simple slope of Eq. (19.5) by restructuring it. However, this method is incorrect and cannot be used for equations that have curvilinear effects. We have to use simple calculus to derive the simple slope. As we all know, the first partial derivative with respect to (w.r.t) X is the simple slope of regression of Y on X. This technique can be applied to the interaction equation we dealt with earlier as well. So, going back to Eq. (19.5), the first partial derivative is:

$$\partial\, \mathrm{Y}/\partial\, \mathrm{X} = a + 2\, b\, \mathrm{X} \tag{19.6}$$

As seen in Eq. (19.6), the simple slope depends or is conditional on the value of X itself. This along with the sign of "a" and "b" is precisely the reason for the effect of X on Y being curvilinear and not just linear. Again, going to back to our steps in choosing legit values of X, we can take one standard deviation below and above the mean (which is zero) and compute the values of the simple slope.

If a manager would like to investigate how price interacts with the curvilinear relationship of advertisement on sales, we can follow the same steps we did in interaction. However, to derive the simple slope, we have to use calculus and not just simple restructuring of the equation for reasons stated above. Interaction effects on curvilinear relationship are complex and advanced concepts. So we urge the interested readers to peruse the resources in the reference section for a complete understanding. We provide some ideas to help understand some basics of interaction effects in curvilinear relationships below.

The full equation after including price (Z) would be:

$$\mathrm{Y} = i_0 + a\, \mathrm{X} + b\, \mathrm{X}^2 + \mathrm{c}\, \mathrm{Z} \tag{19.7}$$

Now unlike simple interaction effects, researchers have to think deeply on whether price (Z) interacts with just advertisement (X) or Z interacts with the higher order term (X^2) as well. Conceptually, these two effects are different and so the researcher has to specify the correct equation based on the hypothesis being tested. We should also remember to always include the lower order terms while trying to introduce higher order terms that have interaction effects. Let us consider one example, where we introduce XZ term. Then Eq. (19.7) will change to the following:

$$\mathrm{Y} = i_0 + a\, \mathrm{X} + b\, \mathrm{X}^2 + \mathrm{c}\, \mathrm{Z} + d\, \mathrm{XZ} \tag{19.8a}$$

To derive the simple slope, which shows the effect of X on Y, we take the first partial derivative w.r.t X:

$$\partial\, \mathrm{Y}/\partial\, \mathrm{X} = a + 2\, b\, \mathrm{X} + d\, \mathrm{Z} \tag{19.8b}$$

It can be seen from the above equation that not only does the effect depend on X (due to the presence of the term 2 *b* X) but it also depends on Z (due to the term *d* Z). This means that the curvilinear effect is conditional on Z as well.

Furthermore, if a manager hypothesizes that Z alters both the level and the shape of the curvilinear effect, it can be tested by introducing the X^2Z term. This term is the product of X^2 and Z. Then Eqs. (19.8a and 19.8b) will change to the following:

$$\mathrm{Y} = i_0 + a\, \mathrm{X} + b\, \mathrm{X}^2 + \mathrm{c}\, \mathrm{Z} + d\, \mathrm{XZ} + \mathrm{e}\, \mathrm{X}^2\mathrm{Z} \tag{19.9a}$$

Please note the presence of all lower order terms (which are X, X^2, Z, and XZ) of the highest order term (X^2Z). Neglecting any the lower order terms leads to incorrect interpretation of equation (19.9a and 19.9b). Again, to calculate the simple slope we do the following:

$$\partial\, \mathrm{Y}/\partial\, \mathrm{X} = a + 2\, b\, \mathrm{X} + d\, \mathrm{Z} + 2\, e\, \mathrm{XZ} \tag{19.9b}$$

This shows that Z affects not just the level but the shape of the curve as well.

Mediation

Mediation is used to study relationships among a set of variables by estimating a system of equations. The objective of a mediation analysis is to extract the mechanism behind the effect of X on Y. Mediation analysis is useful in understanding the intervening mechanism that actually causes the effect of X on Y (MacKinnon 2008; Preacher and Hayes 2004). To find this intervening mechanism, an intervening variable called the mediator (M) is introduced. We can imagine this mediation as a pathway: X→M→Y. Therefore, the total effect of X on Y is now partitioned into an indirect effect via M and a direct effect of X on Y.

In mediation analysis, there is a series of regressions or a set of equations that has to be estimated. First is the effect of X on the mediator M (X→M, slope *a*), next is the effect of mediator on the dependent variable Y (M→Y, slope *b*), and finally the direct effect of X on Y (X→Y. slope c'). The researcher should note that now there are two dependent variables, namely, Y and the newly introduced intervening mediator M. It can be easily understood through the diagram given below (Fig. 19.4).

The diagram (Fig. 19.4) can be represented by the following set of regression equations:

$$\mathrm{Y} = i_0 + b\, \mathrm{M} + c'\, \mathrm{X} \tag{19.10}$$

$$\mathrm{M} = j_0 + a\, \mathrm{X} \tag{19.11}$$

where i_0 *and* j_0 are intercepts and are ignored hereafter for pedagogical purpose.

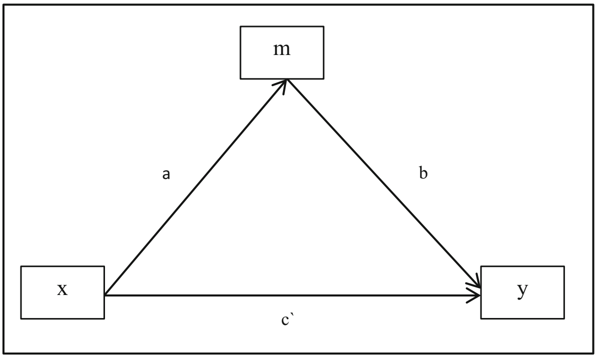

Fig. 19.4 Mediation effect through "M"

Now the mediating effect, also called the indirect effect of X on Y via M, can be obtained by inserting Eq. (19.11) in Eq. (19.10):

$$Y = ab\,X + c'\,X \tag{19.12}$$

This shows that the indirect effect of X on Y is the product of two slopes captured as "a $*$ b" and the direct effect of X on Y is "c$'$."

The slopes "a," "b," and "c$'$" can be estimated using any statistical software and the product of the slopes "a $*$ b" can be computed to find the strength of the mediation. However, we recommend using software that supports path analysis or structural equation modeling (SEM) techniques (STATA or Mplus or R) to estimate both the equations simultaneously without having to run multiple regressions. This practice can help the researcher easily progress from the simple model narrated here to more complex mediational models that are often the case in research. One important question still remains—how to test for the statistical significance of the product term "a $*$ b"? The statistical test of the simple slopes "a," "b," and "c$'$" using standard errors is straightforward. However, for the product of the slopes, recent research in mediational analysis suggests that deriving the standard error (using a technique called delta method) for a higher order product term is inaccurate and that statistical significance has to be tested using nonparametric techniques like bootstrapping. The intuition behind this is that though the individual slopes "a" and "b" are assumed to be normally distributed, the product of these slopes cannot be. Hence, as the derivation of standard errors through parametric methods (like OLS or maximum likelihood) assumes normality, researchers are advised to use nonparametric resampling procedures like bootstrapping to derive the standard errors. The logic behind bootstrapping procedure is straightforward. The estimates of the indirect effect "a $*$ b" are repeatedly obtained from samples drawn from the original sample with replacement. The standard deviation of those estimates is the standard error, which is then used to build a nonsymmetric 95% confidence interval to test the significance of the indirect effect. If the 95% confidence interval of the indirect effect does not contain zero the effect is statistically significant. In this case, we can conclude that there is a significant mediation effect.

Let us go back to the example of advertisement (X) having an effect on product sales (Y). Now a manager might be interested in knowing the reason behind the "increase in advertisement having a positive effect on sales." One could argue that, as advertisement increases (X), customers' awareness (M) about the product also increases, which encourages them to buy more, leading to an increase in product sales (Y). Therefore, here the mediator is consumer awareness (M). If somehow the manager can measure or capture this variable, it can be used, along with the X and Y variables, to test this hypothesized mediation effect. But what is the managerial implication? What really is the "extra" insight derived by doing this mediation analysis? An important insight is that the manager is now able to identify one relevant mechanism that actually causes the effect of X on Y. The manager can then increase resource allocation to marketing strategies that help improve consumer awareness levels. Furthermore, this manager could be intrigued to investigate whether there are other mediators or intervening mechanisms like awareness. Understanding and finding the reasons behind effects are critical to decision making. Armed with the knowledge of the mechanism behind the cause-and-effect relationship, executives can make informed decisions which lead to greater success.

In this section, we have covered simple mediation analysis. Interested readers are strongly recommended to consult the resources provided in the reference section for advanced mediation topics like mediation analysis with multiple mediators and multiple dependent variables. One interesting and important advance in mediation analysis in recent times has been the ability to integrate moderation with mediation analysis. This is called moderated mediation analysis (Preacher et al. 2007). This is an important advancement because it resonates well with the purpose of theory testing in marketing or in general any social science domain. Theory is important primarily for two reasons: to provide arguments to explain the reason behind why and how a variable affects another and to uncover the boundary condition under which this effect could change. This is essentially the spirit of any moderated mediation test as it helps to simultaneously understand the mechanism (mediation) and the conditional effect (moderation) of that mediation pathway. That is, to investigate if the indirect effect of '$X \rightarrow M \rightarrow Y$' as a unit is conditional on the level of a moderator variable, say Z. This would imply how the product term "*ab*" is conditional on legit values of Z (say low and high levels of Z). For example, considering the example of how "*consumer awareness*" *(M)* channelizes the effect of advertisement (X) on product sales (Y), a manager can also understand how the strength of this pathway varies depending on whether the product price (Z) is low or high. As the mediation pathway is conditional on another variable, or mediation effect is moderated by another variable, this test is termed moderated mediation test. The theory and empirics behind moderated mediation has exploded in recent times. Readers are requested to consult the references for a deeper understanding of this technique.

2.2 Data Envelopment Analysis (DEA)

Suppose you are a manager of a big MNC and you want to identify the efficient units (finance, marketing, operations, HR, international business, etc.) within your organization. You may want to decide on allocating resources, hiring, increasing the benefits to the employees of a specific unit, based on the efficiency of the unit. How are you going to decide on the efficiency? You may use observational data for such decisions. However, such an approach may not be great while designing critical strategies such as resource allocation or hiring.

A very prominent analytic tool that can be used in such decision making is "data envelopment analysis" (Cook and Seiford 2009; Cooper et al. 2004). DEA is designed to help managers measure and improve the performance of their organizations. As the quest for efficiency is a never-ending goal, DEA can help capture the efficiency of each unit and suggest the potential factors that may be responsible for making units efficient.

DEA allows managers to take into account all the important factors that affect a unit's performance to provide a complete and comprehensive assessment of efficiency (Charnes et al. 1978). DEA does this by converting multiple inputs and outputs into a single measure of efficiency. By doing so, it identifies those units that are operating efficiently and those that are not. The efficient units, that is, units making best use of input resources to produce outputs, are rated as being 100% efficient, while the inefficient ones obtain lower scores.

Let us understand the technical aspects of DEA in a simple way. Suppose, a manager of ICICI bank wants to measure the efficiency of HR, finance, marketing, operation, public relation, and accounting departments to allocate resources for the following year (Sherman and Franklin 1985). Managers generally have information regarding the number of employees, total issues handled, raises given, and the contribution of these units toward the operational profits of the firm. Managers can use DEA with number of employees, total issues handled, and raises given as the inputs and contribution of these units toward the operational profits as the output. DEA analysis will provide a score of either 1 or less than 1 to each unit. A score of 1 means the unit is perfectly efficient and anything less than 1 means that the unit has the room to grow and be more efficient. Based on the relative importance of the inputs, units may be asked to work on specific inputs such that the units become efficient. Further, managers can now allocate the resources they want to the efficient units.

DEA also provides guidance to managers on the reduction/increment required in the inputs of the units to become efficient, helping managers answer questions, such as "How well are the units doing?" and "How much could they improve?" It suggests performance targets, such as marketing unit should be able to produce 15% more output with their current investment level or HR unit should be able to reduce churn by 25% and still produce the same level of outputs. It also identifies the best performing units. One of the most interesting insights from DEA is that one can test the operating practices of such units and establish a guide to "best practices" for others to emulate.

Let us understand using a practical example. Suppose a bank such as ICICI wants to measure the efficiencies of their branches across India. The goal of the management is to ensure that each of these branches achieves the best possible performance; the problem though is deciding what that means and how best to go about measuring it. The outputs of the branches such as sales, sales growth, accounts, and market share can be studied and compared. Similarly, a branch's inputs, such as staff, office space, and materials costs can be measured. Managers can then develop performance ratios such as sales per member of staff or profit per unit of office space utilized.

However, all these attempts to measure performance may not produce a clear picture as branches may exhibit considerable variation depending on the performance indicator chosen. This is where DEA helps as it provides a more comprehensive measure of efficiency by taking into account all the important factors that affect a branch's performance.

Although the choice of inputs and outputs largely depends on the availability of the data, the beauty of DEA lies in its ability to handle multiple inputs and multiple outputs to give a comprehensive score about the efficiency of the units. DEA can be used within a specific organization (e.g., measuring the efficiency of various divisions of ICICI bank); across organizations in an industry (e.g., comparing the efficiency of all the private sector banks in India); and across industries (e.g., measuring efficiency of banks, FMCG companies). However, such analysis will require availability of similar inputs and outputs for all organizations.

Technical Details of DEA

DEA can be analyzed either as constant return to scale (CRS) or variable return to scale (VRS) (Banker and Thrall 1992; Banker et al. 2004; Seiford and Zhu 1999). Let us first consider the CRS model. Let us assume that there are N decision-making units (DMUs) (DMUs are various banks in our example) with K inputs and 2 outputs (profits and customer satisfaction). These are represented by x_i and y_i for *i*th DMU. The purpose of DEA is to construct a nonparametric envelopment frontier over the data points such that all observed points lie on or below the production frontier (Charnes et al. 1985; Boussofiane et al. 1991; Chen and Soo 2010). We can express the problem as a linear program as follows:

$$\begin{aligned}&\max_{\theta,\lambda}\theta\\&\text{such that } -\theta y_i + Y\lambda \geq 0, \lambda \geq 0\\&\quad x_i - X\lambda \geq 0, \lambda \geq 0\end{aligned}$$

X is the K x N input matrix and Y is 2 x N output matrix; θ is a scalar and λ is an N x 1 vector of constants. The efficiency score for *i*th DMU will be the value of ($1/\theta$). The DMU with efficiency score = 1 will be called efficient DMU. To get the efficiency score for each DMU (i.e., various divisions of ICICI bank) in the sample, we will need to solve the linear programing for the specific DMU under consideration. Note that we can use CRS DEA when the underlying assumption is that the DMUs are operating at an optimal scale. However, this assumption might not hold in reality, most DMUs in a sample may not work at the optimal scale. In such a context, CRS may not be an ideal tool. To avoid any potential scale effects, one may use VRS. One can compute the technical efficiency with VRS that excludes the scale effects. The CRS linear programming problem can be modified to account for VRS by adding the convexity constraint $e_\lambda = 1$ (where e_λ is N X 1 vector of ones). The additional constraints give the frontier piecewise linear and concave characteristics.

$$\begin{aligned} &\max_{\theta,\lambda}\theta \\ &\text{such that } \; -\theta y_i + Y\lambda \geq 0, e_\lambda = 1, \lambda \geq 0 \\ &\qquad\qquad\qquad x_i - X\lambda \geq 0, e_\lambda = 1, \lambda \geq 0 \end{aligned}$$

The CRS and VRS provide different values for the efficiency score of a DMU; it indicates the presence of scale inefficiency in the DMU. One can compute the scale inefficiency by taking the ratio of efficiency score obtained from CRS to the efficiency score obtained from VRS (Caudill et al. 1995; Gagnepain and Ivaldi 2002; Greene 2010).

Example

Suppose that we are interested in evaluating the efficiency of the hospital units (Ouellette and Valérie 2004) of a chain based on a number of characteristics: the total number of employees, the size of units in square meters, the number of patients each unit serves, total number of specialists, total revenue, and satisfaction. It becomes obvious that finding the most efficient units requires us to compare records with multiple features.

To apply DEA, we must define our inputs (X) and outputs (Y). In the case of a hospital chain, X can be the total number of employees, the size of units in square meters, the number of patients each unit serves, total number of specialists; and Y can be total revenue and satisfaction. If we run DEA, we will estimate the output to input ratio for every hospital under the ideal weights (ideal weights are weights that consider the values that each unit puts on inputs and outputs). Once we have their ratios, we will rank them according to their efficiency (Banker and Morey 1986).

STATA/R Code

DEA can be analyzed using multiple statistical programing software. Here, we are providing the syntax required to conduct the analysis in STATA[2] or R[3]. Although STATA does not have a built-in function, one can use user-written command (dea) to do the analysis. In case you find it difficult, please type “help dea” in STATA command window and you will get a step-by-step explanation for the analysis.

[2]https://www.cgdev.org/sites/default/files/archive/doc/stata/MO/DEA/dea_in_stata.pdf (accessed on Jan 30, 2019).

[3]https://www.rdocumentation.org/packages/TFDEA/versions/0.9.8.3/topics/DEA (accessed on Jan 30, 2019).

To do the analysis in STATA, you need to download user-written command (type `net install st0193`).

```
"dea ivars=ovars [if] [in] [, options]"
```

Options:

- rts(crs|vrs|drs|nirs) specifies the returns to scale. The default is rts(crs)
- ort(in|out) specifies the orientation. The default is ort(in)
- stage(1|2) specifies the way to identify all efficiency slacks. The default is stage (2)

To do the analysis in R, you need to use "*Benchmarking*" package.

```
> install.packages("Benchmarking")  # to install package for the
  first time
> library(Benchmarking)   # to load package
# DEA code
> eff <- dea(x,y, RTS="crs")   # where x is input vector and y
  is output vector
> eff   # where "eff" will give the efficiency of each unit
# RTS options allow us to specify which return of scale we want
  •   "crs" - constant return to scale
  •   "vrs" - varying return to scale
  •   "drs" - decreasing return to scale
  •   "irs" - increasing return to scale
```

DEA in Practice

There are eight units of a restaurant chain (largely in Northern India) and the manager wants to measure the efficient unit for the best restaurant award. The manager has insights only on the total number employees and the revenue (in 100,000) of each unit as follows:

Units	1	2	3	4	5	6	7	8
Employee	5	6	11	15	20	9	12	10
Revenue	4	4.8	8	14	18	5	11	9

In order to find the efficiency of each unit, we first calculate the efficiency by dividing the revenue by employee as follows:

Units	1	2	3	4	5	6	7	8
Employee	5	6	11	15	20	9	12	10
Revenue	4	4.8	8	14	18	5	11	9
Revenue/employee	0.8	0.8	0.727	0.93	0.9	0.55	0.91	0.9

As per the efficiency measurement, we find that Unit 4 is the efficient one and Unit 6 is the least efficient one.

Now, to compute the relative efficiency, we need to divide the efficiency of the units by the efficiency of the most efficient units, that is, the relative efficiency of units is measured by taking the ratio of efficiency of each unit and the efficiency of most efficient unit as shown below.

$0 \leq$ revenue *per employee for each* unit/revenue *per employee for the most efficient* unit ≤ 1.

Units	1	2	3	4	5	6	7	8
Employee	5	6	11	15	20	9	12	10
Revenue	4	4.8	8	14	18	5	11	9
Revenue/employee	0.8	0.8	0.727	0.93	0.9	0.55	0.91	0.9
Relative efficiency	0.860215	0.860215	0.78172	1	0.967742	0.591398	0.978495	0.967742

Data envelopment analysis shows that the Unit 4 is the most efficient relative to all other units; and the Unit 6 is the least efficient. That means Unit 4 will be on the frontier and rest will be within the frontier. One can also show the frontier pictorially.

2.3 Stochastic Frontier Analysis (SFA)

Stochastic frontier analysis is a parametric approach, largely used to estimate the production or costs in economics (Baccouche and Kouki 2003; Fenn et al. 2008). Data envelopment analysis and SFA are competing approaches (Jacobs 2001; Cullinane et al. 2006); however, there is no single approach that unifies both. If a manager wants to find what causes the inefficiency in a firm in its operation, the manager may want to adopt SFA. SFA relies on the assumption that decision-making units (such as banks and hospitals) behave suboptimally and they can maximize or minimize their respective objective functions (costs, profits, operational efficiency, etc.) in order to improve (Parsons 2002).

Let us now discuss the components of an SFA (Kumbhakar and Lovell 2003; Bera and Sharma 1999). The stochastic production frontier was first proposed independently by Aigner et al. (1977) and Meeusen and Van den Broeck (1977). In their specification (which is largely different from the standard production function), there are two distinct error terms in the model and can be shown as follows:

$$y_i = \beta x_i - u_i + v_i$$

where x_i are inputs, y_i is the output; u_i captures inefficiency, shortfall from maximal output govern by production function βx_i; and v_i is the error term or the outside influences beyond the control of the producer. In sum, the SFA has two components: a stochastic production frontier serving as a benchmark against which firm efficiency is measured, and a one-sided error term with independent and identical distribution across observations and captures technical inefficiency across production units.

If a manager allows the inefficiencies to depend on the firm level factors (employees, experts, size, alliances, acquisitions, etc.), the manager can examine the determinants of inefficiencies. Such understanding helps to implement different policy interventions to improve efficiency. Managers can also modify the inputs and incorporate in the production function to reduce the inefficiencies.

Now the question is whether one should use SFA over DEA? (Wadud and White 2000; Koetter and Poghosyan 2009). Imagine that there are random variations in the inputs. These variations can make the DEA analysis unstable. A potential advantage of the SFA over DEA is that random variations in inputs can be accommodated.[4] Although there are some benefits, SFA also suffers from multiple disadvantages including its complications in handling multiple outputs. Further, it also requires stochastic multiple output distance functions, and it raises problems for outputs that take zero values.

Technical Details of SFA

The technological relationship between a few input variables and corresponding output variables is given by a "*production function*" (Aigner et al. 1997). Econometrically, if we use data on observed outputs and inputs, the production function will indicate the average level of outputs that can be produced from a given level of inputs (Schmidt 1985). One can estimate production functions at either an individual or an aggregate level.

The literature on production function suggests that the implicit assumption of production functions is that all firms are producing in a technically efficient manner, and the representative (average) firm therefore defines the frontier (Førsund et al. 1980; Farrell and Fieldhouse 1962). The estimation of the production frontier assumes that the boundary of the production function is defined by "best practices" units. It therefore indicates the maximum potential output for a given set of inputs along a straight line from the origin point. The error term in the SFA model represents any other reason firms would be away from (within) the boundary. Observations within the frontier are deemed "inefficient (Hjalmarsson et al. 1996; Reinhard et al. 2000)."

The SFA Model and STATA/R Code

Restating the stochastic frontier model:

$$y_i = \beta x_i - u_i + v_i, u_i = |\mathrm{U}|$$

where |U| is the aggregated (in)efficiency.

[4] http://www.fao.org/docrep/006/Y5027E/y5027e0d.htm (accessed on Jul 6, 2018).

In this area of study, estimation of the model parameters is usually not the primary objective. Estimation and analysis of the inefficiency of individuals in the sample and of the aggregated sample are usually of greater interest.

STATA has a built-in command to estimate SFA.[5] This is essentially a regression analysis where the error term consists of a random error and an inefficiency term, and again can be estimated for both production and cost functions. The syntax of the frontier command is:

```
frontier depvar [indepvars] [if] [in] [weight] [, options]
```

The most important options are:

Distribution (`distname`) specifies the distribution for the inefficiency term as half-normal (`hnormal`), `exponential`, or truncatednormal (`tnormal`). The default is `hnormal`.

`cm (varlist)` fits conditional mean model; may be used only with distribution (`tnormal`). `uhet(varlist)` explanatory variables for technical inefficiency variance function

`vhet(varlist)` explanatory variables for idiosyncratic error variance function

`cost fit` cost frontier model; default is production frontier model

`vce(vcetype)` vcetype may be `oim`, `opg`, `bootstrap`, or `jackknife`.

To do the analysis in R, you can use "Benchmarking" package and sfa function as below.

```
> sfa(x,y)       # where x is a n x k matrix of k inputs
for n units, and y is a n x 1 vector of outputs for n units
> summary(sfa(x,y))
> e <- sfa(x,y)              # Estimate efficiency for each unit
> eff(e)
```

SFA in Practice

A manager can find the effect of inefficiency and the drivers of inefficiency by adopting an SFA approach. For example, given the availability of the data on the improper implementation of the marketing-mix elements, and the within-unit corruption, a manager can model the inefficiency as a function of improper implementation and corruption and see their effect on the overall productivity (e.g., revenue contribution) of each unit. In marketing, there has been several applications of SFA; see, for example, Feng and Fay (2016) who apply it to evaluate salesperson capability, Parsons (2002) who comments upon its application to sales people and retails outlets and observes that unlike explaining just the mean performance SFA provides explanation for the gap between the unit and the nest performer, and Vyt (2008) who uses SFA for comparing retailers' geo-marketing.

[5] https://www.stata.com/manuals13/rfrontier.pdf (accessed on Jan 30, 2019).

2.4 Conjoint Analysis

Conjoint analysis is a marketing research technique to determine consumer preferences and potential customers. Knowing customers' preference provides invaluable information about how customers think and make their decisions before purchasing products. Thus, it helps firms devise their marketing strategies including advertising, promotion, and sales activities (Ofek and Toubia 2014).

Conjoint analysis can be defined as a technique to analyze and determine how consumers value various attributes of a product and the trade-off they are willing to make among the features that comprise a product. These Attributes are inherent characteristics/features of a product. For example, price, color, RAM, camera, screen size, and battery power can be the attributes of a smartphone. Different values that a product attribute can take represent the various levels of that attribute. Black, gold, and white may be levels of the attribute color. Similarly, attribute RAM may have the levels 4GB, 8GB, and 16GB. The importance of each attribute can be inferred by experimentally manipulating the features of a product in various product bundles and observing the consumer's ratings for that product or choices among competing products in the market. The insights provided by conjoint analysis help in developing clearly defined products by identifying those features that are more appealing to the consumers. Conjoint analysis can be productively employed in the following areas:

- Designing products that maximize the measured utilities for customers in a target segment
- Modifying existing products and developing new products
- Selecting market segments for which a given product delivers high utility
- Planning competitive strategy
- Analyzing pricing policies

This technique is extremely useful in new product development process. When firms clearly define products at early stages they are 3.3 times more likely to be successful. It is often used for several types of products such as consumer goods, electrical and electronic products, retirement housing and air travel.

Types of Conjoint Analysis

The basic principle behind any type of conjoint analysis is the same—products are broken down into product attributes, and customers are faced with trade-offs when deciding which combination of attributes and their corresponding levels to purchase. There are different types of conjoint analyses depending on the response type, questioning approach, type of design and whether all the attributes are included in every question of the survey questionnaire. The four types of conjoint analysis based on response type are—(1) rating-based conjoint, (2) best–worst conjoint, (3) rank-based conjoint, and (4) choice-based conjoint. Standard conjoint and adaptive choice-based conjoint are based on the questioning approach. Generic conjoint and brand-specific conjoint are based on the type of design. Full-profile conjoint and partial-profile conjoint are classified depending on the inclusion of all attributes in every question of the survey.

Choice-based conjoint analysis is the most common type of conjoint analyses. In choice-based conjoint, respondents are presented with several questions with each question comprising 2–5 products and are asked to choose their preferred option. The results are then used to calculate a numerical value (known as a “utility score,” “utility,” or “partworth”) that measures how much each attribute and level influenced the customer’s decision to make that choice.

Conducting a Conjoint Analysis

A conjoint analysis comprises three stages:

Experimental Design: This stage is the most crucial part of the study as it ensures that the study includes all attributes and the values of attributes (levels) that will be tested. Three steps in the design are:

- Selection of attributes relevant to the product or service—Attributes are selected on the basis of required information. Only attributes that can be controlled by the firm should be included. Selecting too many attributes should be avoided as it leads to longer questionnaires, which are time consuming and difficult to answer. Ambiguous attributes must be avoided. For example, selecting vibe as an attribute for smartphone may lead to misinterpretation on the part of respondents.
- Selection of levels for the attributes—Levels for the attributes must be understandable and clear to the respondents. Ambiguity must be avoided while specifying levels. For example, specifying 4.5″–5.5″ as a smartphone screen size level should be avoided because it leaves room for interpretation. For quantitative variables the distance between two levels should not be so large that the evaluation of different levels becomes too easy.
- Development of product bundles to be evaluated—In this step, the number of product bundles that the respondent would evaluate are decided based on fractional factorial designs which are used to reduce the number of products.

Data Collection: The type of conjoint analysis used dictates the nature of data to be collected. The goal is to obtain respondents’ preferences for a carefully selected set of product bundles (Profiles). Generally, this data is collected by using a paper based or online survey. The data size depends upon the number of attribute levels to be tested. The more the number of levels of attributes more is the data to be collected for better accuracy. The first step in this process is to design a data collection procedure using one of the following methodologies:

- Pairwise evaluations of product bundles: The respondent considers two or more products at the same time and allocates 100 points among the options.
- Rank-order product bundles: The respondent sorts the products presented and assigns rank 1 to the most preferred option and a rank equal to the number of products presented to the least preferred option.

- Product evaluation on a rating scale: The respondent evaluates each product on a scale (e.g., 1–100), giving a rating of 100 to the most preferred product.
- Choice method: The respondent considers multiple products that are defined on all attributes in the study and chooses the best in the best scaling method and chooses the best and worst product in best–worst scaling method.

Decision Exploration: The last stage in conducting conjoint analysis is decision exploration, that is, evaluating product design options by segmenting customers based on their partworth functions, testing the likely success of a new product by simulating the market conditions and transforming the partworths into product choices that consumers are most likely to purchase.

Interpreting Conjoint Results

The conjoint output gives estimated utilities or partworths corresponding to average consumer preferences for the level of any given attribute.

Let us illustrate the partworths estimation to determine a consumer's preference for pizza. We define four attributes—cheese type, toppings, drink, and price with 3, 3, 2, and 3 levels, respectively, for each attribute.

Cheese	Cheddar
	Mozzarella
	Parmesan
Toppings	Onion
	Mushroom
	Pepperoni
Drink	Coke
	Pepsi
Price	$5
	$7
	$10

Consumer's preferences are quantified using linear regression for various attributes of pizza. In the survey, the respondent indicates her preferences for various pizzas defined by different combinations of the given attributes and their levels.

Sample responses of a single consumer for multiple pizza profiles

Cheddar	Mozzarella	Parmesan	Onion	Mushrooms	Pepperoni	Coke	Pepsi	Five$	Seven$	Ten$	Rating
0	0	1	1	0	0	0	1	0	0	1	4
0	1	0	0	0	1	0	1	0	0	1	1
0	0	1	0	1	0	0	1	1	0	0	6
1	0	0	0	1	0	0	1	0	1	0	1
0	1	0	1	0	0	1	0	0	1	0	3
0	0	1	0	0	1	1	0	0	1	0	6
1	0	0	0	0	1	1	0	0	0	1	6
0	0	1	0	1	0	1	0	1	0	0	7
1	0	0	1	0	0	1	0	1	0	0	2
0	1	0	0	0	1	0	1	1	0	0	1
0	1	0	0	1	0	1	0	0	0	1	3
1	0	0	1	0	0	0	1	0	1	0	2

Now, the approach is to regress the consumer's preferences for pizza based on her ratings of the various attribute levels.

We run the regression on categorical dummy variables. We set one attribute level as the baseline and remove the corresponding level from regression. Our baseline in this example is a pizza with Cheddar cheese, onion toppings, coke, and a price of 5$.

The utility of this baseline pizza is captured in the intercept from the regression output. The other coefficients give the partworths for various attribute levels. Depending on the type of information needed further insights can be derived from the partworths calculated above.

Conjoint Analysis Applications

There are many possible applications of conjoint analysis. However, the four common applications are trade-off analysis, market share forecasting, determining relative attribute importance, and comparing product alternatives. All other applications are generally variants of these four applications.

Trade-off analysis: Utilities from conjoint analysis are used to analyze whether average consumers would be willing to give up on one particular attribute to gain improvements in another.

For the given example, the partworths are estimated as follows:

Partworths			
Intercept		3.28	Best Bundle
Cheese	Cheddar	0	Parmesan
	Mozzarella	−1.11	
	Parmesan	2.78	
Toppings	Onion	0	Pepperoni
	Mushroom	0.78	
	Pepperoni	0.89	
Drink	Coke	0	Coke
	Pepsi	−2	
Price	Dollar5	0	10$
	Dollar7	−0.11	
	Dollar10	0.44	

When analyzing a group of consumers, the partworths may be interpreted as the average preferences in the group. For example, the utilities can be used to justify changes in one particular attribute at the expense of another.

Market share forecasting: This relies on the use of a multinomial logit model. To use conjoint output for market share prediction, the following conditions must be satisfied: The firm must know that we do not need other competing products in the market besides its own offering that a customer is likely to consider before selecting a product in the category. Each of these competitive products' important features must be included in the experimental design. The utilities of the competing product should also be evaluated. Based on product utilities, the market share for product i can be calculated as:

$$Share_i = \frac{e^{Ui}}{\sum_{j=1}^{n} e^{Uj}}$$

U_i is the estimated utility of product i,
U_j is the estimated utility of product j,
n is the total number of products in the competitive set, including product i.

Determining attribute importance: The range of estimated partworths within a given attribute tells how significant the attribute is in the decision process of a consumer. To calculate the importance of any attribute, the difference between the highest and lowest utility level of that attribute is divided by the sum of the differences between the highest and lowest utility level of all the attributes including the one for which the importance is being calculated. The resulting number is generally interpreted as the % decision weight of an attribute in the overall choice process. Attribute importance is measured using the following metric:

$$I_i = \frac{U_i - \overline{U}_i}{\sum_{i=1}^{n} \left(U_i - \overline{U}_i\right)}$$

I_i = importance of any given attribute i
U_i = highest utility within a given attribute
$\overline{U}_i$ = lowest utility level within a given attribute
n = total number of attributes

In the example discussed above, attribute importance for cheese can be calculated as follows:

$$\text{Range} = \text{Max (levels)} - \text{Min (levels)} = 2.78 - (-1.11) = 3.89$$

Attribute importances		
Attribute	Range	Importance
Cheese	3.89	0.53
Toppings	0.89	0.12
Drink	2	0.27
Price	0.55	0.08

$$\begin{aligned}\text{Importance of Cheese } &= \text{Range/ (Sum of ranges of all the attributes)}\\ &= 3.89/\,(3.89 + 0.89 + 2 + .55) = 0.53\end{aligned}$$

Similarly, the relative importance of other attributes can be calculated as shown in the table above.

Comparing product alternatives: Conjoint analysis can also be used to determine consumer choice between two alternative products. Based on the utility derived from each product offering, we can predict how a consumer would choose between the hypothetical profiles.

Let us consider two product profile ratings of a single respondent and predict her choice among these two profiles.

Product	Cheddar	Mozzarella	Parmesan	Onion	Mushrooms	Pepperoni	Coke	Pepsi	Five$	Seven$	Ten$	Rating
A	0	0	1	1	0	0	0	1	0	0	1	4
B	0	1	0	0	0	1	0	1	0	0	1	1

Utility of a product profile is calculated using the estimated partworths. The profile elements are each multiplied by the corresponding partworths and summed up to get the utility. The calculated utilities are then compared and the one with the higher value would be the consumer's choice.

Coefficients of the attribute levels:

Cheddar	Mozzarella	Parmesan	Onion	Mushrooms	Pepperoni	Coke	Pepsi	Five$	Seven$	Ten$
0	−1.11	2.78	0	0.78	0.89	0	−2	0	−0.11	0.44

$$\begin{aligned}\text{Utility of Product A} = {} & 3.28 + 0.00^*0 + (-1.11)^*0 + (2.78) \\ & ^*1 + (0)^*1 + (0.78)^*0 + (0.89)^*0 + (0)^*0 + (-2.00) \\ & ^*1 + 0^*0 + (-0.11)^*0 + 0.44^*1 = 4.5\end{aligned}$$

$$\begin{aligned}\text{Utility of Product B} = {} & 3.28 + 0.00^*0 + (-1.11)^*1 + (2.78)^*0 + (0) \\ & ^*0 + (0.78)^*0 + (0.89)^*1 + (0)^*0 + (-2) \\ & ^*1 + 0^*0 + (-0.11)^*0 + (0.44)^*1 = 1.5\end{aligned}$$

$$\text{Utility of Product A} > \text{Utility of Product B}$$

Therefore, the predicted choice of this particular consumer is Product A.

Product	Cheddar	Mozzarella	Parmesan	Onion	Mushrooms	Pepperoni	Coke	Pepsi	Five$	Seven$	Ten$	Utility	Choice
A	0	0	1	1	0	0	0	1	0	0	1	4.5	A
B	0	1	0	0	0	1	0	1	0	0	1	1.5	

Conjoint analysis has several advantages, such as uncovering hidden drivers that may not be apparent to the respondents themselves and evaluating the choice at an individual level. It can be used to obtain brand equity by determining the popularity of a brand. However, it has certain disadvantages such as added complexity in the experimental design because of the inclusion of a large number of attributes. This increases respondent's fatigue in taking the survey and thus compromises the accuracy of the result. Also, the validity of conjoint analysis depends on the completeness of attributes. Conjoint analysis equates a customer's overall utility for a product with the sum of her utilities for the component parts. Hence, a highly valued option on one attribute can compensate for unattractive options on another attribute and thus give misleading results.

2.5 Customer Analytics

Customer Lifetime Value (CLV)

Customers are the basis of a firm's existence. A firm creates and provides value (product or service) to customers and in return customers provide value (revenue/profits) to firms. This section will focus on one such measure for analyzing the value provided by customers to firms—customer lifetime value (CLV).

Before we understand CLV, let us look at some other customer metrics that were (are still being) used extensively before paving the way for the much superior one of CLV. One way to determine customer value is to look at metrics such as average customer revenue or average customer profit. Unless used for a specific segment, these metrics put all customers at equal level, which is incorrect and certainly not of interest to marketers who celebrate customer heterogeneity. Some customers do provide more profits than others and some customers actually lead to losses. Another interesting way to analyze the value provided by a customer is to look at their purchase behavior through RFM analysis (recency, frequency, monetary value). In RFM analysis, one looks at the following parameters:

- Recency—How recently was the last purchase made?
- Frequency—How frequently do they purchase?
- Monetary value—How much do they spend?

Customers can be rated on each parameter and classified from most valuable (highest recency, frequency, and monetary value) to least valuable (lowest recency, frequency, and monetary value). This is a simple model that can help segment customers and predict their future behavior. But all these are backward looking metrics—that is, these metrics do not take into account future potential of a customer. These methods were good enough when we did not have access to great data and advanced statistical software to analyze data. Now we have the capability to collect and analyze data at the lowest level, which allows more sophisticated analysis. CLV is one such metric that directly accounts for future value. CLV can be defined as the total financial contribution from the current period into the future—that is, revenues minus costs—of a customer over his/her future lifetime with the company and therefore reflects the future profitability of the customer (Kumar 2010). It is a forward looking customer metric that not only takes into account the current value but also the future value provided by customers. It provides a dollar value for customer relationship. It helps distinguish customers according to the value provided by them over the life of their business with the firm. Future marketing strategies can then be planned accordingly for both current and future customers.

Now that we have established that CLV is an important metric we will try to calculate it. CLV is calculated by discounting all the current and future profits expected from the customer's side. It deals with profit margins from each customer instead of revenue. It also takes into account the percentage of customers retained by the firm. CLV can be calculated as:

$$\text{CLV} = \sum_{t} \frac{m_t r_t}{(1+i)^t} \tag{19.13}$$

where m_t is the profit margin during year/duration t

r_t is the retention rate during year/duration t

i is the constant discount rate

t is year/duration

When the onetime acquisition cost of a customer is subtracted from Eq. (19.13) we get the CLV of that particular customer. If we assume an infinite horizon, a constant profit margin, a constant retention rate, and a constant discount rate Eq. (19.13) can be simplified to:

$$CLV = m\left(\frac{r}{1+i-r}\right) \tag{19.14}$$

Here $\left(\frac{r}{1+i-r}\right)$ is the margin multiple and depends on the retention rate and discount rate. The higher the retention rate and lower the discount rate, the more valuable is a customer. Table 19.2 provides the margin multiple based on typical retention and discount rates. Table 19.2 is a very useful "back of the envelope" calculation of CLV!

CLV is the maximum value provided by customers to the firm. Hence, it can be used as the upper limit of any customer-centric activity. For example, if a potential customer's CLV is $5 then a firm should not spend more than $5 on acquiring this customer. This holds for customer retention and development activities as well. As mentioned before, the dollar value provided by CLV can help distinguish cohorts of customers. So it also helps distinguish future prospects that are similar to currently profitable customers.

As with any metric, CLV also has some limitations. First, it is very difficult to calculate profit margins at individual level. Sometimes, revenue and/or cost cannot be attributed to a single customer, making it difficult to calculate profit margins. Similarly, it is difficult to accurately calculate retention rates because the rate calculation requires sophisticated analysis. A small increase in retention rate can

Table 19.2 Margin multiple

	Discount rate (i)			
Retention rate (r)	10%	12%	14%	16%
60%	1.20	1.15	1.11	1.07
70%	1.75	1.67	1.59	1.52
80%	2.67	2.50	2.35	2.22
90%	4.50	4.09	3.75	3.46

substantially increase the margin multiple, which in turn increases the CLV. Hence, accuracy of retention rate is extremely important for calculating CLV.

Despite the limitations mentioned above, CLV helps in making important marketing decisions. It is a customer metric that takes into account both the current and future profitability of a customer. But CLV should not be the only decision-making criteria. Firms should take into account factors such as reference/influence of the customer, and brand reputation, along with the CLV to make marketing decisions.

Customer Referral Value (CRV)

There are many resource-intensive ways to go about acquiring new businesses. While most firms are engaged in the traditional ways of attracting customers and hence gaining profits, there are effective ways to bring the business at no or little cost. One such way is customer referral. Think about your friends who are loyal customers to Tata Motors. Tata Motors can use your friends to refer you for a new model of car, or simply to the firm itself. In such situations, will Tata Motors incur cost to acquire? Probably not . . . or may be a little. Conditional on the tangible and intangible aspects of reference, you may end up buying a new car from Tata Motors. And that is where the power of metrics such as customer referral value (CRV) lies.

Let us think about a few firms that engage in the referral programs: *Dropbox referral program* secured 4 million users in just 15 months. Inspired by PayPal, who literally gave free money for referrals, *Dropbox* added double-sided referral programs, where both referrer and referee get rewarded. Amazon prime, PayPal, Airbnb, and Uber have recently seen huge success of referral programs in their business strategies.[6] Whether it is a B2B or a B2C business, customer referral has seen significant success in recent times. According to Linkedln, 84% of B2B buying decisions start with a referral. Further, customer referral has a good conversion rate. The probability of a referred customer getting converted is 30% higher as compared to a lead generated through traditional channels.[7]

Now the question is how to capitalize on the referrals and design business strategies. Literature on customer management suggests that firms should engage in measuring the values of each referral and then decide the follow-up strategies (e.g., Kumar 2010; Kumar et al. 2007). Accordingly, customer referral value is defined as an estimate of lifetime values of any type-one referrals—people who would have not purchased or become customers without referrals (e.g., Kumar et al. 2007). Firms should also include the value of type-two referrals—people who would have become customers anyway. This has implications for managing marketing efforts to acquire new customers.

CRV is more complicated than computing customer lifetime value (CLV). Computation of CRV requires the estimation of the average number of successful referrals a customer makes after providing some incentive from the firm's side. For

[6]https://www.referralcandy.com/blog/47-referral-programs/ (accessed on May 19, 2018).

[7]https://influitive.com/blog/9-stellar-referral-program-examples/ (accessed on May 19, 2018).

that, we need to look at the past behavior, which must include enough variance in the number of referrals for proper empirical modeling and accuracy. Computation of the CRV requires understanding of the time that can go by and still be sure that a customer's referrals are actually prompted by a firm's referral incentives. Further, it is critical to understand the conversion rate of referrals to actual customers. Finally, a customer's referral value is the present value of her type-one referrals plus present value of her type-two referrals (Kumar et al. 2007). Customer referral value of a customer is the monetary value associated with the future profits given by each referred prospect, discounted to present value.

CRV can be calculated by summing up the value of the customers who joined because of the referral and the value of the customers that would have joined anyway discounted to present value. We can compute the CRV of customer[8] i as

$$\mathrm{CRV_i} = \sum_{t=1}^{T}\sum_{y=1}^{n1}\frac{\left(A_{ty} - a_{ty} + M_{ty} + ACQ1_{ty}\right)}{(1+r)^t} + \sum_{t=1}^{T}\sum_{y=1}^{n2}\frac{\left(ACQ_{ty}\right)}{(1+r)^t}$$

where

A_{ty} = contribution margin by customer y who otherwise would not buy the product

a_{ty} = cost of the referral for customer y

$ACQ1_{ty}$ = savings in acquisition cost from customers who would not join w/o the referral

$ACQ2_{ty}$ = savings in acquisition cost from customers who would have joined anyway

T = number of periods that will be predicted into the future (e.g., years)

n1 = number of customers who would not join w/o the referral

n2 = number of customers who would have joined anyway

M_{ty} = marketing costs required to retain customers

A firm can first compute the CRV of their customers and then categorize them based on the value of CRV. A firm may find a particular group of customers to have significantly higher CRV than that of others. The firm can market to or provide incentives to this set of customers to increase the referrals and hence new customer acquisitions. Moreover, a firm can look at profiles, similar to that of the high CRV group, who have not referred yet and induce them to refer by providing some incentives. Given the available data, CRV can be computed with any standard statistical software.

Customer Influence Value (CIV)

Imagine yourself to be looking at the best car that is affordable within your budget. Also, imagine that you do not have much knowledge on the technical aspects of a car. Probably you will ask your friends and colleagues or search online for advice.

[8] Formula for computing CRV is adopted from Kumar et al. (2007).

What else can you do to reach to your decision? With growing power of social platforms (online or offline), you may want to post a question on Facebook or ask a question to an expert auto-blogger, or follow someone on Instagram whose ideas, comments, and feedback about the auto industry influence you. Just about everything from big firms to kids have some sort of strategies, tips, experiences, and attribution that drive others' decision, sales, impression, etc. That said, most firms cannot fully grasp the value of societal connections strategies and tactics, which goes beyond traditional marketing ploys and tactics.[9] It is critical to compute this influence while determining the value of your customers. Social influence can play a significantly larger role in your decision to buy a new car or putting your kids in a particular school, or deciding on which dating app to go for. Hence, value of a customer can go beyond her purchase value, or referral values. In addition to CLV and CRV, value of a customer can stem from her influence on other customers. Value of a customer's influence refers to the monetary value of the profits associated with the purchases generated by a customer's social media influence on other acquired customers and prospects, discounted to present value (Kumar 2013).

Understanding CIV can be of great value for most firms. For an ice cream retailer, Kumar and Mirchandani (2012) show that a firm can harness the true value of customer influence. They design a seven-step process to identify the influencers in online social network, observe their influences over time, and substantially improve the firm performance. Indeed customers' influence has significant value for firms and firms should measure and implement CIV in their business strategies. For a detailed understanding of CIV computation, refer Kumar et al. (2013), Kumar and Mirchandani (2012).

3 Applications

Marketing Analytics has evolved over the last century of applications, research and data collection. Some might even say that marketing was the first consumer of large "business" data! Wedel and Kannan (2016) provide a readable summary of the evolution of marketing analytics as well as pose several questions for researchers. They classify the applications into customer relationship management (CRM), marketing mix analytics, personalization, and privacy & data security. In this book, too, there are many applications of the tools: the chapter on social media analytics has applications to online advertising, A/B experiments, and digital attribution; the chapters on forecasting analytics, retail analytics, pricing analytics, and supply chain analytics contain applications to their specialized settings, such as demand forecasting, assortment planning, and distribution planning; and the case study "InfoMedia Solutions" contains an application to media-mix planning. Other

[9]CLV: http://www.customerlifetimevalue.co/ and CIV: https://www.mavrck.co/resources/ (accessed on Sep 15, 2018).

applications covered in this book are recommendation engines, geo-fencing, market segmentation, and search targeting. The rapid changes in data availability and tools will continue to spur the development of new applications in marketing. To keep abreast of the new developments, researchers may follow topics in "Marketing Science Institute[10]" and their research priorities.

Electronic Supplementary Material

All the datasets, code, and other material referred in this section are available in www.allaboutanalytics.net.

- Data 19.1: exercise_inter.csv
- Data 19.2: exercise_curvilinear.csv
- Data 19.3: exercise_mediation.csv
- Data 19.4: ABC_hospital_group.csv
- Data 19.5: restaurant_chain_data.csv ("DEA in practice" section)
- Data 19.6: pizza.csv ("Conjoint Analysis Interpretation" section)
- Data 19.7: product_profile_ratings.csv ("Comparing product alternatives" section in conjoint analysis)

Exercises

Ex. 19.1 Use the data file titled "*exercise_inter.csv*" and answer the following questions:

a) Why are independent variables mean-centered? Mean-center advertising, discount, and promotion variables.
b) Is the effect of advertising on sales contingent on the level of discount? Plot a graph to interpret the interaction effect.
c) Is the effect of advertising on sales contingent on the level of promotion? Plot a graph to interpret the interaction effect.

Ex. 19.2 Use the data file titled "*exercise_curvilinear.csv*" and answer the following questions:

a) Is the effect of advertising on profit curvilinear or linear? Plot a graph if the relationship is curvilinear.
b) Is the effect of sales promotion on profit curvilinear or linear? Plot a graph if the relationship is curvilinear.
c) Is the effect of rebates and discount on profit curvilinear or linear? Plot a graph if the relationship is curvilinear.

[10]www.msi.org (accessed on Jul 6, 2018).

Ex. 19.3 Use "*exercise_mediation.csv*" data and answer the following questions:

a) Does recall mediate the effect of advertising on market share?
b) Why is bootstrapping used in mediation analysis?

Ex. 19.4 Find the most efficient hospital unit of ABC hospital group given the following information:

Unit	1	2	3	4	5	6	7	8
Profit (crore)	120	160	430	856	200	320	189	253
Number of Specialists	150	100	120	180	220	90	140	160
Area (sq feet)	21,000	32,650	40,000	18,780	19,870	50,000	33,000	19,878

References

Aigner, D. J., Lovell, C. A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production functions. *Journal of Econometrics, 6*(1), 21–37.

Aiken, L.S., and West, S.G, 1991. *Multiple regression: Testing and interpreting interactions*.

Baccouche, R., & Kouki, M. (2003). Stochastic production frontier and technical inefficiency: A sensitivity analysis. *Econometric Reviews, 22*(1), 79–91.

Banker, R. D., Cooper, W. W., Seiford, L. M., Thrall, R. M., & Zhu, J. (2004). Returns to scale in different DEA models. *European Journal of Operational Research, 154*, 345–362.

Banker, R. D., & Morey, R. (1986). Efficiency analysis for exogenously fixed inputs and outputs. *Operation Research, 34*, 513–521.

Banker, R. D., & Thrall, R. M. (1992). Estimation of returns to scale using data envelopment analysis. *European Journal of Operational Research, 62*(1), 74–84.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173.

Bera, A. K., & Sharma, S. C. (1999). Estimating production uncertainty in stochastic frontier production function models. *Journal of Productivity Analysis, 12*(2), 187–210.

Boussofiane, A., Dyson, R. G., & Thanassoulis, E. (1991). Applied data envelopment analysis. *European Journal of Operational Research, 52*(1), 1–15.

Caudill, S. B., Ford, J. M., & Gropper, D. M. (1995). Frontier estimation and firm-specific 1076 inefficiency measure in the presence of heteroskedasticity. *Journal of Business & Economic Statistics, 13*(1), 105–111.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*, 429–444.

Charnes, A., Clark, T., Cooper, W. W., & Golany, B. (1985). A developmental study of data envelopment analysis in measuring the efficiency of maintenance units in the U.S. air forces, in: R. Thompson and R.M. Thrall (eds.). *Annals of Operational Research, 2*, 95–112.

Chen, C.-F., & Soo, K. T. (2010). Some university students are more equal than others: efficiency evidence from England. *Economics Bulletin, 30*(4), 2697–2708.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook, W. D., & Seiford, L. M. (2009). Data envelopment analysis (DEA)–Thirty years on. *European Journal of Operational Research, 192*(1), 1–17.

Cooper, W. W., Seiford, L. M., & Zhu, J. (2004). *Data envelopment analysis. Handbook on data envelopment analysis* (pp. 1–39). Boston, MA: Springer.

Cullinane, K., Wang, T. F., Song, D. W., & Ji, P. (2006). The technical efficiency of container ports: comparing data envelopment analysis and stochastic frontier analysis. *Transportation Research Part A: Policy and Practice, 40*(4), 354–374.

Farrell, M. J., & Fieldhouse, M. (1962). Estimating efficient production functions under increasing returns to scale. *Journal of the Royal Statistical Society. Series A (General), 125*, 252–267.

Farris, P. W., Bendle, N. T., Pfeifer, P. E., & Reibstein, D. J. (2010). *Marketing metrics: The definitive guide to measuring marketing performance, Introduction* (pp. 1–25). London: Pearson.

Feng, C., & Fay, S. A. (2016). Inferring salesperson capability using stochastic frontier analysis. *Journal of Personal Selling and Sales Management, 36*, 294–306.

Fenn, P., Vencappa, D., Diacon, S., Klumpes, P., & O'Brien, C. (2008). Market structure and the efficiency of European insurance companies: A stochastic frontier analysis. *Journal of Banking & Finance, 32*(1), 86–100.

Førsund, F. R., Lovell, C. K., & Schmidt, P. (1980). A survey of frontier production functions and of their relationship to efficiency measurement. *Journal of Econometrics, 13*(1), 5–25.

Gagnepain, P., & Ivaldi, M. (2002). Stochastic frontiers and asymmetric information models. *Journal of Productivity Analysis, 18*(2), 145–159.

Greene, W. H. (2010). A stochastic frontier model with correction for sample selection. *Journal of Productivity Analysis, 34*(1), 15–24.

Hjalmarsson, L., Kumbhakar, S. C., & Heshmati, A. (1996). DEA, DFA and SFA: a comparison. *Journal of Productivity Analysis, 7*(2-3), 303–327.

Jacobs, R. (2001). Alternative methods to examine hospital efficiency: Data envelopment analysis and stochastic frontier analysis. *Health Care Management Science, 4*(2), 103–115.

Koetter, M., & Poghosyan, T. (2009). The identification of technology regimes in banking: Implications for the market power-fragility nexus. *Journal of Banking & Finance, 33*, 1413–1422.

Kumar, V. (2010). *Customer relationship management*. Hoboken, NJ: Wiley Online Library.

Kumar, V. (2013). *Profitable customer engagement: Concept, metrics and strategies*. Thousand Oaks, CA: SAGE Publications India.

Kumar, V., Andrew Petersen, J., & Leone, R. P. (2007). How valuable is word of mouth? *Harvard Business Review, 85*(10), 139.

Kumar, V., Bhaskaran, V., Mirchandani, R., & Shah, M. (2013). Practice prize winner-creating a measurable social media marketing strategy: Increasing the value and ROI of intangibles and tangibles for Hokey Pokey. *Marketing Science, 32*(2), 194–212.

Kumar, V., & Mirchandani, R. (2012). Increasing the ROI of social media marketing. *MIT Sloan Management Review, 54*(1), 55.

Kumar, V., & Sharma, A. (2017). Leveraging marketing analytics to improve firm performance: insights from implementation. *Applied Marketing Analytics, 3*(1), 58–69.

Kumar, V., Sharma, A., Donthu, N., & Rountree, C. (2015). Practice prize paper-implementing integrated marketing science modeling at a non-profit organization: Balancing multiple business objectives at Georgia Aquarium. *Marketing Science, 34*(6), 804–814.

Kumbhakar, S. C., & Lovell, C. K. (2003). *Stochastic frontier analysis*. Cambridge: Cambridge university press.

Lilien, G. L., Rangaswamy, A., & De Bruyn, A. (2013). *Principles of marketing engineering*. State College, PA: DecisionPro.

Lilien, G. L. (2011). Bridging the academic–practitioner divide in marketing decision models. *Journal of Marketing, 75*(4), 196–210.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Abingdon: Routledge.

Meeusen, W., & van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review, 18*(2), 435–444.

Ofek, E., & Toubia, O. (2014). *Conjoint analysis: A do it yourself guide*. Harvard Business School, note, 515024.

Ouellette, P., & Vierstraete, V. (2004). Technological change and efficiency in the presence of quasi-fixed inputs: A DEA application to the hospital sector. *European Journal of Operational Research, 154*(3), 755–763.

Parsons, J. L. (2002). Using stochastic frontier analysis for performance measurement and benchmarking. *Advances in Econometrics, 16*, 317–350.

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31*(4), 437–448.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, 36*(4), 717–731.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42*(1), 185–227.

Reinhard, S., Lovell, C., & Thijssen, G. (2000). Environmental efficiency with multiple environmentally detrimental variables; estimated with SFA and DEA. *European Journal of Operational Research, 121*(3), 287–303.

Schmidt, P. (1985). Frontier production functions. *Econometric Reviews, 4*(2), 289–328.

Seiford, L. M., & Zhu, J. (1999). An investigation of returns to scale under data envelopment analysis. *Omega, 27*, 1–11.

Sherman, H. D., & Gold, F. (1985). Bank branch operating efficiency: Evaluation with data envelopment analysis. *Journal of Banking & Finance, 9*(2), 297–315.

Venkatesan, R., Farris, P., & Wilcox, R. T. (2014). *Cutting-edge marketing analytics: Real world cases and data sets for hands on learning*. London: Pearson Education.

Vyt, D. (2008). Retail network performance evaluation: A DEA approach considering retailers' geomarketing. *The International Review of Retail, Distribution and Consumer Research*, 235–253.

Wadud, A., & White, B. (2000). Farm household efficiency in Bangladesh: a comparison of stochastic frontier and DEA methods. *Applied Economics, 32*(13), 1665–1673.

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing, 80*, 97–121.

Other Resources

Kristopher, J. Preacher (2018). *Preacher's website*. Retrieved May 19, 2018, from http://quantpsy.org/medn.htm.

Retrieved May 19, 2018, from www.conjoint.online.

Software: STATA, Mplus. Retrieved May 19, 2018, from www.statmodel.com.