DEMAND FORECASTING IN MARKETING: METHODS, TYPES OF DATA, AND FUTURE RESEARCH

Carla Freitas Silveira Netto – Universidade Federal do Rio Grande do Sul - Brazil – carla.netto@gmail.com

Vinicius Andade Brei – Universidade Federal do Rio Grande do Sul – Brazil – brei@ufrgs.br

The first author has a scholarship from CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico

Abstract

Demand forecasts are fundamental to plan and deliver products and services. Despite such relevance, marketers have difficulty to choose which forecast method is the best for their organizations. One possible explanation for this baffling task is that the literature is not clear about demand forecasting methods' classifications, approaches, complexity, requirements, and efficiency. This theoretical paper tries to improve this scenario, reviewing the state of the art about demand forecasting in marketing. More specifically, we focus on: (1) the most frequently used models by academics and practitioners; (2) different classifications and approaches of those models, especially the ones based on statistics/mathematics and big-data; (3) challenges of big data/computer based forecasting; (4) types of data used; and (5) research gaps and suggestions of future research on demand forecasting in marketing. The most important research gaps are related to the types of models applied in marketing literature (structural). Besides simpler, easier to implement models, further research is necessary to develop forecasting techniques that incorporate dynamic effects, primary data, and nonparametric approaches more efficiently. The literature also evidences some gaps concerning the optimal use of types of data and data sources. Of foremost importance are data sets about durable goods, location/geographical data, big data, and the combination of different data sets. Based on the state of the art about forecasting methods, types and use of data, and research gaps found, we present suggestions for future research. New studies about demand forecasting in marketing should focus on durables goods and other types of less frequently purchased products. They could also combine different sources of data, such as free public data, firm property data, commercially available market research, big data, and primary data (e.g., surveys and experiments). Future studies should also analyze how to improve the use of location/geographical data, incorporating their dynamic perspective, without creating barriers to the method implementation. We also discuss how marketing and computer science should be integrated to fulfill those gaps.

Key words - Demand; Marketing; theoretical paper

1. INTRODUCTION

Demand forecasts are important to the most basics processes in any organization. To plan and deliver products and services is necessary to know what the future might hold. However, a demand forecast is important to plan all business decisions: sales, finance, production management, logistics and also marketing (Canitz, 2016). To be able to predict next purchases is a valuable thing to marketing more than for other fields in social sciences (Chintagunta & Nair, 2011). As Beal & Wilson (2015) states,

> making the best possible forecasts using data that are readily available can help businesses provide consumers with the right product at the right place at the right time and at the right price. Forecasting helps change data into information which can help businesses become more profitable [...] Thus, forecasting knowledge and ability

> should be an essential skill set of all marketing majors (Beal & Wilson, 2015, p. 115).

To select the most appropriate forecasting technique from the range available is challenging. According to Armstrong (2001) the ways of selecting forecasting methods are: convenience (inexpensive, but risky); market popularity (what others do); what experts advise; statistical criteria; track record; and guidelines from prior research.

This theoretical paper has the purpose to review the literature (the area guidelines from prior research) about demand forecasting. This review focuses on: (1) the models of demand forecasting in marketing (literature and practice); (2) different classifications and approaches of those models; (3) challenges of big data/computer based forecasting; (4) types of data used in demand forecasting models in marketing; and (5) research gaps on demand forecasting in marketing.

The choice of the method is usually based on familiarity and not on what is more appropriate to the market studied or the data (Canitz, 2016). So, to select the method (and the respective technique) it is important to consider not only the characteristics of the market studied, but also the characteristics of the available data. The first criteria to select a method is related to the amount of objective data available (Armstrong, 2001). This will define if it is to follow a qualitative/judgmental approach or a quantitative one. There are fields of knowledge that have searched for improvements on judgmental methods. Operational research is one of them. This area has combined quantitative methods with qualitative ones (Fildes, Nikolopoulos, Crone, & Syntetos, 2008), such as: Delphi, intentions-to-buy surveys, and also the combination of individuals' forecasts (as sales staff opinions).

For this theoretical paper, we consider that judgments or domain knowledge should be used to create hypothesis and add structure to the model, but not to override the forecast after it is done. Domain knowledge improves forecast accuracy and reduces the need to do such adjustments (Chase Jr, 2013). The general principle that is followed is to select a method that is “structured, quantitative, causal, and simple” (Armstrong, 2001,p.373).

Another reason is that the articles found in marketing focus on structured, quantitative and causal models (e.g., Albuquerque & Bronnenberg, 2012; Allenby, Garratt, & Rossi, 2010; Bollinger & Gillingham, 2012; Che, Chen, & Chen, 2012; Chen, Wang, & Xie, 2011; Ching, Clark, Horstmann, & Lim, 2015; Draganska & Klapper, 2011; Jing & Lewis, 2011; Liu, Singh, & Srinivasan, 2016; Luan & Sudhir, 2010; Mehta & Ma, 2012; Mukherjee & Kadiyali, 2011; Narayanan & Nair, 2013; Petrin & Train, 2010; Shah, Kumar, & Zhao, 2015; Shriver, 2015; Stephen & Galak, 2012; Yang, Zhao, Erdem, & Zhao, 2010; Zhang & Kalra, 2014). For those reasons, we focus on quantitative methods of demand forecasting.

Regarding quantitative methods, Singh (2016) divide the research on forecasting in four types: behavioral-focused (judgmental adjustments to statistical forecasts); business performance focused (impact of forecasting practices on performance); statistics/mathematics-focused (time-series and causal); and big-data-based (the newest research stream). As mentioned before, judgmental adjustments are beyond the scope of this study. Business performance is not analyzed either since the goal is not to discuss the advantages or difficulties to implement the process of forecasting in companies. Therefore, in the remaining of this theoretical paper the focus will be on statistics/mathematics and big-data based forecasting.

This paper will describe the classification of models and types of data used in demand forecasting in marketing, since “the proliferation of data, contexts, and motivations has now resulted in large classes of demand models, differing both in their properties and in their intended use” (Chintagunta & Nair, 2011, p.977). It unfolds as follows: first the classification of models of demand forecasting in marketing literature are discussed, divided in two approaches: statistics/mathematics and big-data based researches. In forecasting practice the focus is on the statistics/mathematics techniques applied. After that, we introduce the types of data used in models of demand forecasting in marketing. Finally, the gaps found in the literature are presented and summarized.

2. STATISTICS/MATHEMATICS-FOCUSED METHODS OF DEMAND FORECASTING IN MARKETING LITERATURE

Demand systems can be divided into two: demand in characteristics space and in product space. Demand in characteristics space assumes that consumers choose products by groups of characteristics. These models are flexible and usually outperform the models of the product space system. An issue is the assumption that consumers choose no more than one good (Nevo, 2011). Within characteristics space system, discrete choice models are more popular in the academic literature (Nevo, 2011). Discrete-choice models are popular in marketing because “much of micro data in marketing involve consumers choosing from a fixed set of alternatives within a category” (Chintagunta & Nair, 2011, p.981).

Demand in product space, on the other hand, considers that consumers decide first by categories, then by segments and finally by brands. Therefore, products, not characteristics, are grouped into these models. Examples are: linear expenditure models; Rotterdam model; Translog Model; and AIDS. Product space systems are simpler to estimate, requiring mostly linear methods, which save computational time. On the other hand, the products need to be classified into segments that are frequently hard to justify. It also assumes that consumers buy a number of products of all brands, when consumers, in reality, may consume more than one brand, but not all of them (Nevo, 2011).

One common disadvantage of both systems of models is that they are static and for many markets the demand is dynamic. This means that they do not consider the possibility of consumers' decisions in the present affecting posterior decision or that the present decision is affected by expectations of the future (Nevo, 2011).

Another way of classifying is suggested by Roberts (1998). The author divides the models of demand forecasting by level (individual or aggregate) and if they are applied for new products or not. Models for new products at the individual level are used to forecast the market share and they apply discrete choice analysis (Roberts, 1998). According to Roberts (1998), the focus of forecasting new products has been on the aggregate-level, applying diffusion models and are, in reality, frequently applied only post-launch, studying what made the diffusion possible because pre-launch forecasts are challenging.

Post-launch models can be at the individual level and also apply discrete choice models. The individual level data comes from scanner data that is frequently used to analyze consumer preference and response to marketing instruments (Roberts, 1998). At the aggregate level, marketing has focused in the "study of advertising effects and other marketing mix variables on sales" (Roberts, 1998, p.172).

Types of demand analysis can also be divided based on their goals, according to Chintagunta & Nair (2011): forecasting, measurement, and testing. These authors subdivide these goals on their respective models: descriptive models (for stable environments), structural models, and reduce-form causal effects. By descriptive models, they mean models that focus on forecasting sales across time on the bases of variables available today (e.g., current marketing mix variables and sales). The emphasis of these models is not on causality (Chintagunta & Nair, 2011). These models "cannot literally test a theory about consumer or firm behavior — only an econometric representation of the theory can serve as the basis for such a test" (Reiss, 2011, p.952).

Structural models, on the other hand, use the theory to predict phenomena. These models should "demonstrate that the theory, combined with the chosen econometric specification, can explain key patterns in sample" (Chintagunta & Nair, 2011, p.979). They "combine mathematical, economic, or marketing models of behavior with statistical assumptions to derive estimable empirical models" (Reiss, 2011, p.951). Discrete choice models are examples of structural models. They can be causally interpreted but have limitations as well (Chintagunta & Nair, 2011; Reiss, 2011):

- It is challenging to find the best combination of theory, data and econometric specification;
- They are time consuming;
- They make it difficult to build simple models that are realistic and can be estimated with the data available;
- They have results that may be overly affected by strong assumptions;
- The assumptions of distributions are made for computationally convenience and most times do not have economic defense.

Finally, reduce-form causal models, used for the goals of measurement and testing, are different from descriptive models since the latter does not imply causality. They are also diverse from structure models, as "fewer distributional and specification assumptions are required because simulating radically different counterfactuals is not a goal of the analysis" (Chintagunta & Nair, 2011). The similarities are that structural and reduce-form models imply causality and require theory. The authors conclude mentioning that "whereas a true reduced form is derived from a structural model, this term is now routinely [and not correctly] used to describe descriptive (linear) regressions" (Reiss, 2011, p.962).

In the marketing literature the models are mostly structural (e.g., Albuquerque & Bronnenberg, 2012; Allenby et al., 2010; Bollinger & Gillingham, 2012; Che et al., 2012;

Chen et al., 2011; Ching et al., 2015; Draganska & Klapper, 2011; Jing & Lewis, 2011; Liu et al., 2016; Luan & Sudhir, 2010; Mehta & Ma, 2012; Mukherjee & Kadiyali, 2011; Narayanan & Nair, 2013; Petrin & Train, 2010; Shah et al., 2015; Shriver, 2015; Stephen & Galak, 2012; Yang et al., 2010; Zhang & Kalra, 2014). There are also some models that apply a reduced form (e.g., Briesch, Dillon, & Fox, 2013; Chung, Derdenger, & Srinivasan, 2013) or a combination of reduced form and structural model (Chung et al., 2013). Some studies apply structural models with a Bayesian approach (e.g., Aribarg, Arora, & Kang, 2010; Arora, Henderson, & Liu, 2011; Che et al., 2012; J. Chung & Rao, 2012; Feit, Wang, Bradlow, & Fader, 2013; Rooderkerk, Van Heerde, & Bijmolt, 2011; Zhao, Yang, Narayan, & Zhao, 2013; Zhao, Zhao, & Helsen, 2011).

For that reason another important distinction must be made between the use of classical and Bayesian statistics in demand forecasting models. This is important since "the vast majority of the recent Bayesian literature in marketing emphasizes the value of the Bayesian approach to inference, particularly in situations with limited information" (Rossi & Allenby, 2003, p.317). Bayesian statistics is commonly used in marketing, partially due to computing developments that have made it accessible (Allenby, Bakken, & Rossi, 2004). For example, Markov Chain Monte Carlo (MCMC) simulation made it easier to estimate complex models of behavior that would not be possible with other methods (Allenby et al., 2004).

According to Allenby et al. (2004), Fildes et al., (2008), and Rossi & Allenby (2003), the advantages of the Bayesian approach are: it is able to reflect heterogeneity in consumer preferences; the developed models are more realistic; it allows disaggregate analysis; the Hierarchical Bayes methods have predictive superiority due to avoiding the restrictive analytic assumptions that alternative methods impose; it allows studies of high-dimensional data and complex relationships; and "instead of a point estimate of values for each respondent, we usually end up with a distribution of estimates for each respondent" (Allenby et al., 2004, p.25). This can be informative about uncertainty, but, as the authors mention, can also complicate the analysis.

One important difference between classical and Bayesian statistics is that the former says nothing about how to incorporate different sources of data, such as expert's information and other data sets (Rossi & Allenby, 2003). That is important because "a major challenge facing marketing practitioners is the merging of information acquired across a variety of different data sets" (Rossi & Allenby, 2003, p.321).

Reiss (2011) explains an experimental approach in demand forecasting studies to overcome problems with endogeneity bias. There is a bias when omitted variables (unobserved factors) are correlated with the error and one or more independent variables in a regression (Rossi, 2014). The idea is to run experiments manipulating one variable (price, for example) independently of the others (e.g.; promotions), controlling this issue. This approach includes: lab experiments, randomized field experiments, natural experiments and instrumental variables (IV) methods.

Marketing researches prefer the more realistic field experiments that have external validity, but is more difficult to control for confounds (Reiss, 2011). Natural experiments are still rare in marketing, but are commonly applied in economics (Chen et al., 2011). These types of experiments investigate effects of variables not in control of the researches, such as government intervention or policy changes, for example (Chen et al., 2011; Reiss, 2011). Instrumental variables (IV) are observables variables that are correlated with X variable, but are not a part of the structural equation, not affecting Y (Rossi, 2014). IV and experiments do not solve this problem without support from theory (Reiss, 2011). The researcher should not use invalid or week IVs. If there is concern about an unobserved variable, it should be measured (Rossi, 2014).

After this discussion about the classifications and types of models of demand forecast used in marketing literature, the focus will be on practice. The next section will briefly discuss the differences in techniques used in the marketing literature and practice.

3. STATISTICS/MATHEMATICS-FOCUSED DEMAND FORECASTING TECHNIQUES USED IN PRACTICE

Statistics/mathematics-focused demand forecasting techniques used in practice can be divided in three streams (Chase Jr, 2013): time-series (also called extrapolation methods), causal and weighted combined forecasting. Time-series techniques are those that identify patterns (trend, seasonality, cyclical, and randomness) and predict those into the future. These techniques have higher predictive accuracy in stable markets. Examples include: moving average; simple exponential smoothing; Holt's two parameter; and Winters' three parameter. Time-series are simple to develop and requires minimal amount of data but are unable to predict sudden changes in demand. It can also predict only one to three periods ahead with any degree of accuracy (Chase Jr, 2013).

ARIMA are a more advanced time-series technique that combines time series and regression elements (Chase Jr, 2013). This technique is more accurate to predict demand in the long term. It can model trend/cycle, seasonality, as well as other factors influencing demand (explanatory variables), but requires more data and is more complex to develop (Chase Jr, 2013).

Causal techniques assume that future sales are related to changes in other variables (price, promotions, among others). Examples are regression and ARIMAX (an extension of ARIMA). These techniques require more data and are more complicated to develop (Chase Jr, 2013). On the other hand, they can include intervention variables (using dummy variables).

Weighted combined forecasting combine methods (i.e., time series, causal, and/or judgmental) and create a single forecast (Chase Jr, 2013). This can be done by given each equal or different weight. This combination outperforms most single forecasts, since biases among methods will compensate one another (Chase Jr, 2013).

Even believing that forecasting accuracy is important, practitioners still practice simpler forecasting methods (Fildes et al., 2008). One reason is that the models that result from marketing research are hard to implement in practice. This is a significant challenge of structural models, because “the large number of observations in practice, the large array of state and control variables, and the frequency of decision making can render the application of structural models infeasible” (Mela, 2011, p.974). In the next section we discuss big-data based/computer intensive methods of demand forecasting that are applied in the literature with potential to become techniques used in marketing practice as well.

4. BIG-DATA BASED / COMPUTER-INTENSIVE METHODS OF DEMAND FORECASTING IN MARKETING

Methods applied to demand forecasting are related to the characteristics of the variables and data sets that became available to marketing scientists. As scanner data enabled marketing to apply structured causal models from fields like transportation science and economics, now big data makes it possible to apply methods from machine learning (Chintagunta, Hanssens, & Hauser, 2016).

Big-data is more about new types of variables, size of data sets used but also different methods applied. With bigger data sets and a higher number of different attributes applied in forecasting studies, “analysis using conventional statistical methods [are] impractical or even infeasible because of computer constraints” (Fildes et al., 2008, p.1165).

Computer intensive methods, “applies statistical and machine learning algorithms for discovering valid, novel and potentially useful predictive information from large data sets in unstructured problems” (Fildes et al., 2008, p. 1165). They are created in different areas such as statistics and machine learning. Statistics is the intellectual base of these applications, “since the notion of finding useful patterns from data for prediction has long been a statistical endeavor” (Fildes et al., 2008, p.1156).

Most studies applying data mining tools use current customers’ data. The focus is on segmenting or classifying them, in CRM related topics, such as up/cross-selling, churn analysis, credit scoring, among others (Fildes et al., 2008; Ngai, Xiu, & Chau, 2009). One example of this application is the study of Sundsøy, Bjelland, Iqbal, de Montjoye, & others (2014) that compared different machine learning algorithms to the judgment of the marketing team in identifying customers that were more likely to buy mobile internet service.

On table 1 a few of the most influential algorithms are presented with their definitions, advantages and some limitations. The algorithms described on table 1 are not representative of all existing algorithms only of those found in the literature and used for demand forecasting.

Table 1 – Some of the most influential data mining algorithms

	Definition	Advantages	Limitations
SVM	The aim is to find the best classification function to distinguish between members of different classes	One of the most robust and accurate; has strong theorical background; requires less examples for training; have performed well where complex relationships between the input attributes and the output attribute exist; used for nonlinear environments; can be used as	Difficult to directly interpret the findings; does not yield probability estimates for hypothesis testing or predictive/posterior bounds; do not have an easy software

		structural gap identifier	application
CART (Classifi-cation and Regression Trees)	Decision tree method – a binary recursive partitioning procedure	Capable of processing continuous and nominal attributes both as targets and predictors; the prediction for a particular case can be traced back	Are always binary; trees with many layers are not easy to interpret

Font: Adapted from Ali, Sayın, Van Woensel, & Fransoo (2009); Cui & Curry (2005); Wu et al. (2008).

Two examples of sales forecasting studies that used the data mining techniques of table 1 are Ali et al. (2009) and Sun, Choi, Au, & Yu (2008). Although these studies are related to the marketing area they were not published on marketing journals. Ali et al. (2009) forecast sales of a grocery store, in the presence of promotions. The authors compared different methods of forecasting: exponential smoothing; stepwise linear regression; support vector regression (SRM's regression version) and CART (regression trees).

As a result, they found that, for periods with no information about marketing instruments (in this case, the promotions), time series techniques performed better. In this case, machine learning techniques only matched the performance of other models (Ali et al., 2009). The reason is that time-series techniques perform well if sales are stable, which corroborate to one of the advantages of these techniques mentioned on the last section. For periods with promotions, on the other hand, regression trees performed better, but is a more complex technique and demands more intense data manipulation.

The authors propose a combination of these techniques in a forecasting system, applying time-series for periods when there is no promotion and regression tree model in periods with promotions. According to the authors,

> using more sophisticated input variables only helps when the prediction technique has the capacity to take advantage of them appropriately. Obviously, the use of more advanced techniques comes at a cost, namely the use of more extensive data

> (resulting in data preparation cost) and the maintenance of more complicated models. However, since the improvement is substantial, the benefits are likely to outweigh the costs involved (Ali et al., 2009, p.12341).

Sun et al. (2008) applied a neural network technique called extreme learning machine to forecast sales in fashion retailing. The advantages of that technique are the higher generalization performance and that it avoids difficulties of other learning methods (stopping criteria and long computation time, for example). The disadvantage is that “its solution is usually different from time to time because the input weights and hidden biases are randomly chosen” (Sun et al., 2008, p.412).

More recently (Liu et al., 2016) used big data from different sources (Twitter, Google Search, IMDB reviews, among others) to predict TV shows ratings. Cloud computing was applied, so it would be possible to process the large amount of data and to prepare the unstructured data was used text mining (machine learning technique). Alternative machine learning models were tested, but they were outperformed by the structured model chosen by the authors.

Although some efforts were already made in marketing to apply big-data based/computer intensive methods on demand forecasting, there is still space for evolution Cui & Curry (2005). To do so marketing researchers need to learn about (or associate with) other disciplines, like data science and machine learning (Chintagunta et al., 2016).

Finally, is important to keep in mind that the method should be chosen after the research problem is defined and the data to solve is accessible. These two criteria should determine the model and not the opposite (Reiss, 2011). The next section will review the types of data commonly used in demand forecasting marketing literature.

5. TYPES OF DATA USED IN DEMAND FORECASTING MODELS IN MARKETING

Choosing the data to structured causal models is related to the theory and the research problem wished to address. Researchers study the theory, the market, and develop a model based on assumptions. From this model researchers know what data they need to gather. Mela (2011) outlines a process to data procurement for structural models that consist in determining the data necessary to the research problem, finding the source (or sources) of data and then negotiate, acquire, and check the data. Research problems can also be derived from access to new types of data that were not available before. New types of data bring new opportunities to apply marketing theory.

In demand forecasting practice, according to Chase Jr (2013), the most common data sources used are: customer orders, customer shipments (or replenishment); point-of-sales data; promotions; price; Consumer Price Index; gross national product growth; and Consumer Confidence Index. The marketing literature, on the other hand, focus on marketing instruments, such as promotions, price, and their relation to sales or market share. The basic model studied in marketing, according to Fildes et al. (2008), is:

$$\text{Marketing response}_{ijt} = f\,(\text{marketing instruments; seasonality; exogenous factors}) \qquad (1)$$

i = brand or SKU
j = store
t = time
Marketing response = sales or market share
Marketing instruments = price, promotion, competition, among others.

Marketing theory has many potential explanatory variables to put on this basic model (Fildes et al., 2008). Researches should caution that a greater number of variables in a model can explain better, “but the extra variables actually make the model worse at forecasting”

(Chase Jr, 2013, p.86). For that reason, it is recommended to keep the model specification simple.

In marketing literature there is an emphasis on disaggregate analysis. This is due to the fact that the area has access to consumer panels linked with data on marketing instruments (Chintagunta & Nair, 2011). This means that the models in marketing deal with censored or truncated data (many zeros in the data). Also, it means that models on marketing are linked to economic theory, structural work, with emphasis on heterogeneity across consumers and considers that products are also differentiated (Chintagunta & Nair, 2011).

Mela (2011) divide the types of data in: firm property data, free public data, and commercially available market research. Public data and commercially available data are less customized to the research problem. On one hand, public data is free and commercially available data is often expensive. However, the latter is much quicker to obtain and easier to use (Mela, 2011). Firm property data is of great importance not only because it is rich but also because it is rare to have access. Also having access to people directly involved in the decision making on the firms can bring insights into the “rules of the game” (Mela, 2011, p.969). Those are important to specify the model.

Some of the variables applied to the basic model (1) that can be found in the marketing literature are: price; discounts; sales; market share; demographics (income; household size); attributes of products and services; competition; periods (of stockouts, of release of products, or of product-harm crisis); promotion (advertising expenditures, earned media; celebrity endorsements); geographical distance (based on zip-codes); number of alternatives; number of points-of-sales; word-of-mouth data (volume, variance and valence of posts from experts, peers, and critics); among others. The data sources of those variables are, as divided by Mela (2011):

- Firm property data, mainly scanner panel data (e.g., Albuquerque & Bronnenberg, 2012; Bollinger & Gillingham, 2012; Che et al., 2012; Yubo Chen et al., 2011; Yuxin Chen & Steckel, 2012; Ching et al., 2015; J. Chung & Rao, 2012; Draganska & Klapper, 2011; Feit et al., 2013; Jing & Lewis, 2011; Luan & Sudhir, 2010; Mehta & Ma, 2012; Narayanan & Nair, 2013; Shah et al., 2015; Stephen & Galak, 2012; Zhang & Kalra, 2014; Zhao et al., 2013)
- Free public data, from online word of mouth and Census (e.g., Albuquerque & Bronnenberg, 2012; Yuxin Chen & Steckel, 2012; P. K. Chintagunta, Gopinath, & Venkataraman, 2010; J. Chung & Rao, 2012; Gopinath, Chintagunta, & Venkataraman, 2013; X. Liu et al., 2016; Luan & Sudhir, 2010; Petrin & Train, 2010; Shriver, 2015; Stephen & Galak, 2012; Zhao et al., 2013)
- Commercially available market research (e.g., Bollinger & Gillingham, 2012; Briesch et al., 2013; Ching et al., 2015; P. K. Chintagunta et al., 2010; Derdenger & Kumar, 2013; Draganska & Klapper, 2011; Gopinath et al., 2013; Liu et al., 2016; Mukherjee & Kadiyali, 2011; Petrin & Train, 2010; Shriver, 2015; Yang et al., 2010; Zhao et al., 2011)
- Primary data, collected for the study, essentially experiments and surveys (e.g., Allenby et al., 2010; Aribarg et al., 2010; Arora et al., 2011; Yubo Chen et al., 2011; J. Chung & Rao, 2012; Rooderkerk et al., 2011)

As mentioned before, methods applied to demand forecasting are related to the characteristics of the variables and data sets that became available to marketing scientists. As developments are still occurring in big data, giving access to structured and unstructured data, many recent studies apply it on forecasting. One reason for it is that “anonymized web search data has proven to be helpful for other forecasts as well since online activity often is a good leading proxy for purchases and actions of the general public" (LaRiviere, McAfee, Rao, Narayanan, & Sun, 2016).

The bigger influx of information, though, is for those businesses that are already online. The connection between online and offline behavior is still a challenge (Liu et al., 2016), but there are some studies that attempt to find this relation. Chevalier & Mayzlin (2006) and Zhao et al. (2013) developed models to predict book sales with data from online book reviews. Dewan & Ramaprasad (2014) and Dhar & Chang (2009), studied the effect of blog posts and social network on sales of these music's singles and albums. Xiong & Bharadwaj (2014) used prerelease buzz to forecast video games sales. Godes & Mayzlin (2004) and Liu et al. (2016) investigate the effect of user generated content on TV ratings. Lastly, another relation investigated by authors is between word of mouth and sales in the movie industry (Chintagunta et al., 2010; Dellarocas, Zhang, & Awad, 2007; Gopinath et al., 2013; Karniouchina, 2011; Liu, 2006; Moon, Bergey, & Iacobucci, 2010; Onishi & Manchanda, 2012).

Studies on demand forecasting in marketing have aimed for contributions on method. Thus, they propose different approaches (often Bayesian) or modified versions of models to accommodate the data available and solve problems of endogeneity in structural models (e.g., Luan & Sudhir, 2010; Narayanan & Nair, 2013; Petrin & Train, 2010). The studies also aspire for substantial contribution applying marketing theory to econometric models, such as: brand awareness, choice complexity, satisfaction, word of mouth, conjoint purchase or promotion expenditures. On the next section, the focus will be on reviewing suggestions that have been already made, as contributions of future research on marketing literature for demand forecasting.

6. RESEARCH GAPS ON DEMAND FORECASTING IN MARKETING

Most of the articles on demand forecasting in marketing literature are structural models, as described in section 2. These models have many benefits but they also bring some challenges. The literature evidences that the challenge is "to develop structural models that provide realistic descriptions of the environments in which firms market products and services" (Reiss, 2011, p.963). To the practice of marketing they are challenging since the implementation of the models is not easy. Models like the basic model of demand studied in marketing (1), are not adopted by firms, and also

> evidence of improved accuracy is lacking […] there remains a need for simple operational models that include key marketing instruments and that are downwardly compatible in the product hierarchy (from category to brand to SKU). While the intellectual framework has been effectively laid down […] the practical questions examining the benefits in terms of forecast accuracy and price promotion planning and the level of complexity valuable in modeling the problem remain under-researched. (Fildes et al., 2008, p.1165).

Regarding the literature, models of demand forecasting that include marketing factors are econometrically endogenous. Therefore, the correction of such endogeneity is one of the contributions that authors seek. One way to address this issue is to use instrumental variables that are hard to choose, to justify, and do not correct fully the problem. They correct the intercept endogeneity, but not the slope endogeneity, which, according to Luan & Sudhir (2010), is lacking marketing researchers' attention. The authors define slope endogeneity as "private information possessed by managers about the heterogeneous effects of marketing-mix variables on sales" (Luan & Sudhir, 2010, p.444).

Another way to correct endogeneity is "to address the codetermination is to impose restrictions from an assumed model of supply […] into the demand estimation step"

(Chintagunta & Nair, 2011, p.980). Even if this is one way to address the issue, the authors consider that models in marketing literature are better at explaining demand than supply data and propose this as an area for future research. They also alert that supply side models are harder to estimate with the current data and computing power (Chintagunta & Nair, 2011).

Since the ability to work with big data “is not in many marketing researchers’ skill sets” (Feit et al., 2013, p.363), marketing area should gain knowledge in computer science. One of the opportunities for future researches on demand forecasting in marketing is the use of different computer intensive based methods, such as support vector machines and neural networks (Fildes et al., 2008).

The combination of data from new sources is another opportunity. Especially real-time data that is already used in “areas as diverse as stock-price trading, electricity load forecasting […] and could be applied more widely” (Fildes et al., 2008, p.1166-1167). These new sources of data need to be combined because databases do not have all the information about a customer’s behavior, even the best ones (Chen & Steckel, 2012). It remains an opportunity to combine data such as purchase, sensor, market, temporal (weather, traffic, etc.) and unstructured from social/mobile/digital/ecommerce (Chase Jr, 2013).

This is important for products that are not frequently purchased (durables), for example, that do not have rich data sets as those from package consumer goods (Chen & Steckel, 2012). Also methods can be combined to improve accuracy (Armstrong, 2001; Fildes et al., 2008), and “the more that data and methods differ, the greater the expected improvement in accuracy over the average of the individual forecasts” (Armstrong, 2001, p.419).

Another important opportunity concerns location (geographical) variables. This type of data is abundant nowadays because of mobile devices and applications (apps) that keep location metadata stored. They have been used, for example, to predict population movements

(Lu, Wetter, Bharti, Tatem, & Bengtsson, 2013). Also, data from satellite on the location of night lights has been used as an alternative to measure economic growth (Henderson, Storeygard, & Weil, 2012).

In marketing literature location data (although not metadata or satellite data, but zip codes and distances to stores/services) was used in models of demand forecasting about hotels (Zhang & Kalra, 2014); gas stations (Chan, Padmanabhan, & Seetharaman, 2007); fuel adoption (Shriver, 2015); drug prescriptions (Stremersch, Landsman, & Venkataraman, 2013); solar panels (Bollinger & Gillingham, 2012); organic products (Sridhar, Bezawada, & Trivedi, 2012); and the car industry (Albuquerque & Bronnenberg, 2012; Bucklin, Siddarth, & Silva-Risso, 2008; Narayanan & Nair, 2013).

Metadata has been used in other areas and proven to be able to uniquely identify individuals, "this means that knowing four random spatiotemporal points […] is enough to uniquely reidentify 90% of the individuals and to uncover all of their records" (De Montjoye, Hidalgo, Verleysen, & Blondel, 2013, p.537). There is still potential to explore location information such as studying similarities among customers preferences based on geographic location (Chung & Rao, 2012) or migration patterns influence on their purchases (Bronnenberg, Dubé, & Gentzkow, 2012), for example.

A good example of the use of location in models is Bucklin et al. (2008) study. Three measures of car dealers' concentration, accessibility and spread were developed based on the geographic locations of buyers and new car dealers on a choice model. The authors found that these three measures were significantly related to new car choice. These types of measures help the firms to decide the effects of opening or closing points-of-purchase. As the authors state, "practitioners also need to be able to assess how changes in distribution (e.g., the size and structure of a dealer network) might affect demand for their products and services" (Bucklin et al., 2008, p.473-474). According to the authors, only product, price, and promotion variables were incorporated as attributes in utility before their study. They conclude that “the empirical relationship between market share (or sales) for a product and its level of distribution intensity remains an open question in the case of consumer durables” (Bucklin et al., 2008, p.474).

Even if new methods are not applied, there is still chance for improvement in structural models that are the most commonly applied in marketing literature. Chintagunta & Nair (2011) suggest three new directions**:** dynamics; use of data on unobservables (primary data); and nonparametric approaches.

Dynamics means models that consider purchase on present impacts purchase in the future, such as: storability and durability. To model non-frequently purchased products such as durables is a challenge not only because of its dynamics but also because scanner data are not so common for these products (Zhao et al., 2011). The decisions concerning these type of products are also more sophisticated (involving more people and more time to decide), although empirical research on the subject is not frequent in marketing (Ni, Neslin, & Sun, 2012). For those reasons, “a thorough understanding of consumer decisions with respect to durables will help develop and test both economic and consumer behavior theories, and it will have important implications for managerial decisions” (Ni et al., 2012, p.1008).

Experience goods are also a dynamic problem because “purchase today provides a signal about quality, which updates the future information set” (Chintagunta & Nair, 2011, p.991). To close the discussion of dynamics issues, complementarities refer to products that are purchased only after another one is, so the choice of the complementary product is dependent on the choice of the first product (common in technological products due to compatibility issues).

Another direction is related to improve demand models with primary data that reduce possible confounds such as experiments mentioned on section 2. The final direction mentioned by Chintagunta & Nair (2011) is the use of nonparametric approaches due to the increased access to larger data sets. Some of these directions are actually overlapped, since dynamic problems such as intermittent demand (periods of no demand followed by periods of highly variable demand, which introduces lumpiness), can be improved by nonparametric approaches (Fildes et al., 2008). Also, Allenby et al. (2010) warn that current research is dominated by linear utility specifications and that it is necessary in structural models to use utility specifications that have more realistic assumptions.

Summarizing, the gaps found in the literature are:

- Develop models more realistic and simpler to implement;
- Find new approaches on correcting for endogeneity;
- Develop better supply side models;
- Gain knowledge in computer science to be able to work with big data;
- Incorporate new sources of data;
- Built models with location/geographical variables;
- Combine data sets and methods;
- Improve models with primary data;
- Consider dynamic issues: durable, experience goods, and complementarities;
- Use nonparametric approaches and more realistic assumptions on utility specifications.

7. CONCLUDING REMARKS

As the title of this theoretical paper informs, the focus was on demand forecasting methods, types of data and future research in marketing. To do that, we described the classification of models of demand forecasting in marketing. They were divided in two approaches: statistics/mathematics and big-data/computer intensive based methods. Two sections were made to discuss statistics/mathematics based methods, one about the literature and the other about practice. The section about demand forecasting in marketing literature explains different classifications of methods that were found: by demand systems; by level of data (aggregate or individual) and product launch period (pre or post-launch); and by their goal (forecasting, measurement, and testing). Types of model (descriptive, structural and reduced-form) were also described.

In the marketing literature the models are mostly structural (e.g., Albuquerque & Bronnenberg, 2012; Allenby et al., 2010; Bollinger & Gillingham, 2012; Che et al., 2012; Yubo Chen et al., 2011; Ching et al., 2015; Draganska & Klapper, 2011; Jing & Lewis, 2011; X. Liu et al., 2016; Luan & Sudhir, 2010; Mehta & Ma, 2012; Mukherjee & Kadiyali, 2011; Narayanan & Nair, 2013; Petrin & Train, 2010; Shah et al., 2015; Shriver, 2015; Stephen & Galak, 2012; Yang et al., 2010; Zhang & Kalra, 2014). Some also apply a reduced form model (e.g., Briesch et al., 2013; Chung et al., 2013) or a combination of reduced form and structural model (e.g., Chung et al., 2013).

The third section was about the techniques used in demand forecasting practice: time-series, causal and weighted combined forecasting. The main purpose of this section is to contrast the simple techniques used by practitioners in demand forecasting to the more sophisticated models available in marketing literature. A reason for that is that the models that result from marketing research are hard to implement in practice.

The section about big-data/computer intensive methods describes some algorithms that were already used in demand forecasting: support vector regression (SRM's regression version); CART (regression trees) and extreme learning machine (Ali et al., 2009; Sun et al., 2008). It was also discussed that this methods are used more intensively in the CRM literature and in other fields. In conclusion, it is stated that the application of such methods in marketing's demand forecasting studies should still evolve.

After that the types of data used in models of demand forecasting in marketing were introduced. They are divided in: firm property data that consists, mainly, in scanner panel data; free public data, from online word-of-mouth and Census; commercially available market research; and primary data, essentially experiments and surveys. On the last part of the section the contribution of big data to demand forecasting studies was discussed.

On the last section, the gaps found in the literature are presented and summarized. The gaps are related to the types of models applied in marketing literature (structural). There is a need for simpler models that are easier to implement; models that incorporate dynamic effects; that use primary data; and nonparametric approaches. Types of data and data sources used also brings some gaps to be explored in future research: data sets about durable goods; location/geographical data; big data; and the combination of different data sets.

Combining what was detailed in this theoretical paper about methods, data and the gaps found, it is possible to suggest possibilities for new studies on the field. A future study to contribute to the area would be one in the market of durables (non frequent purchase product). It would also combine different sources of data: free public data; firm property data; commercially available market research; big data; and primary data (surveys and experiments). The model would use location/geographical data, considering dynamic issues and it would be easy to implement. To do all of that, it would use knowledge from computer science. A possibility would be complementing structural models with SVM, as suggested by (Cui & Curry, 2005). Or, at least, the new study would combine two or more of those suggestions made by the authors.

REFERENCES

Albuquerque, P., & Bronnenberg, B. J. (2012). Measuring the impact of negative demand shocks on car dealer networks. *Marketing Science*, *31*(1), 4–23.

Ali, Ö. G., Sayın, S., Van Woensel, T., & Fransoo, J. (2009). SKU demand forecasting in the presence of promotions. *Expert Systems with Applications*, *36*(10), 12340–12348.

Allenby, G. M., Bakken, D. G., & Rossi, P. E. (2004). The HB revolution how Bayesian methods have changed the face of marketing research. *Marketing Research*, *16*, 20–25.

Allenby, G. M., Garratt, M. J., & Rossi, P. E. (2010). A model for trade-up and change in considered brands. *Marketing Science*, *29*(1), 40–56.

Aribarg, A., Arora, N., & Kang, M. Y. (2010). Predicting joint choice using individual data. *Marketing Science*, *29*(1), 139–157.

Armstrong, J. S. (2001). *Principles of forecasting: a handbook for researchers and practitioners* (Vol. 30). Springer Science & Business Media.

Arora, N., Henderson, T., & Liu, Q. (2011). Noncompensatory dyadic choices. *Marketing Science*, *30*(6), 1028–1047.

Beal, M., & Wilson, J. H. (2015). forecasting in the marketing curriculum. *International Journal of Business, Marketing, & Decision Science*, *8*(1).

Bollinger, B., & Gillingham, K. (2012). Peer effects in the diffusion of solar photovoltaic panels. *Marketing Science*, *31*(6), 900–912.

Briesch, R. A., Dillon, W. R., & Fox, E. J. (2013). Category positioning and store choice: The role of destination categories. *Marketing Science*, *32*(3), 488–509.

Bronnenberg, B. J., Dubé, J.-P. H., & Gentzkow, M. (2012). The evolution of brand preferences: Evidence from consumer migration. *The American Economic Review*, *102*(6), 2472–2508.

Bucklin, R. E., Siddarth, S., & Silva-Risso, J. M. (2008). Distribution intensity and new car choice. *Journal of Marketing Research*, *45*(4), 473–486.

Canitz, H. (2016). Overcoming Barriers to Improving Forecast Capabilities. *Foresight: The International Journal of Applied Forecasting*, (41), 26–34.

Chan, T. Y., Padmanabhan, V., & Seetharaman, P. B. (2007). An econometric model of location and pricing in the gasoline market. *Journal of Marketing Research*, *44*(4), 622–635.

Chase Jr, C. W. (2013). *Demand-driven forecasting: a structured approach to forecasting*. John Wiley & Sons.

Che, H., Chen, X., & Chen, Y. (2012). Investigating effects of out-of-stock on consumer stockkeeping unit choice. *Journal of Marketing Research*, *49*(4), 502–513.

Chen, Y., & Steckel, J. H. (2012). Modeling credit card share of wallet: Solving the incomplete information problem. *Journal of Marketing Research, 49*(5), 655–669.

Chen, Y., Wang, Q., & Xie, J. (2011). Online social interactions: A natural experiment on word of mouth versus observational learning. *Journal of Marketing Research*, *48*(2), 238–254.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, *43*(3), 345–354.

Ching, A. T., Clark, R., Horstmann, I., & Lim, H. (2015). The effects of publicity on demand: The case of anti-cholesterol drugs. *Marketing Science*, *35*(1), 158–181.

Chintagunta, P., Hanssens, D. M., & Hauser, J. R. (2016). Editorial—Marketing Science and Big Data. *Marketing Science, 35*(3), 341–342.

Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science*, *29*(5), 944–957.

Chintagunta, P. K., & Nair, H. S. (2011). Structural workshop paper-discrete-choice models of consumer demand in marketing. *Marketing Science, 30*(6), 977–996.

Chung, J., & Rao, V. R. (2012). A general consumer preference model for experience products: application to internet recommendation services. *Journal of Marketing Research*, *49*(3), 289–305.

Chung, K. Y., Derdenger, T. P., & Srinivasan, K. (2013). Economic value of celebrity endorsements: Tiger Woods' impact on sales of Nike golf balls. *Marketing Science*, *32*(2), 271–293.

Cui, D., & Curry, D. (2005). Prediction in marketing using the support vector machine. *Marketing Science*, *24*(4), 595–615.

De Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific Reports*, *3*, 1376.

Dellarocas, C., Zhang, X. M., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, *21*(4), 23–45.

Derdenger, T., & Kumar, V. (2013). The dynamic effects of bundling as a product strategy. *Marketing Science*, *32*(6), 827–859.

Dewan, S., & Ramaprasad, J. (2014). Social media, traditional media, and music sales. *Mis Quarterly*, *38*(1), 101–121.

Dhar, V., & Chang, E. A. (2009). Does chatter matter? The impact of user-generated content on music sales. *Journal of Interactive Marketing*, *23*(4), 300–307.

Draganska, M., & Klapper, D. (2011). Choice set heterogeneity and the role of advertising: An analysis with micro and macro data. *Journal of Marketing Research*, *48*(4), 653–669.

Feit, E. M., Wang, P., Bradlow, E. T., & Fader, P. S. (2013). Fusing aggregate and disaggregate data with an application to multiplatform media consumption. *Journal of Marketing Research*, *50*(3), 348–364.

Fildes, R., Nikolopoulos, K., Crone, S. F., & Syntetos, A. A. (2008). Forecasting and operational research: a review. *Journal of the Operational Research Society*, *59*(9), 1150–1172.

Godes, D., & Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing Science*, *23*(4), 545–560.

Gopinath, S., Chintagunta, P. K., & Venkataraman, S. (2013). Blogs, advertising, and local-market movie box office performance. *Management Science*, *59*(12), 2635–2654.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *The American Economic Review*, *102*(2), 994–1028.

Jing, X., & Lewis, M. (2011). Stockouts in online retailing. *Journal of Marketing Research*, *48*(2), 342–354.

Karniouchina, E. V. (2011). Impact of star and movie buzz on motion picture distribution and box office revenue. *International Journal of Research in Marketing*, *28*(1), 62–74.

LaRiviere, J., McAfee, P., Rao, J., Narayanan, V. K., & Sun, W. (2016, May 25). Where Predictive Analytics Is Having the Biggest Impact. Retrieved September 7, 2016, from https://hbr.org/2016/05/where-predictive-analytics-is-having-the-biggest-impact

Liu, X., Singh, P. V., & Srinivasan, K. (2016). A structured analysis of unstructured big data by leveraging cloud computing. *Marketing Science*, *35*(3), 363–388.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing*, *70*(3), 74–89.

Lu, X., Wetter, E., Bharti, N., Tatem, A. J., & Bengtsson, L. (2013). Approaching the limit of predictability in human mobility. *Scientific Reports*, *3*. Retrieved from http://www.nature.com/srep/2013/131011/srep02923/full/srep02923.html?message-global=remove&v=1392752313000/_/jcr:system/jcr:versionStorage/92/26/0c/92260c04-2e93-4f47-bdee-ca79793ce3c1/1.2/jcr:frozenNode

Luan, Y. J., & Sudhir, K. (2010). Forecasting marketing-mix responsiveness for new products. *Journal of Marketing Research*, *47*(3), 444–457.

Mehta, N., & Ma, Y. (2012). A multicategory model of consumers' purchase incidence, quantity, and brand choice decisions: Methodological issues and implications on promotional decisions. *Journal of Marketing Research*, *49*(4), 435–451.

Mela, C. F. (2011). Structural Workshop Paper-Data Selection and Procurement. *Marketing Science*, *30*(6), 965–976.

Moon, S., Bergey, P. K., & Iacobucci, D. (2010). Dynamic effects among movie ratings, movie revenues, and viewer satisfaction. *Journal of Marketing*, *74*(1), 108–121.

Mukherjee, A., & Kadiyali, V. (2011). Modeling multichannel home video demand in the US motion picture industry. *Journal of Marketing Research*, *48*(6), 985–995.

Narayanan, S., & Nair, H. S. (2013). Estimating causal installed-base effects: A bias-correction approach. *Journal of Marketing Research*, *50*(1), 70–94.

Nevo, A. (2011). Empirical Models of Consumer Behavior. *Annual Review of Economics*, *3*(1), 51–75. https://doi.org/10.1146/annurev-economics-061109-080402

Ngai, E. W., Xiu, L., & Chau, D. C. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, *36*(2), 2592–2602.

Ni, J., Neslin, S. A., & Sun, B. (2012). Database submission-the isms durable goods data sets. *Marketing Science*, *31*(6), 1008–1013.

Onishi, H., & Manchanda, P. (2012). Marketing activity, blogging and sales. *International Journal of Research in Marketing*, *29*(3), 221–234.

Petrin, A., & Train, K. (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research*, *47*(1), 3–13.

Reiss, P. C. (2011). Structural Workshop Paper-Descriptive, Structural, and Experimental Empirical Methods in Marketing Research. *Marketing Science*, *30*(6), 950–964.

Roberts, J. H. (1998). Marketing approaches to forecasting problems. *Journal of Forecasting*, *17*(3–4), 169–174.

Rooderkerk, R. P., Van Heerde, H. J., & Bijmolt, T. H. (2011). Incorporating context effects into a choice model. *Journal of Marketing Research*, *48*(4), 767–780.

Rossi, P. E. (2014). Invited Paper—Even the Rich Can Make Themselves Poor: A Critical Examination of IV Methods in Marketing Applications. *Marketing Science*, *33*(5), 655–672.

Rossi, P. E., & Allenby, G. M. (2003). Bayesian statistics and marketing. *Marketing Science*, *22*(3), 304–328.

Shah, D., Kumar, V., & Zhao, Y. (2015). Diagnosing Brand Performance: Accounting for the Dynamic Impact of Product Availability with Aggregate Data. *Journal of Marketing Research*, *52*(2), 147–165.

Shriver, S. K. (2015). Network effects in alternative fuel adoption: Empirical analysis of the market for ethanol. *Marketing Science*, *34*(1), 78–97.

Singh, S. (2016). Forecasting: Academia versus Business. *Foresight: The International Journal of Applied Forecasting*, (41), 46–47.

Sridhar, K., Bezawada, R., & Trivedi, M. (2012). Investigating the drivers of consumer cross-category learning for new products using multiple data sets. *Marketing Science*, *31*(4), 668–688.

Stephen, A. T., & Galak, J. (2012). The effects of traditional and social earned media on sales: A study of a microlending marketplace. *Journal of Marketing Research*, *49*(5), 624–639.

Stremersch, S., Landsman, V., & Venkataraman, S. (2013). The relationship between DTCA, drug requests, and prescriptions: Uncovering variation in specialty and space. *Marketing Science*, *32*(1), 89–110.

Sun, Z.-L., Choi, T.-M., Au, K.-F., & Yu, Y. (2008). Sales forecasting using extreme learning machine with applications in fashion retailing. *Decision Support Systems*, *46*(1), 411–419.

Sundsøy, P. al, Bjelland, J., Iqbal, A. M., de Montjoye, Y.-A., & others. (2014). Big Data-Driven Marketing: How machine learning outperforms marketers' gut-feeling. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction* (pp. 367–374). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-05579-4_45

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., … others. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*(1), 1–37.

Xiong, G., & Bharadwaj, S. (2014). Prerelease buzz evolution patterns and new product performance. *Marketing Science*, *33*(3), 401–421.

Yang, S., Zhao, Y., Erdem, T., & Zhao, Y. (2010). Modeling the intrahousehold behavioral interaction. *Journal of Marketing Research*, *47*(3), 470–484.

Zhang, W., & Kalra, A. (2014). A Joint Examination of Quality Choice and Satisfaction: The Impact of Circumstantial Variables. *Journal of Marketing Research*, *51*(4), 448–462.

Zhao, Y., Yang, S., Narayan, V., & Zhao, Y. (2013). Modeling consumer learning from online product reviews. *Marketing Science*, *32*(1), 153–169.

Zhao, Y., Zhao, Y., & Helsen, K. (2011). Consumer learning in a turbulent market environment: Modeling consumer choice dynamics after a product-harm crisis. *Journal of Marketing Research*, *48*(2), 255–267.