Marketing Analytics for Data-Rich Environments

Author(s): Michel Wedel and P.K. Kannan

Source: *Journal of Marketing*, Vol. 80, No. 6, AMA and MSI Special Issue of "Journal of Marketing" | Mapping the Boundaries of Marketing: What Needs to Be Known (November 2016), pp. 97-121

Published by: Sage Publications, Inc.

Stable URL: https://www.jstor.org/stable/44134975

Accessed: 24-01-2020 20:23 UTC

Michel Wedel & P.K. Kannan

# Marketing Analytics for Data-Rich Environments

The authors provide a critical examination of marketing analytics methods by tracing their historical development, examining their applications to structured and unstructured data generated within or external to a firm, and reviewing their potential to support marketing decisions. The authors identify directions for new analytical research methods, addressing (1) analytics for optimizing marketing-mix spending in a data-rich environment, (2) analytics for personalization, and (3) analytics in the context of customers' privacy and data security. They review the implications for organizations that intend to implement big data analytics. Finally, turning to the future, the authors identify trends that will shape marketing analytics as a discipline as well as marketing analytics education.

*Keywords*: big data, marketing analytics, marketing mix, personalization, privacy

*Online Supplement*: http://dx.doi.org/10.1509/jm.15.0413

Data has been called "the oil" of the digital economy. The routine capture of digital information through online and mobile applications produces vast data streams on how consumers feel, behave, and interact around products and services as well as how they respond to marketing efforts. Data are assuming an increasingly central role in organizations, as marketers aim to harness data to build and maintain customer relationships; personalize products, services, and the marketing mix; and automate marketing processes in real time. The explosive growth of media, channels, digital devices, and software applications has provided firms with unprecedented opportunities to leverage data to offer more value to customers, enhance their experiences, increase their satisfaction and loyalty, and extract value. Although big data's potential may have been overhyped initially, and companies may have invested too much in data capture and storage and not enough in analytics, it is becoming clear that the availability of big data is spawning data-driven decision cultures in companies, providing them with competitive advantages, and having a significant impact on their financial performance. The increasingly widespread recognition that big data can be leveraged effectively to support marketing decisions is highlighted by the success of industry leaders. Entirely new forms of marketing have emerged, including recommendations, geo-fencing, search marketing, and retargeting. Marketing analytics has come to play a central role in these developments, and there is urgent demand for new, more powerful metrics and analytical methods that make data-driven marketing operations more efficient and effective. However, it is yet not sufficiently clear which types of analytics work for which types of problems and data, what new methods are needed for analyzing new types of data, or how companies and their management should evolve to develop and implement skills and procedures to compete in this new environment.

The Marketing Science Institute has outlined the scope of research priorities around these issues.[1] The present article provides a review of research on one of these priorities: analytics for data-rich environments. We have structured our thoughts using the framework in Figure 1. At the center is the use of analytics to support marketing decisions, which is founded, on the one hand, on the availability of data and, on the other hand, on advances in analytical methods. Key domains for analytics applications are (1) customer relationship management (CRM), with methods that help acquisition, retention, and satisfaction of customers to improve their lifetime value to the firm[2]; (2) the marketing mix, with methods, models, and algorithms that support the allocation of resources to enhance the effectiveness of marketing effort; (3) personalization of the marketing mix to individual consumers, in which significant advances have been made as a result of the development of various approaches to capture customer heterogeneity; and (4) privacy and security, an area that is of increasing concern to firms and regulators. These domains lead to two pillars of the successful development and implementation of marketing analytics in firms: (1) the adoption of organizational structures and cultures that foster data-driven decision making and (2) the education and training of analytics professionals.

Michel Wedel is PepsiCo Chaired Professor of Consumer Science and Distinguished University Professor, Robert H. Smith School of Business, University of Maryland, College Park (e-mail: mwedel@rhsmith.umd.edu). P.K. Kannan is Ralph J. Tyser Professor of Marketing Science, Robert H. Smith School of Business, University of Maryland, College Park (e-mail: pkannan@rhsmith.umd.edu).

[1] See http://www.msi.org/articles/marketers-top-concerns-frame-2014-16-research-priorities/.

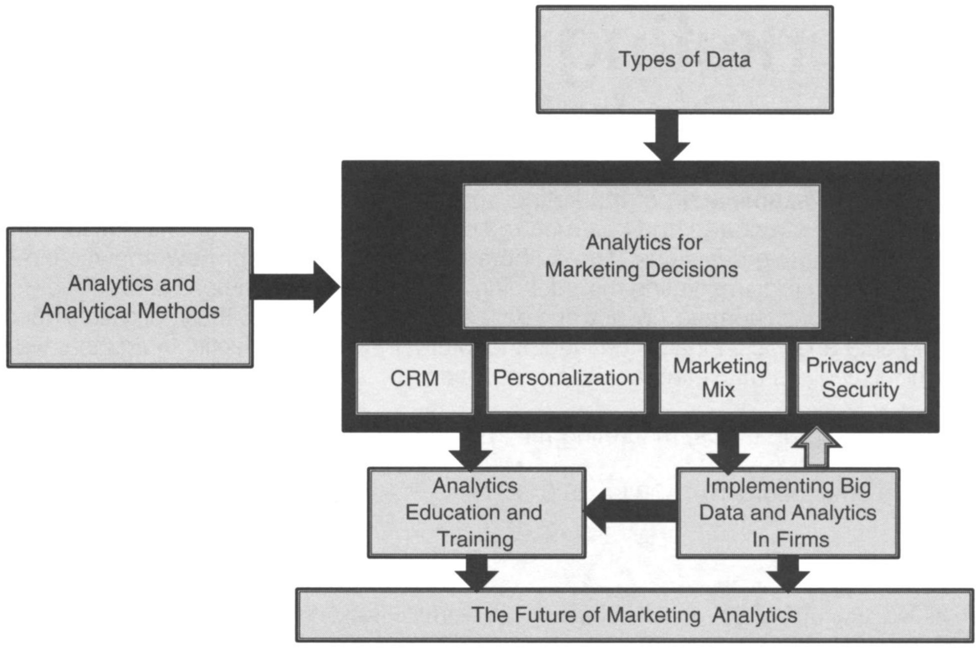
[2] We do not focus on CRM issues beyond personalization in this article, because another article in this issue covers CRM in depth.


**ISSN: 0022-2429 (print)**
**1547-7185 (electronic)**



***Journal of Marketing: AMA/MSI Special Issue***
**Vol. 80 (November 2016), 97–121**
**DOI: 10.1509/jm.15.0413**

**FIGURE 1**
**Article Framework**




Notes: Marketing data and analytical methods are used in four main areas of marketing decisions. Their implementation in firms depends on firm culture and organizational structure and poses requirements for education and training, which will shape the future of marketing analytics.

The agenda for this article is as follows. Using the framework in Figure 1, we provide a brief review of the history of marketing data and analytics, followed by a critical examination of the extent to which specific analytical methods are applicable in data-rich environments and support marketing decision making in core domains. This analysis leads to the identification of future research directions. We choose to focus on (1) analytics for optimizing marketing-mix spending, (2) analytics for personalization of the marketing mix, and (3) analytics in the context of data security and customer privacy. We review the implications for implementing big data analytics in organizations and for analytics education and training. In doing so, we identify trends that will shape marketing as a discipline, and we discuss actual and aspired interconnections between marketing practice and academia.

# A Brief History of Marketing Data and Analytics

Marketing analytics involves collection, management, and analysis—descriptive, diagnostic, predictive, and prescriptive—of data to obtain insights into marketing performance, maximize the effectiveness of instruments of marketing control, and optimize firms' return on investment (ROI). It is interdisciplinary, being at the nexus of marketing and other areas of business, mathematics, statistics, economics, econometrics, psychology, psychometrics, and, more recently, computer science. Marketing analytics has a long history, and as a result of explosive growth in the availability of data in the digital economy in the last two decades, firms have increasingly recognized the key competitive advantages that analytics may afford, which has propelled its development and deployment (Davenport 2006).

### *Available Data*

The history of the systematic use of data in marketing starts around 1910 with the work of Charles Coolidge Parlin for the Curtis Publishing Company in Boston (Bartels 1988, p. 125). Parlin gathered information on markets to guide advertising and other business practices, prompting several major U.S. companies to establish commercial research departments. Duncan (1919) emphasized the use of external in addition to internal data by these departments. Questionnaire survey research, already conducted in the context of opinion polls by Gallup in the 1820s, became increasingly popular in the 1920s (Reilly 1929). Around that time, concepts from psychology were being brought into marketing to foster greater understanding of the consumer. Starch's (1923) attention, interest, desire, action (AIDA) model is a prime example, and he is credited for the widespread adoption of copy research. This era also saw the first use of eye-tracking data (Nixon 1924).

In 1923, A.C. Nielsen founded one of the first market research companies. Nielsen started by measuring product sales in stores, and in the 1930s and 1950s, he began assessing radio and television audiences. In 1931, the market research firm Burke was founded in the United States, and it initially did product testing research for Procter & Gamble. In 1934, the market research firm GfK was established in Germany. The next decade saw the rise of field experiments and the increased use of telephone surveys (White 1931).

Panel data became increasingly popular, at first mostly for measuring media exposure, but in the 1940s firms began using panel data to record consumer purchases (Stonborough 1942). George Cullinan, who introduced the "recency, frequency, monetary" metrics that became central in CRM (Neslin 2014), stimulated the use of companies' own customer data beginning in 1961. In 1966, the Selling Areas Marketing Institute was founded, which focused on warehouse withdrawal data. The importance of computers for marketing research was first recognized around that time as well (Casher 1969).

Beginning in the late 1970s, geo-demographic data were amassed from government databases and credit agencies by the market research firm Claritas, founded on the work by the sociologist Charles Booth around 1890. The introduction of the Universal Product Code and IBM's computerized point-of-sale scanning devices in food retailing in 1972 marked the first automated capture of data by retailers. Companies such as Nielsen quickly recognized the promise of using point-of-sale scanner data for research purposes and replaced bimonthly store audits with more granular scanner data. Soon, individual customers could be traced through loyalty cards, which led to the emergence of scanner panel data (Guadagni and Little 1983). The market research firm IRI, which measured television advertising since the company's founding in 1979, rolled out its in-home barcode scanning service in 1995.

The use of internal customer data was greatly propelled by the introduction of the personal computer to the mass market by IBM in 1981. Personal computers enabled marketers to store data on current and prospective customers, which contributed to the emergence of database marketing, pioneered by Robert and Kate Kestnbaum and Robert Shaw (1987). In 1990, CRM software emerged, for which earlier work on sales force automation at Siebel Systems paved the way. Personal computers also facilitated survey research through personal and telephone interviewing.

In 1995, after more than two decades of development at the Defense Advanced Research Projects Agency and other organizations, the World Wide Web came into existence, and this led to the availability of large volumes of marketing data. Clickstream data extracted from server logs were used to track page views and clicks using cookies. Click-through data yielded measures of the effectiveness of online advertising. The Internet stimulated the development of CRM systems by firms such as Oracle, and in 1999 Salesforce was the first company to deliver CRM systems through cloud computing.

Google was founded in 1998, and it championed keyword search and the capture of search data. Search engines had been around since the previous decade; the first file transfer protocol search engine Archie was developed at McGill University. The advent of user-generated content, including online product reviews, blogs, and video, resulted in increasing volume and variety of data. The launch of Facebook in 2004 opened up an era of social network data. With the advent of YouTube in 2005, vast amounts of data in the form of user-uploaded text and video became the raw material for behavioral targeting. Twitter, with its much simpler 140-character messages, followed suit in 2006. Smartphones had existed since the early 1990s, but the introduction of the Apple iPhone in 2007, with its global positioning system (GPS) capabilities, marked the onset of the capture of consumer location data at an unprecedented scale.

### Analytics

The initiative of the Ford Foundation and the Harvard Institute of Basic Mathematics for Applications in Business (in 1959/1960) is widely credited for having provided the major impetus for the application of analytics to marketing (Winer and Neslin 2014). It led to the founding of the Marketing Science Institute in 1961, which has since had a continued role in bridging marketing academia and practice. Statistical methods (e.g., analysis of variance) had been applied in marketing research for more than a decade (Ferber 1949), but the development of statistical and econometric models tailored to specific marketing problems took off when marketing was recognized as a field of decision making through the Ford/Harvard initiative (Bartels 1988, p. 125). The development of Bayesian decision theory at the Harvard Institute (Raiffa and Schlaifer 1961) also played a role, exemplified by its successful application to, among other things, pricing decisions by Green (1963). Academic research in marketing began to focus more on the development of statistical models and predictive analytics. Although it is not possible to review all subsequent developments here (for an extensive review, see Winer and Neslin 2014), we note a few landmarks.

New product diffusion models (Bass 1969) involved applications of differential equations from epidemiology. Stochastic models of buyer behavior (Massy, Montgomery, and Morrison 1970) were rooted in statistics and involved distributional assumptions on measures of consumers' purchase behavior. The application of decision calculus (Little and Lodish 1969; Lodish 1971) to optimize spending on advertising and the sales force became popular after its introduction to marketing by Little (1970). Market share and demand models for store-level scanner data (Nakanishi and Cooper 1974) were derived from econometric models of demand. Multidimensional scaling and unfolding techniques, founded in psychometrics (Coombs 1950), became an active area of research, with key contributions by Green (1969) and DeSarbo (DeSarbo and Rao 1986). These techniques paved the way for market structure and product positioning research by deriving spatial maps from proximity and preference judgments and choice. Conjoint analysis (Green and Srinivasan 1978) and, later, conjoint choice analysis (Louviére and Woodworth 1983) are unique contributions that evolved from work in psychometrics by Luce on the quantification of psychological attributes (Luce and Tukey 1964). Scanner panel–based multinomial logit models (Guadagni and Little 1983) were built directly on research in econometrics by McFadden (1974). The nested logit model that captures hierarchical consumer decision making was introduced in marketing (Kannan and Wright 1991), and it was recognized that models of multiple aspects of consumer behavior (e.g., incidence, choice, timing, quantity) could be integrated (Gupta 1988). This proved to be a powerful insight for models of recency, frequency, and monetary metrics (Schmittlein and Peterson 1994). Whereas previous methods to identify competitive market structures were based on estimated cross-price elasticities, models that derive competitive maps from panel choice data were developed on the basis of the notion that competitive market structures arise from consumer perceptions of substitutability, revealed through their choices of products (Elrod 1988). Time-series methods (DeKimpe and Hanssens 1995) enabled researchers to test whether marketing instruments resulted in permanent or transient changes in sales.

Heterogeneity in the behaviors of individual consumers became a core premise on which marketing strategy was based, and the mixture choice model was the first to enable managers to identify response-based consumer segments from scanner data (Kamakura and Russell 1989). This model was generalized to accommodate a wide range of models of consumer behavior (Wedel and DeSarbo 1995). Consumer heterogeneity was represented in a continuous fashion in hierarchical Bayes models (Rossi, McCulloch, and Allenby 1996). Although scholars initially debated which of these two approaches best represented heterogeneity, research has shown that the approaches each match specific types of marketing problems, with few differences between them (Andrews, Ainslie, and Currim 2002). It can be safely said that the Bayesian approach is now one of the dominant modeling approaches in marketing, offering a powerful framework to develop integrated models of consumer behavior (Rossi and Allenby 2003). Such models have been successfully applied to advertisement eye tracking (Wedel and Pieters 2000), e-mail marketing (Ansari and Mela 2003), web browsing (Montgomery et al. 2004), social networks (Moe and Trusov 2011), and paid search advertising (Rutz, Trusov, and Bucklin 2011).

The derivation of profit-maximizing decisions, inspired by the work of Dorfman and Steiner (1954) in economics, formed the basis of the operations research (OR) approach to optimal decision making in advertising (Parsons and Bass 1971), sales force allocation (Mantrala, Sinha, and Zoltners 1994), target selection in direct marketing (Bult and Wansbeek 1995), and customization of online price discounts (Zhang and Krishnamurthi 2004). Structural models founded in economics include approaches that supplement aggregate demand equations with supply-side equilibrium assumptions (Chintagunta 2002), based on the work of the economists Berry, Levinsohn, and Pakes (1995). A second class of structural models accommodates forward-looking behavior (Erdem and Keane 1996), based on work in economics by Rust (1987). Structural models allow for predictions of agent shifts in behavior when policy changes are implemented (Chintagunta et al. 2006).

### From Theory to Practice

Roberts, Kayande, and Stremersch (2014) empirically demonstrate the impact of these academic developments on marketing practice. Through interviews among managers, they find a significant impact of several analytics tools on firm decision making. The relevance of these developments for the practice of marketing is further evidenced by examples of companies that were founded on academic work. Early cases of successful companies include Starch and Associates, a company that specialized in ad copy testing based on Starch's academic work, and John D.C. Little and Glen L. Urban's Management Decision Systems, which was later sold to IRI. Zoltman and Sinha's work on sales force allocation was implemented in practice through ZS Associates. Claes Fornell's work on the measurement of satisfaction led to the American Consumer Satisfaction Index, produced by his company, CFI Group. MarketShare, the company cofounded by Dominique Hanssens, successfully implemented his models on the long-term effectiveness of the marketing mix. Jan-Benedict E.M. Steenkamp founded AiMark, a joint venture with GfK that applies academic methods and concepts particularly in international marketing. Virtually all of these companies became successful through the application of analytics.

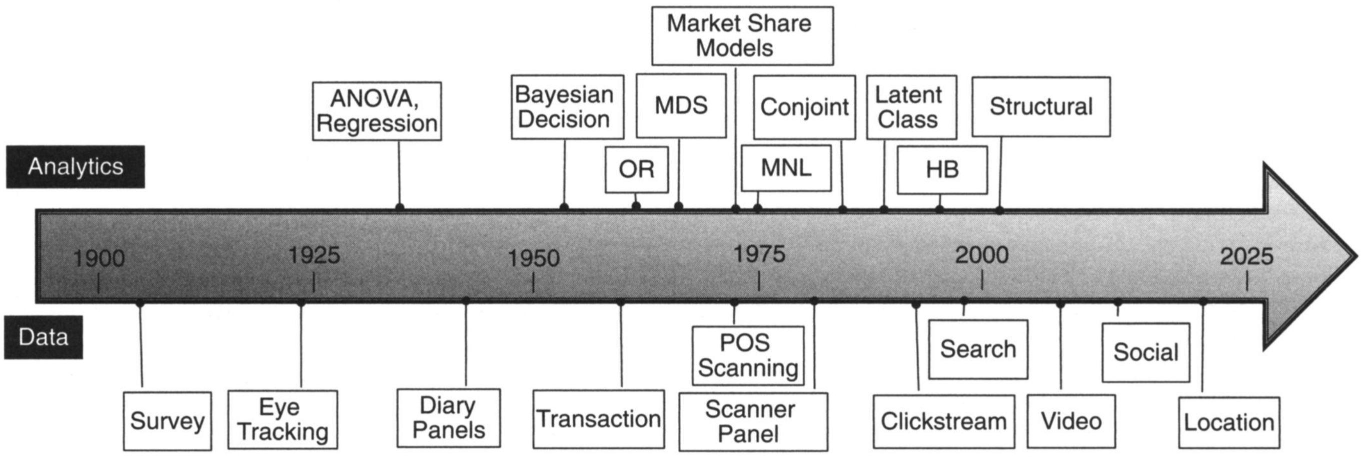
Examples of companies with very close ties to academia include Richard M. Johnson's Sawtooth Software, which specializes in the design and analysis of and software for conjoint studies, and Steven Cohen and Mark Garratt's In4mation Insights, which applies comprehensive Bayesian statistical models to a wide range of applied problems including marketing-mix modeling. In some cases, marketing academia lags behind developments in practice and so focuses instead on the impact and validity of these developments in practice. In other cases, academics are coinvestigators who rely on data and problems provided by companies and work together with these companies to develop implementable analytics solutions. Yet, as we discuss next, in an increasing number of application areas in the digital economy, academics are leading the development of new concepts and methods.

### Synthesis

The development of data-driven analytics in marketing from around 1900 until the introduction of the World Wide Web in 1995 has progressed through approximately three stages: (1) the description of observable market conditions through simple statistical approaches, (2) the development of models to provide insights and diagnostics using theories from economics and psychology, and (3) the evaluation of marketing policies, in which their effects are predicted and marketing decision making is supported using statistical, econometric, and OR approaches. In many cases throughout the history of marketing analytics, soon after new sources of data became available, methods to analyze them were introduced or developed (for an outline of the history of data and analytical methods, see Figure 2; Table 1 summarizes state-of-the-art approaches). Many of the methods developed by marketing academics since the 1960s have now found their way into practice and support decision making in areas such as CRM, marketing mix, and personalization and have increased the financial performance of the firms deploying them.

Since 2000, the automated capture of online clickstream, messaging, word-of-mouth (WOM), transaction, and location data has greatly reduced the variable cost of data collection and has resulted in unprecedented volumes of data that provide insights on consumer behavior at exceptional levels of depth and granularity. Although academics have

**FIGURE 2**
**An Outline of the Timeline of Marketing Data and Analytics**




Notes: ANOVA = analysis of variance; MDS = multidimensional scaling; POS = point of sale; MNL = multinomial logit model; HB = hierarchical Bayes. This timeline summarizes the availability of new marketing data and the development of the major classes of marketing models. As new types of data became available, new models to analyze them followed.

taken up the challenge to develop diagnostic and predictive models for these data in the last decade, these developments are admittedly still in their infancy. On the one hand, descriptive metrics displayed on dashboards are popular in practice. This could be the result of constraints on computing power, a need for rapid real-time insights, a lack of trained analysts, and/or the presence of organizational barriers to implementing advanced analytics. In particular, unstructured data in the form of blogs, reviews, and tweets offer opportunities for deep insights into the economics and psychology of consumer behavior, which could usher in the second stage in digital marketing analytics once appropriate models

**TABLE 1**
**Marketing Analytics: State-of-the-Art Approaches and Their Applications**

| Area of Focus | Developments and State-of-the-Art Approaches |
|---|---|
| Data | |
| Structured data | • A plethora of descriptive, diagnostic, predictive, and prescriptive methods for analytics in many areas of marketing are available<br>• Approaches to deal with big data include those using Bayesian methods, data aggregation and data compression methods, sampling and variable selection methods, approximations and model simplifications, efficient Markov chain Monte Carlo algorithms, and parallel computing<br>• Field and quasi-experiments, instrumental variables (IV) and instrument-free approaches to endogeneity, regression discontinuity approaches |
| Unstructured data | • Mostly descriptive and diagnostic analytical methods, predictive and prescriptive methods still play catch-up<br>• Text mining and machine learning approaches<br>• Incorporating structure through metrics for text, audio, image, and video data; eye tracking, face recognition, and other neurodata |
| Marketing-mix modeling | • Modeling effects of social networks, keyword search, online WOM, trending, and mobile/location within the marketing mix<br>• Analysis of entire path to purchase, attribution modeling<br>• Incorporating specific institutional settings and contexts to enhance estimation of structural models and their policy simulations; better instrumental variables to address endogeneity; field and quasi-experiments for causal effects |
| Personalization | • Online and mobile personalization of the marketing mix<br>• Dealing with missing observations and incorporating receptivity into recommendations<br>• Adaptive personalization approaches—learning and adapting to users' changes in preferences in a continuous automated cycle |
| Security and privacy | • Research into the effects of privacy and security regulations and policies on consumer behavior and competition between firms<br>• Models to analyze minimized and anonymized data |

are developed and applied. On the other hand, machine learning methods from computer science (including deep neural networks and cognitive systems, which we discuss subsequently; see Table 1) have become popular in practice but have been infrequently researched in marketing academia. Their popularity may stem from their excellent predictive performance and black-box nature, which allows for routine application with limited analyst intervention. The question is whether marketing academics should jump on the machine learning bandwagon, something they may have been reluctant to do because these techniques do not establish causal effects or produce generalizable theoretical insights. However, combining these approaches with more classical models for marketing analytics may address these shortcomings and hold promise for further research (Table 2). It is reasonable to expect that the third step in the evolution of analytics in the digital economy—the development of models to generate diagnostic insights and support real-time decisions from big data—is imminent. However, marketing academia will need to develop analytical methods with a keen eye for data volume and variety as well as speed of computation, components that have thus far been largely ignored (see Table 2). In the remainder of this article, we review recent developments and identify potential barriers and opportunities toward successful implementation of analytics to support marketing decisions in data-rich environments.

# Data and Analytics

## *Types of Data*

Big data is often characterized by the four "Vs": volume (from terabytes to petabytes), velocity (from one-time snapshots to high-frequency and streaming data), variety (numeric, network, text, images, and video), and veracity (reliability and validity). The first two characteristics are important from a computing standpoint, and the second two are important from an analytics standpoint. Sometimes a fifth "V" is added: value. It transcends the first four and is important from a business standpoint. Big data is mostly observational, but surveys, field experiments, and lab experiments may yield data of large variety and high velocity. Much of the excitement surrounding big data is exemplified by the scale and scope of observational data generated by the "big three" of big data: Google, Amazon, and Facebook. Google receives more than 4 million search queries per minute from the 2.4 billion Internet users around the world and processes 20 petabytes of information per day. Facebook's 1.3 billion users share 2.5 million pieces of content each minute. Amazon has created a marketplace with 278 million active customers from which it records data on online browsing and purchasing behavior. These and other firms have changed the landscape of marketing in the last decade through the generation, provision, and utilization of big data.

**TABLE 2**
**Marketing Analytics: Issues for Further Research**

| Area of Focus | Promising and Important Issues for Research |
|---|---|
| Data | |
| Structured data | • Behavioral targeting with cross-device data; mobile, location-based, and social analytics<br>• Fusing data generated within the firm with data generated outside the firm; integrating "small stats on big data" with "big stats on small data" approaches<br>• Combining machine learning approaches with econometric and theory-based methods for big data applications; computational solutions to marketing models for big data |
| Unstructured data | • Development of diagnostic, predictive, and prescriptive approaches for analysis of large-scale unstructured data<br>• Approaches to analyze unstructured social, geo-spatial, mobile data and combining them with structured data in big data contexts<br>• Using, evaluating, and extending deep learning methods and cognitive computing to analyze unstructured marketing data |
| Marketing-mix modeling | • Aligning analysis of disaggregate data with that of aggregate data and including unstructured data in the analysis of the marketing mix<br>• New techniques and methods to accurately measure the impact of marketing instruments and their carryover and spillover across media and devices using integrated path-to-purchase data<br>• Dynamic, multi–time period and cross-category optimization of the marketing mix<br>• Approaches to incorporate different planning cycles for different marketing instruments in media-mix models |
| Personalization | • Automated closed-loop marketing solutions for digital environments; fully automated marketing solutions<br>• Personalization and customization techniques using cognitive systems, general artificial intelligence, and automated attention analysis; personalization of content<br>• Mobile, location-based personalization of the marketing mix |
| Security and privacy | • Methods to produce and handle data minimization and data anonymization in assessing marketing-mix effectiveness and personalization<br>• Distributed data solutions to enhance data security and privacy while maximizing personalized marketing opportunities |

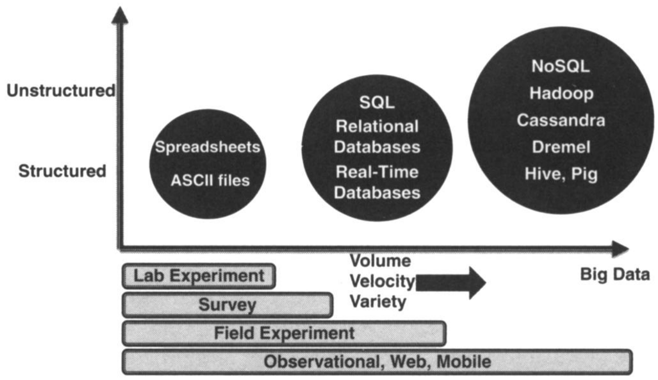
Emerging solutions to link customer data across online and offline channels and across television, tablet, mobile, and other digital devices will further contribute to the availability of data. Moreover, in 2014, well over 15 billion devices were equipped with sensors that enable them to connect and transfer data over networks without human interaction. This "Internet of Things" may become a major source of new product and service development and generate massive data in the process.

Surveys have become much easier to administer with the advances in technology allowing for online and mobile data collection (e.g., Amazon Mechanical Turk). Firms continuously assess customer satisfaction; new digital interfaces require this to be done with short surveys to reduce fatigue and attrition. For example, loyalty is often evaluated with single-item Net Promoter Scores. As a consequence, longitudinal and repeated cross-section data are becoming more common. Mittal, Kumar, and Tsiros (1999) use such data to track the drivers of customer loyalty over time. To address the issue of shorter questionnaires, analytic techniques have been developed to create personalized surveys that are adaptive on the basis of the responses to earlier questions (Kamakura and Wedel 1995) as well as the design of tailored split-questionnaires for massive surveys (Adigüzel and Wedel 2008).

Digital technologies facilitate large-scale field experiments that produce big data and have become powerful tools for eliciting answers to questions on the causal effects of marketing actions. For example, large-scale A/B testing enables firms to "test and learn" for optimizing website designs, (search, social, and mobile) advertising, behavioral targeting, and other aspects of the marketing mix. Hui et al. (2013) use field experiments to evaluate mobile promotions in retail stores. Alternatively, natural (or quasi-) experiments capitalize on exogenous shocks that occur naturally in the data to establish causal relations, but often more extensive analytical methods (including matching and instrumental variables methods) are required to establish causality. For example, Ailawadi et al. (2010) show how quasi-experimental designs can be used to evaluate the impact of the entry of Wal-Mart stores on retailers, using a before-and-after design with a control group of stores matched on a variety of measures. Another way to leverage big data to assess causality is to examine thin slices of data around policy changes that occur in the data, which can reveal the impact of those changes on dependent variables of interest through so-called regression discontinuity designs (Hartmann, Nair, and Narayanan 2011).

Finally, lab experiments typically generate smaller volumes of data, but technological advances have allowed for online administration and collection of audio, video, eye-tracking, face-tracking (Teixeira, Wedel, and Pieters 2010), and neuromarketing data obtained from electroencephalography and brain imaging (Telpaz, Webb, and Levy 2015). Such data are collected routinely by firms such as Nielsen, and they often yield p > n data with more variables than respondents. Meta-analysis techniques can be used to generalize findings across large numbers of these experiments (Bijmolt, Van Heerde, and Pieters 2005).

**FIGURE 3**
**Managing Types of Data**




Notes: The vertical axis shows the degree of structure in the data and the horizontal axis shows the dimensions resulting in big data. Software to manage these data appear in the core of the figure.

### *Software for Big Data Processing*[3]

Figure 3 provides an overview of the classes of marketing data discussed previously and methods to store and manipulate it. For small to medium-sized structured data, the conventional methods such as Excel spreadsheets; ASCII files; or data sets of statistical packages such as SAS, S-Plus, STATA, and SPSS are adequate. SAS holds up particularly well as data size increases and is popular in many industry sectors (e.g., retailing, financial services, government) for that reason. As the number of records goes into the millions, relational databases such as MySQL (used by, e.g., Wikipedia) are increasingly effective for data manipulation and for querying. For big and real-time web applications in which volume, variety, and velocity are high, databases such as NoSQL are the preferred choice because they provide a mechanism for storage and retrieval of data that does not require tabular relations like those in relational databases, and they can be scaled out across commodity hardware. Apache Cassandra, an open-source software initially developed by Facebook, is a good example of such a distributed database management system. Hadoop, originally developed at Yahoo!, is a system to store and manipulate data across a multitude of computers, written in the Java programming language. At its core are the Hadoop distributed file management system for data storage and the MapReduce programming framework for data processing. Typically, applications are written in a language such as Pig, which maps queries across pieces of data that are stored across hundreds of computers in a parallel fashion and then combines the information from all to answer the query. SQL engines such as Dremel (Google), Hive (Hortonworks), and Spark (Databricks) allow very short response times. For postprocessing, however, such high-frequency data are often still stored in relational databases with greater functionality.

[3]The Web Appendix provides links to explanations of terms used in this and other sections.

C++, Fortran, and Java are powerful and fast low-level programming tools for analytics that come with large libraries of routines. Java programs are often embedded as applets within the code of web pages. R, used by Google, is a considerably slower but often-used open-source, higher-level programming language with functionality comparable to languages such as MATLAB. Perl is software that is suited for processing unstructured clickstream (HTML) data; it was initially used by Amazon but has been mostly supplanted by its rival Python (used by Dropbox), which is a more intuitive programming language that enables MapReduce implementation. Currently, academic research in marketing analytics already relies on many of these programming languages, and R seems to be the most popular. Much of this software for big data management and processing likely will become an integral part of the ecosystem of marketing academics and applied marketing analysts in the near future.

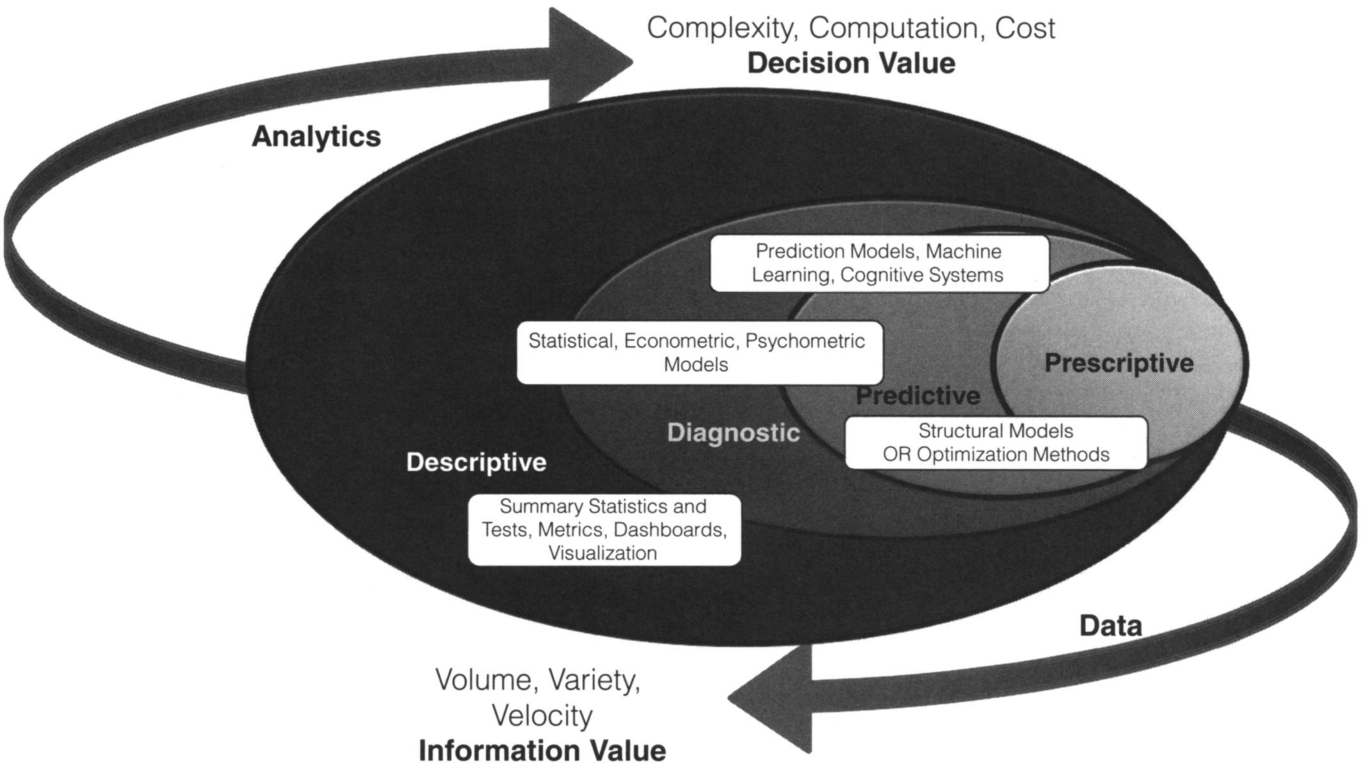
### *Volume, Variety, Velocity: Implications for Big Data Analytics*

The question is whether better business decisions require more data or better models. Some of the debate surrounding that question originates in research at Microsoft, in which Banko and Brill (2001) showed that in the context of text mining, algorithms of different complexity performed similarly, but adding data greatly improved performance. Indeed, throughout the academic marketing literature, complex models barely outperform simpler ones on data sets of small to moderate size. The answer to the question is rooted in the bias–variance trade-off. On the one hand, bias results from an incomplete representation of the true data-generating mechanism (DGM) by a model because of simplifying assumptions. A less complex model (one that contains fewer parameters) often has a higher bias, but a model needs to simplify reality to provide generalizable insights. To quote statistician George Box, "All models are wrong, but some are useful." A simple model may produce tractable closed-form solutions, but numerical and sampling methods allow for examination of more complex models at higher computational cost. Model averaging and ensemble methods such as bagging or boosting address the bias in simpler models by averaging many of them (Hastie, Tibshirani, and Friedman 2008). In marketing, researchers routinely use model-free evidence to provide confidence that more complex models accurately capture the DGM (see, e.g., Bronnenberg, Dubé, and Gentzkow 2012). Field experiments are increasingly popular because data quality (veracity) can substitute for model complexity: when the DGM is under the researchers' control, simpler models can be used to make causal inferences (Hui et al. 2013). Variance, on the other hand, results from random variation in the data due to sampling and measurement error. A larger volume of data reduces the variance. Complex models calibrated on smaller data sets often over-fit the data (i.e., they capture random error rather than the DGM). The notion that more data reduces error is well known to benefit machine learning methods such as neural networks, which are highly parameterized (Geman, Bienenstock, and Doursat 1992). However, not all data are created equal. A larger volume of data reduces variance, and even simpler models will fit better. Yet as data variety increases and data become richer, the underlying DGM expands. Much of the appeal of big data in marketing is that it provides traces of consumer behaviors (e.g., activities, interests, opinions, interactions) that were previously costly to observe even in small samples. To fully capture the information value of these data, more complex models are needed. Those models will support deeper insights and better decisions, while, at the same time, large volumes of data will support such richer representations of the DGM. However, these models come at greater computational costs.

Many current statistical and econometric models and the estimation methods used in the marketing literature are not designed to handle large volumes of data efficiently. Solutions to this problem involve data reduction, faster algorithms, model simplification, and/or computational solutions, which we discuss next. To fully support data-driven marketing decision making, the field of marketing analytics needs to encompass four levels of analysis: (1) descriptive data summarization and visualization for exploratory purposes, (2) diagnostic explanatory models that estimate relationships between variables and allow for hypothesis testing, (3) predictive models that enable forecasts of variables of interest and simulation of the effect of marketing control settings, and (4) prescriptive optimization models that are used to determine optimal levels of marketing control variables. Figure 4 shows that the feasibility of these higher levels of analysis decreases as a function of big data dimensions. It illustrates that the information value of the data grows as its volume, variety, and velocity increases but that the decision value derived from analytical methods increases at the expense of increased model complexity and computational cost.

In the realm of structured data, in which many of the advances in marketing analytics have been so far, all four levels of analysis are encountered. Many of the developments in marketing engineering (Lilien and Rangaswamy 2006) have been in this space as well, spanning a very wide range of areas of marketing (including pricing, advertising, promotions, sales force, sales management, competition, distribution, marketing mix, branding, segmentation and positioning, new product development, product portfolio, loyalty, acquisition, and retention). Explanatory and predictive models, such as linear and logistic regression and time-series models, have traditionally used standard econometric estimation methods such as generalized least squares, method of moments, and maximum likelihood. These optimization-based estimation methods become unwieldy for complex models with a large number of parameters. For complex models, simulation-based likelihood and Bayesian Markov chain Monte Carlo (MCMC) methods are used extensively. Markov chain Monte Carlo is a class of Bayesian estimation methods, the primary objective of which is to characterize the posterior distribution of model parameters. Such methods involve recursively drawing samples of subsets of parameters from their conditional posterior distributions (Gelman et al. 2003). This makes it possible to fit models that generate deep insight into the underlying phenomenon with the aim of generating predictions that generalize across categories, contexts, and markets. Optimization models have been deployed for sales

**FIGURE 4**
**Data and Analytic Approaches**




Notes: The figure shows the size and degree of structure in marketing data from right to left, and the extent to which analytical methods of increasing complexity are applied to that data from left to right.

force allocation, optimal pricing, conjoint analysis, optimal product/service design, optimal targeting, and marketing-mix applications.

There have been an increasing number of marketing analytics applications in the realm of unstructured data. Technological developments in processing unstructured data and the development of metrics from data summaries—such as provided by text-mining, eye-tracking, and pattern-recognition software—allow researchers to provide a data structure to facilitate the application of analytical methods. An example of the use of metrics as a gateway to predictive analytics includes the application by Netzer et al. (2012), who use text mining on user-generated content to develop competitive market structures. Once a data structure is put in place using metrics, researchers can build explanatory, prediction, and optimization models. Although the application of predictive and prescriptive approaches for unstructured data still lags, especially in practice, analyzing unstructured data in marketing seems to boil down to transforming them into structured data using appropriate metrics.

Large-volume structured data comprises four main dimensions: variables, attributes, subjects, and time (Naik et al. 2008). The cost of modeling structured data for which one or more of these dimensions is large can be reduced in one of two ways. First, one or more of the dimensions of the data can be reduced through aggregation, sampling, or selection; alternatively, situation-appropriate simplifications in model specifications can be used. Second, the speed and capacity of computational resources can be increased with approximations, more efficient algorithms, and high-performance computing. Techniques for reducing the dimensionality of data and speeding up computations are often deployed simultaneously, and we discuss these subsequently.

*Aggregation and compression.* Data volume can be reduced through aggregation of one or more of its dimensions, most frequently subjects, variables, or time. This can be done by simple averaging or summing—which, in several cases, yields sufficient statistics of model parameters that make processing of the complete data unnecessary—as well as through variable-reduction methods such as principal component analysis and related methods, which are common in data mining, speech recognition, and image processing. For example, Naik and Tsai (2004) propose a semiparametric single-factor model that combines sliced inverse regression and isotonic regression. It reduces dimensionality in the analysis of high-dimensional customer transaction databases and is scalable because it avoids iterative solutions of an objective function. Naik, Wedel, and Kamakura (2010) extend this to models with multiple factors and apply it to the analysis of large data on customer churn.

Aggregation of data on different samples of customers (e.g., mobile, social, streaming, geo-demographic) can be accomplished by merging aggregated data along spatial (e.g., designated market area, zip code) or time (e.g., week, month) dimensions or through data-fusion methods (Gilula, McCulloch, and Rossi 2006; Kamakura and Wedel 1997). Data requirements for specific applications can be reduced by fusing data at different levels of aggregation. For example, if store-level sales data are available from a retailer, these could be fused with in-home scanner panel data. This creates new variables that can increase data veracity because the store data has better market coverage but no competitor information, while the reverse is true for the home scanning data. Fusion may also be useful when applying structural models of demand that recover individual-level heterogeneity from aggregate data (store-level demand), in which case the fusion with individual-level data (scanner panel data) can help identify the heterogeneity distribution. Feit et al. (2013) use Bayesian fusion techniques to merge such aggregate data (on customer usage of media over time) with disaggregate data (customers' individual-level usage at each touch point) to make inferences about customer-level behavior patterns.

Bayesian approaches can be used in data compression. For example, in processing data collected over time, a Bayesian model can be estimated on an initial set of data for the first time period to determine the posterior distributions for the parameters. Then, the researcher only needs to retain these posteriors for future usage as priors for the parameters of the model calibrated on new data for subsequent time periods. Oravecz, Huentelman, and Vandekerckhove (2015) apply this method in the context of crowdsourcing. There are several refinements of this general approach. Ridgeway and Madigan (2002) propose first performing traditional MCMC on a subset of the data to obtain an initial estimate of the posterior distribution and then applying importance sampling/resampling to the initial estimates based on the complete data. This procedure can also be applied as new data come in over time. A related technique involves the use of information-reweighed priors, which obviates the need to run MCMC chains each time new data come in. Instead, the new data are used to reweight the existing samples from the posterior distribution of the parameters (Wang, Bradlow, and George 2014). This approach is related to the particle filter applied, for example, by Chung, Rust, and Wedel (2009) to reduce the computational burden in processing sequentially incoming data. All these sequential Bayesian updating techniques substantially reduce the computational burden of estimating complex models with MCMC on large-volume, high-velocity data because they reweigh (or redraw) the original samples of the parameters from their posterior distributions, often with closed-form weights that are proportional to the likelihood computed from the new data. This class of algorithms thus holds promise for big data because it avoids running MCMC chains on the full data or on new data that comes in. In addition, parallelizing these algorithms is much easier than with standard MCMC because they do not involve iterative computations.

*Sampling and selection.* Sampling is mostly applied to subjects, products, or attributes. In many cases, big data internal to the company is composed of the entire population of customers. Using samples of these data allows for classical sampling-based inference. Here, the researcher has full control over the size, nature, and completeness of the sample and can analyze multiple samples. Some of the dominant estimation approaches in marketing academia, in particular maximum likelihood, are developed within a statistical framework that purports to use a sample to make inferences on the population. Yet because, in many cases, big data captures an entire population, statistical inference becomes mute as asymptotic confidence regions degenerate to point masses under the weight of these massive data (Naik et al. 2008). Traditional statistical inference and hypothesis testing lose their appeal because the *p*-value, the probability of obtaining an effect in repeated samples that is at least as extreme as the effect in the data at hand, becomes meaningless in that case. Unless samples of the data are being analyzed, alternative methods are called for. A problem of using samples rather than the complete data, however, is that this approach may limit researchers' ability to handle long-tail distributions and extreme observations, and it is problematic when the modeling focus is on explaining or predicting rare events in the tail of high-dimensional data (see Naik and Tsai 2004). Furthermore, problems with sampling arise when inferences are made on social networks. In this case, a sampling frame may not be available, and simple random and other standard sampling methods may be inefficient or even detrimental to network properties (snowball or random forest samples perform better; see Ebbes, Huang, and Rangaswamy 2015). More importantly, sampling impedes personalization, for which data on each individual customer is needed, and thus eliminates a major point of leverage of big data.

Bayesian statistical inference offers philosophical advantages in big data applications because inference is conditioned on the data and considers parameters random. Inference reflects the researcher's subjective uncertainty about the model and its parameters rather than random variation due to sampling (Berger 1985). This enables the researcher to formulate a probabilistic statement about the underlying truth rather than about the data (e.g., "What is the probability that the null hypothesis is true?"). However, a limitation of many MCMC algorithms is that they are iterative in nature and, therefore, are computationally intense. Solutions to this computational problem (see the previous and following discussions) will render comprehensive statistical modeling of big data feasible, which may then be used to drive metrics on dashboards and displays. It is a promising avenue for further development to combine deep insight with user dashboards, as illustrated by Dew and Ansari (2015), who use semiparametric prediction of customer base dynamics on dashboards for computer games. These developments are important given the ubiquitous use of dashboards as the primary basis for decision making in industry, as is the case at Procter & Gamble, for example.

Selection can be used to reduce the dimensionality of big data in terms of variables, attributes, or subjects. Selection of subjects/customers can be used when interest focuses on specific well-defined subpopulations or segments. Even though big data may have a large number of variables ($p > n$ data), they may not all contribute to prediction. Bayesian additive regression tree approaches produce tree structures that may be used to select relevant variables. In the computationally intense Bayesian variable section approach, the key idea is to use a mixture prior, which enables the researcher to obtain a posterior distribution over all possible subset models. Alternatively, lasso-type methods can be used, which place a Laplace prior on coefficients (Genkin, Lewis, and Madigan 2007). Routines have been developed for the estimation of these approaches using parallel computing (Allenby et al. 2014).

*Approximations and simplifications.* A development employed for big data predictive analytics is the "divide-and-conquer" strategy. Several simpler models are fit to the data, and the results are combined. Examples of this strategy include estimation of logistic regression, or classification and regression trees on subsamples of the data, which then are tied together through bootstrapping, bagging, and boosting techniques (Varian 2014). To allow for statistical inference in the context of structured big data, researchers have used variations of this strategy to overcome the disadvantages of using a single random sample. Within a Bayesian framework, analyses of subsamples of big data with a single or multiple models have been combined using meta-analysis techniques (Bijmolt, Van Heerde, and Pieters 2005; Wang, Bradlow, and George 2014) or model-averaging methods (Chung, Rust, and Wedel 2009).

Another approach to reduce the computational burden of MCMC for big data analytics is to derive analytical approximations to complex posterior distributions in Bayesian models. Bradlow, Hardie, and Fader (2002) and Everson and Bradlow (2002) derive closed-form Bayesian inference for models with nonconjugate priors and likelihood, such as the negative binomial and beta-binomial models, using series expansions. A related technique that uses tractable deterministic approximations to the posterior distribution is variational inference (Braun and McAuliffe 2010). Here, the idea is to develop a (quadratic) approximation to the posterior distribution, the mode of which can be derived in closed form. Another method that promises to speed up the computations of MCMC is scalable rejection sampling (Braun and Damien 2015), which relies on tractable stochastic approximations to the posterior distribution (rather than deterministic approximations, as in variational inference). Taken together, these developments make MCMC estimation of hierarchical models on big data increasingly feasible.

An alternate way to achieve tractability is to simplify the models themselves: the researcher can use simple probability models without predictor variables that allow for closed-form solutions and fast computation. Fader and Hardie's (2009) study is an example in the realm of CRM to assess lifetime value. More work is needed to support the application of model-free methods (Goldgar 2001; Hastie, Tibshirani, and Friedman 2008; Wilson et al. 2010). Model-free methods can reduce computational effort so that big data can be analyzed in real time, but predictive validation is critical—for example, though cross-validation or bagging (Hastie, Tibshirani, and Friedman 2008). In the case of unstructured data, the issue is more complex. Deep neural networks (Hinton 2007) provide good prediction results for voice recognition, natural language processing, visual recognition and classification (especially objects and scenes in images and video), and computer game playing. These neural networks have many hidden layers that can be trained through stochastic gradient descent methods, and they provide viable approaches to the analysis of unstructured data with much predictive power (Nguyen, Yosinski, and Clune 2015). Both Facebook and Google have recently invested in the development and application of such approaches. Marketing models for large-scale unstructured data are still in their infancy, but research is starting to emerge (Lee and Bradlow 2011; Netzer et al. 2012). In this work, the computation of metrics from text, image, and video data using image-processing methods facilitates the application of standard models for structured data. Examples are Pieters, Wedel, and Batra (2010), who use file size of JPEG images as a measure of feature complexity of advertisement images; Landwehr, Labroo, and Herrmann (2011), who apply image morphing to selected design points to compute visual similarity of car images; and Xiao and Ding (2014), who deploy eigenface methods to classify facial features of models in ads.

Relatively little work in the academic marketing literature has addressed deep neural networks and other machine learning methods. This may be because marketing academics favor methods that represent the underlying DGM and support the determination of marketing control variables and thus may shy away from "one solution fits all" models and estimation methods, the identification and convergence properties of which cannot be unequivocally established. Nevertheless, future gains can be made if some of these methods can be integrated with the more theory-driven approaches in marketing. This is a fruitful area for further research.

*Computation.* Many of the statistical and econometric models used in marketing are currently not scalable to big data. MapReduce algorithms (which are at the core of Hadoop) provide a solution and allow for the processing of very large data in a massively parallel way by bringing computation locally to pieces of the data distributed across multiple cores rather than copying the data in its entirety for input into analysis software. For example, MapReduce-based clustering, naive Bayes classification, singular value decomposition, collaborative filtering, logistic regression, and neural networks have been developed. This framework was initially used by Google and has been implemented for multicore desktop grids and mobile computing environments.

Likelihood maximization is well suited for MapReduce because the log-likelihood consists of a sum across individual log-likelihood terms that can easily be distributed and allow for Map() and Reduce() operations. In this context, stochastic gradient descent (SGD) methods are often used to optimize the log-likelihood. Rather than evaluating the gradient of all terms in the sum, SGD samples a subset of these terms at every step and evaluates their gradient, which greatly economizes computations.

Parallelization of MCMC is also an active area of research, and several promising breakthroughs have been made recently (Brockwell and Kadane 2005; Neiswanger, Wang, and Xing 2014; Scott et al. 2013; Tibbitts, Haran, and Liechty 2011). Research also seems to be underway to combine features of SGD and MCMC. With the continued growth of multicore computing, formerly computationally prohibitive MCMC algorithms have now become feasible, as illustrated by their large-scale implementation by the analytics company In4mation Insights. Recent advances in parallelization using graphical processing units that promise to speed up likelihood maximization and MCMC sampling (Suchard et al. 2010) are equally promising but outside of the scope of the present exposition.

### *Synthesis*

Currently, only a few academic marketing applications take advantage of extremely large-scale data, especially rich unstructured data, and tackle the computational challenges that come with it. Marketing applications favor comprehensive statistical and econometric models that capture the DGM in detail but are often computationally (too) burdensome for big data (Table 1). Solutions to big data analytics in the future will use the following:

1. Developments in high-performance computing, including MapReduce frameworks for parallel processing, grid and cloud computing, and computing on graphic cards;
2. Simpler descriptive modeling approaches, such as probability models, or computer science and machine learning approaches that facilitate closed-form computations, possibly in combination with model averaging and other divide-and-conquer strategies to reduce bias;
3. Speed improvements in algorithms provided by variational inference, scalable rejection sampling, resampling and reweighting, sequential MCMC, and parallelization of likelihood and MCMC algorithms; and
4. Application of aggregation, data fusion, selection, and sampling methods that reduce the dimensionality of data.

Research in practice has often deployed a combination of components 1 and 2, focusing on exploration and description and generating actionable insights from unstructured data in real time; such work can be called "small stats on big data." The majority of academic research currently focuses on components 3 and 4: rigorous and comprehensive process models that allow for statistical inference on underlying causal behavioral mechanisms and optimal decision making, mostly calibrated on small to moderately sized structured data; such work can be called "big stats on small data."

Future solutions will likely have an "all-of-the-above" nature (Table 2). One-size-fits-all approaches may not be as effective, and techniques will need to be mixed and matched to fit the specific properties of the problem in question. Therefore, software for big data management and processing and high-performance computing likely will become an integral part of the ecosystem of marketing analysts in the near future.

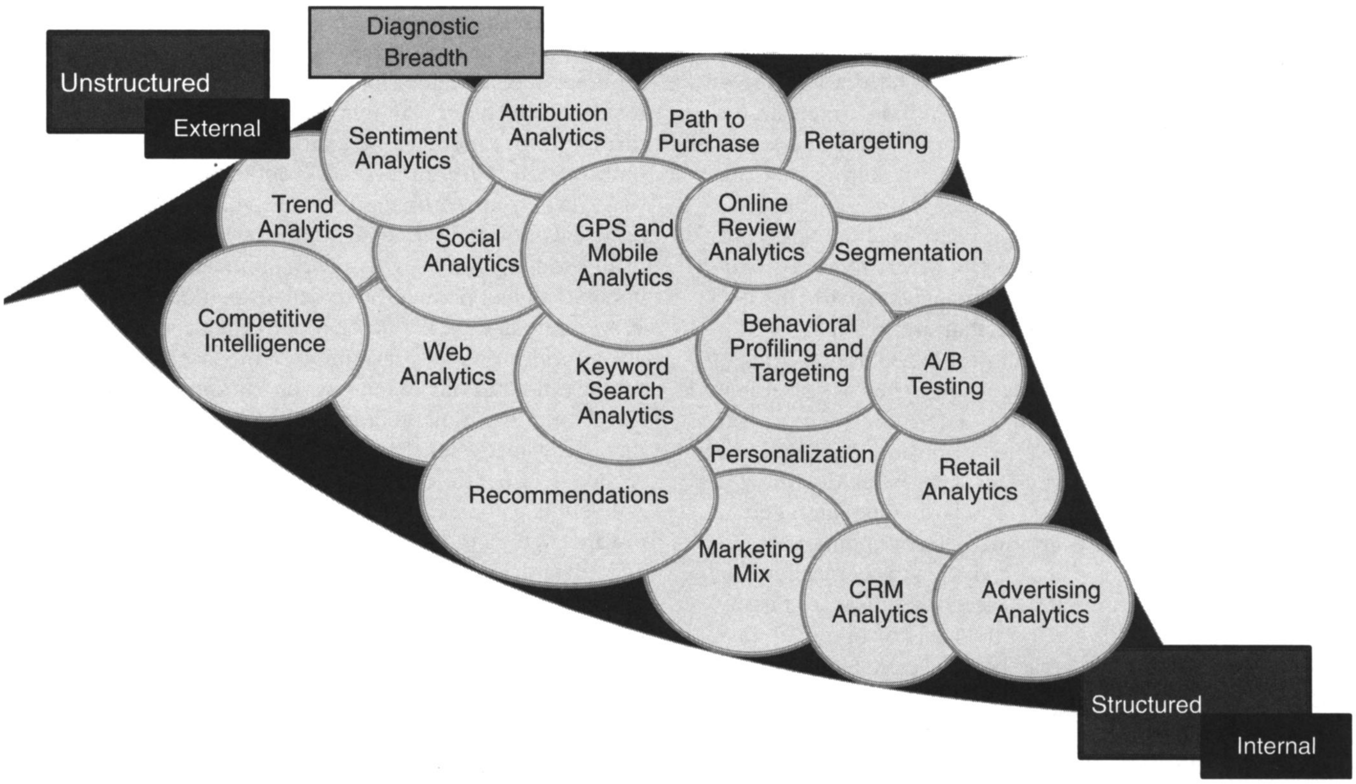
## Analytics and Models

Rich internal and/or external data enable marketing analytics to create value for companies and help them achieve their short-term and long-term objectives. We define marketing analytics as the methods for measuring, analyzing, predicting, and managing marketing performance with the purpose of maximizing effectiveness and return on investment (ROI). Figure 5 shows how big data marketing analytics creates increasing diagnostic breadth, which is often particularly beneficial for supporting firms' long-term objectives.

The following examples of recent research (illustrated in Figure 5) take advantage of new digital data sources to develop tailored analytical approaches that yield novel insights. The analysis of online reviews may help a firm fine-tune its offerings and provide better value to its customers. Chevalier and Mayzlin (2006) demonstrate this for online (book) reviews, which were shown to positively affect book sales. Keyword search analytics may help firms assess profitability of the design of their websites and placement of their ads. For example, Yao and Mela (2011) develop a dynamic structural model to explore the interaction of consumers and advertisers in keyword search. They find that when consumers click more frequently, the position of the sponsored advertising link has a larger effect. Furthermore, the study shows that search tools (e.g., sorting/filtering on the basis of price and ratings) may lead to increased platform revenue and consumer welfare. Analytics for mobile retail data may help a firm provide better recommendations, target promotions, personalize offerings, and increase spending by existing customers. Through field experiments with retail stores, Hui et al. (2013) find that mobile promotions motivate shoppers to travel further inside the store, which induces greater unplanned spending. Social analytics can help firms evaluate and monitor their brand equity and their competitive positions by identifying trending keywords. For example, Nam and Kannan (2014) propose measures based on social tagging data and show how they can be used to track customer-based brand equity and proactively improve brand performance. Competitive intelligence and trend forecasting can help firms identify changes in the environment and set up defenses to retain market share. Along these lines, Du and Kamakura (2012) show how to spot market trends with Google trends data using factor-analytic models. Clickstream data analytics allows for pattern matching between customer and noncustomer behavior to help firms identify segments for behavioral targeting. Trusov, Ma, and Jamal (2016) show how to combine a firm's data with third-party data to improve the recovery of customer profiles. Mobile GPS data analytics provides opportunities to geo-target customers with promotional offers based on situational contexts. Mobile data enable firms to test the efficacy of their targeting of both customers and noncustomers to increase revenues. Using field experiments, Andrews et al. (2015) show that commuters in crowded subway trains are twice as likely to respond to a mobile offer as commuters in noncrowded subway trains.

These illustrative examples make it easy to understand the importance of big data analytics for supporting marketing decision making in a wide range of areas. The marketing engineering approach, championed by Lilien and Rangaswamy (2006), has contributed to widespread recognition that if the problem drives the choice of models, the superior effectiveness of these models, the quality of the insights they

**FIGURE 5**
**The Diagnostic Breadth of Big Data Marketing Analytics**




Notes: The arrow shows the increasing breadth of diagnostic insights as a function of utilization of (mostly structured) internal data and (mostly unstructured) external data.

yield, and the consistency of decisions based on them, are all enhanced. After five decades of development, most marketing strategies and tactics now have their own well-specified data and analytical requirements. Academic marketing research has developed methods that specifically tackle issues in areas such as pricing, advertising, promotions, sales force, sales management, competition, distribution, branding, segmentation, positioning, new product development, product portfolio, loyalty, and acquisition and retention. Several marketing subfields have had extensive development of analytical methods, so that a cohesive set of models and decision making tools is available, including CRM analytics, web analytics, and advertising analytics. Next, we discuss analytics for three closely connected core domains in more detail: marketing-/media-mix optimization, personalization, and privacy and security.

### *Marketing Mix/Media Mix*

Models to measure the performance of the firm's marketing mix, forecast its effects, and optimize its elements date back to the 1960s. We noted some of these landmark developments in the "Analytics" subsection (for reviews, see Gatignon 1993; Hanssens 2014; Hanssens, Parsons, and Schultz 2001; Leeflang et al. 2000; Rao 2014). As new sources of data become available, there are increased opportunities for better and more detailed causal explanations as well as recommendations for optimal actions at higher levels of specificity and granularity. This was the case when scanner data became available (see Wittink et al. 2011), and new sources of digital data will lead to similar developments. For example, digital data on competitive intelligence and external trends can be used to understand the drivers of performance under the direct control of the firm and disentangle them from the external factors such as competition, environmental, economic, and demographic factors and overall market trends. Similarly, field experiments controlling for the impact of external factors allow online and offline retailers to calibrate the effects of price and promotions on demand for their products and improve forecasts of their impact (Muller 2014). Next, we focus on developments in marketing-mix modeling in the era of big data, which involve (1) including information and metrics obtained from new digital data sources to yield better explanations of the effects of marketing-mix elements; (2) attributing marketing-mix effects to new touch points, allocating market resources across classic and new media, and understanding and forecasting the simultaneous impact of marketing-mix elements on performance metrics; and (3) assessing causal effects of marketing control variables through structural representations of consumer behavior, IVs, and field experiments.

*Incorporating new data sources.* Research in marketing-mix allocation significantly benefits from two specific developments in data availability. The first is the increased availability of extensive customer-level data from within firm environments—through conducting direct surveys of customers, measuring attitudes or satisfaction, or recording customer behavior in physical stores and on websites and mobile apps. Hanssens et al. (2014) take advantage of one source of such data—consumer mindset metrics—to better model marketing actions' impact on sales performance. They find that combining marketing-mix and attitudinal metrics in VAR models improves both the prediction of sales and recommendations for marketing-mix allocation. The second development involves using data collected on customers and prospects outside the firm environment in addition to data that are available within the firm. This may alleviate the problem that activities of (potential) customers with competitors are unobservable in internal data and may help fully determine their path to purchase. For example, measures of online WOM (Godes and Mayzlin 2004), online reviews (Chevalier and Mayzlin 2006), or clickstreams (Moe 2003) can be included in marketing-mix models to provide better explanations and predictions of consumer choice and sales. Specifically, Moe (2003) uses clickstream data to categorize visits as buying, browsing, or searching visits on the basis of observed navigational patterns and shows that these different types of visits are associated with different purchase likelihoods. Although significant strides have been made, further research should focus on establishing which specific metrics work and which do not and how they can be best included in models of individual choice, aggregate sales, and market performance.

*Attribution and allocation to new touch points.* Data from new channels and devices are contributing to the development of new ways in which better marketing-mix decisions can be made. For example, while Prins and Verhoef (2007) examine the synergies between direct marketing and mass communications, Risselada, Verhoef, and Bijmolt (2014) take advantage of data from customers' social networks to understand the dynamic effects of direct marketing and social influence on the adoption of a high-technology product. Nitzan and Libai (2011) use data on more than a million customers' individual social networks to understand how network neighborhoods influence the hazard of defection from a service provider. Joo et al. (2014) focus on branded (as opposed to generic) keywords and find that television ads affect the number of related searches online. Similarly, Liaukonyte, Teixeira, and Wilbur (2015), using large-scale quasi-experimental data of television advertising and online shopping frequency at two-minute windows, find that television advertising influences online shopping and that the advertising content plays a key role. These studies highlight the role of cross-media effects in planning the marketing mix. In the context of new devices, Danaher et al. (2015) use panel data to examine the effectiveness of mobile coupon promotions. They find that location and time of delivery of coupons (relative to shopping time) influence redemption. Fong, Fang, and Luo (2015) examine the effectiveness of locational targeting of mobile promotions using a randomized field experiment and investigate targeting at the firm's own location (geo-fencing) versus a competitor's location (geo-conquesting). They find that competitive locational targeting produces increasing returns to the depth of promotional discounts.

The aforementioned research highlights convergence of different media (television, Internet, and mobile) and the resultant spillovers of marketing-mix actions delivered through those media. The availability of individual-level paths to purchase data—across multiple online channels (e.g., display ads, affiliates, referrals, search), across devices (e.g., desktop, tablet, smartphones), or across online and offline touch points—will create significant opportunities to understand and predict the impact of marketing actions at a very granular level. For one, these data have thrust the attribution problem—assigning credit to each touch point for the ultimate conversion—to the forefront. Li and Kannan (2014) propose a methodology to tackle that problem. Like marketing-mix allocation, attribution involves a marketing resource allocation problem. Yet even if the attribution problem is completely solved, it is only an intermediate step toward predicting its effects on the entire customer journey and toward obtaining an optimal allocation of the entire marketing mix. Many challenges can be expected. The modeling must accommodate spillovers across marketing actions and must reconcile more granular online and mobile data (e.g., derived from social networks) with more aggregate offline data and coordinate the different planning cycles for different advertising channels.

In addition, increased options for marketers to influence consumers—such as through firm-generated content in social media and content marketing, in which firms become content creators and publishers—have placed importance on the issue of understanding the individual effects of these options as part of the marketing mix. Newer methods and techniques are needed to accurately measure their impact. For example, Johnson, Lewis, and Nubbemeyer (2015) measure the effect of display ads using a new methodology that facilitates identification of the treatment effects of ads in a randomized experiment. They show it to be better than public service announcements and intent-to-treat A/B tests in minimizing the costs of tests. After such individual effects are measured, optimally allocating budgets across marketing/media-mix elements becomes possible.

Albers (2012) provides guidelines on how practical decision aids for optimal marketing mix allocation can be developed. He points to the need to study managers' behavior to better determine the specification of supply-side models. One of the important payoffs of working in a data-rich environment lies in the creation of decision aids to better budget and better allocate investments across the marketing mix, different products, market segments, and customers. Hanssens (2014) provides a review of optimization algorithms that span single-period and multiperiod approaches and are appropriate for monopolistic and competitive environments. Naik, Raman, and Winer (2005) explicitly model the strategic behavior of a firm that anticipates how competitors will likely make future decisions and reasons backward to deduce its own optimal decision in response. Although most extant work has focused on allocating the budget on single products, Fischer et al. (2011) propose a heuristic approach to solve the dynamic marketing budget allocation problem for multiproduct and multisegment (countries) firms. This approach, implemented at the multinational company Bayer, is an example of a modeling direction that solves pressing practical problems.

*Assessing causality of marketing-mix effects.* Assessing causality in marketing-mix models has received widespread attention in academia but unfortunately has not yet received as much attention in industry. If a marketing control variable is endogenously determined but not accounted for in the model (because of, e.g., missing variables, management actions dependent on sales outcomes), the DGM is not accurately captured. In that case, predictions of the effects of this marketing-mix element will be biased (Rossi 2014). This problem may be alleviated if exogenous IVs that are related to the endogenous control variable can be found. First, the variety in big data might help in finding better IVs, which is necessary because IVs are often problematic. In the case of television advertising, Shapiro (2014) exploits discontinuities in advertising spending in local designated market areas. Regression discontinuity designs that exploit variations in a possibly endogenous treatment variable on either side of a threshold are not economical in their data usage and may, therefore, benefit from large data (Hartmann, Nair, and Narayanan 2011). However, models with IVs do not generally predict better out of sample (Ebbes, Papies, and Van Heerde 2011). Researchers have developed several instrument-free methods to help in situations in which no valid instruments can be found (Ebbes et al. 2005; Park and Gupta 2012). These methods are suitable for automated application in large-scale data-production environments in industry, in which searching for valid instruments on a case-by-case basis is often infeasible. Second, digital data environments allow for field experiments that enable the researcher to assess the causal effects of marketing control variables (for research in this context, see Andrews et al. 2015; Hui et al. 2013). Third, in structural modeling of demand and supply, new types of data can help in calibrating the specifications of the models more precisely and efficiently. Chung, Steenburgh, and Sudhir (2014) provide an illustrative example, in which they estimate a dynamic structural model of sales force response to a bonus-based compensation plan. Rather than assuming the discount factors used by forward-looking sales people, as previous research has done, they estimate them from field data using a combination of exclusion restrictions and a model specific to the institutional setting.

Finally, taking consumers' forward-looking behavior into consideration is important in developing marketing-mix models that account for the idea that consumers may maximize their payoff over a finite or infinite horizon, rather than myopically. Although the identification of these models benefits from increased variation in data of large volume and variety, such models come with computational challenges that still need to be resolved. Liu, Montgomery, and Srinivasan (2015) tackle this problem by building a model of consumers' financial planning decisions based on the assumption that they are forward looking and discount future revenues. The researchers estimate their model with parallel MCMC, which enables them to accommodate individual-level heterogeneity and to design targeted marketing strategies. This work is one of the first applications of a structural model on relatively big data and is a promising development because it is important to account for forward-looking behavior in marketing-mix models, even those calibrated on field experiments.

### Personalization

Personalization takes marketing-mix allocation one step further in that it adapts the product or service offering and other elements of the marketing mix to the individual users' needs (Khan, Lewis, and Singh 2009). There are three main methods of personalization. (1) Pull personalization provides a personalized service when a customer explicitly requests it. An example is Dell, which enables customers to customize the computer they buy in terms of prespecified product features. (2) Passive personalization displays personalized information about products or services in response to related customer activities, but the consumer has to act on that information. For example, Catalina Marketing Services, an industry leader of personalized coupons delivered at the checkout counter of brick-and-mortar retail stores, personalizes coupons on the basis of shoppers' purchase history recorded on their loyalty cards. Recommendation systems represent another example of this approach. (3) Push personalization takes passive personalization one step further by sending a personalized product or service directly to customers without their explicit request. An example of this is Pandora, which creates online or mobile personalized radio stations. The radio stations are individually tailored on the basis of users' initial music selections and similarities between song attributes extracted from the Music Genome database.

For each of these types of personalization, there are three possible levels of granularity: (1) mass personalization, in which all consumers receive the same offering and/or marketing mix, personalized to their average taste; (2) segment-level personalization, in which groups of consumers with homogeneous preferences are identified and the marketing mix is personalized in the same way for all consumers in one segment; and (3) individual-level personalization, in which each consumer receives offerings and/or elements of the marketing mix customized to his or her individual tastes and behaviors. However, the availability of big data with extensive individual-level information does not necessarily make it desirable for companies to personalize at the most granular level. Big data offers firms the opportunity to choose an optimal level of granularity for different elements of the marketing mix, depending on the existence of economies of scale and ROI. For example, a firm such as Ford Motor Company develops a global (mass) brand image; personalizes product and brand advertising to segments of customers; customizes sales effort, prices, and promotions at the individual level; and personalizes in-car experiences using imaging technology.

*Recommendation systems.* Recommendation systems are powerful personalization tools, with best-in-class applications by Amazon and Netflix. There are two basic types of recommendation engines that are based on content filtering or collaborative filtering, but there are also hybrid recommendation systems that combine features of both types. Content filtering involves digital agents that make recommendations based on the similarity between a customer's past preferences for products and services. Collaborative filtering predicts a customer's preferences using those of similar customers. Model-based systems use statistical methods to predict these preferences; the marketing literature has predominantly focused on these (Ansari, Essegaier, and Kohli 2000). Research has demonstrated that model-based systems outperform simpler recommendation engines but do so at the cost of a larger computational burden. It has also shown that because many consumers are unwilling or unable to actively provide product ratings, much of the information in ratings-based recommendation systems is missing. Adequately dealing with this missing information in ubiquitous ratings-based recommendation systems can render recommendations much more effective (Ying, Feinberg, and Wedel 2006). In addition, most systems produce recommendations for consumers on the basis of their predicted preferences or choices but not necessarily on the basis of their predicted responsiveness to the recommendations themselves. Taking receptivity into account in making recommendations by utilizing ideas of response-based segmentation can greatly increase their effectiveness (Bodapati 2008).

Conceptually, personalization consists of (1) learning consumer preferences, (2) adapting offerings to consumers, and (3) evaluating the effectiveness of the personalization. Some of the problems with ratings-based recommendation systems have prompted companies (e.g., Amazon) to use data obtained unobtrusively from customers as input for online and mobile personalization of services. These three stages have long been used in closed-loop marketing (CLM) strategies. In digital environments, CLM can be fully automated in a continuous cycle, which gives rise to adaptive personalization systems.

*Adaptive personalization.* Adaptive personalization systems take personalization a step further by providing dynamically personalized services in real time (Steckel et al. 2005). For example, Groupon personalizes daily deals for products and services from local or national retailers and delivers them by e-mail or on mobile devices; as it collects more data on the individual subscriber, the deals are more accurately personalized. Another example is the buying and selling of online display ad impressions in real-time bidding auctions on ad-exchange platforms. These auctions are run fully automated in the time (less than one-tenth of a second) it takes for a website to load. The winning ad is instantly displayed on the publisher's site. To construct autonomous bidding rules, advertisers (1) track consumers' browsing behavior across websites, (2) selectively expose segments defined on the basis of those behaviors to their online display ads, and (3) record consumers' click-through behavior in response to their ads. This enables ad placement to be targeted across consumers, time, ad networks, and websites at a very high level of granularity. We have mentioned Pandora's adaptive personalization as another example in a previous section. Adaptive personalization thus takes marketing automation to the next stage. Rather than automating simple marketing decisions, it automates CLM's entire feedback loop. Automation offers the additional benefit of speeding up the personalization cycle dramatically. Adaptive personalization systems require minimal proactive user input and are mostly based on observed purchase, usage, or clickstream data. They learn consumer tastes adaptively over time by tracking consumers' changing behaviors. From a consumer's viewpoint, these systems are easy to use: the user only interacts with the service while usage data are recorded, and the service is adapted automatically. Online and mobile adaptive personalization systems implement fully automated CLM strategies by collecting and analyzing data, predicting user behavior, personalizing services, and evaluating the effectiveness of the recommendations in a continuous and automated cycle.

Zhang and Krishnamurthi (2004) were among the first to develop an adaptive personalization approach. They personalize online promotional price discounts by using an integrated purchase incidence, quantity, and timing model that forecasts consumers' response to promotional effort over time, and they employ numerical profit maximization to adaptively determine the timing and depth of personalized promotions. This application is conceptually similar to Catalina Marketing's services in offline stores. In an extension of this work, Zhang and Wedel (2009) investigate the profit implications of adaptive personalization online and offline, comparing three levels of granularity: mass, segment, and individual. The results show that individual-level personalization is profitable, but mostly in the online channel. Chung, Rust, and Wedel (2009) design and evaluate an adaptive personalization approach for mobile music. Their approach personalizes music using listening data as well as the music attributes that are used as predictor variables. They develop a scalable real-time particle-filtering algorithm (a dynamic MCMC algorithm) for personalization that runs on mobile devices. An element of surprise is incorporated through random recommendations, which prevent the system from homing in on a too-narrow set of user tastes. The model is unobtrusive to the users and requires no user input other than the user's listening behavior for songs that are automatically downloaded to the device. Field tests have shown that this system outperforms alternative algorithms similar to those of Pandora.

Hauser et al. (2009) develop a system for adaptive personalization of website design. This approach, which they call "website morphing," involves matching the content, look, and "feel" of the website to a fixed number of cognitive styles. First, the system estimates the probability of each cognitive style segment for website visitors on the basis of initialization data that involve the respondents' clickstreams and judgments of alternative web page morphs. In a second loop, the optimal morph assignment is computed using dynamic programming, maximizing both expected immediate profit and discounted future profit obtained when the user makes a purchase on the website. The system balances the trade-off between exploitation (i.e., presenting product options that best suit users' predicted preferences) and exploration (i.e., introducing surprise to help improve estimation). Morphing may substantially improve the expected profitability realized at the website. Researchers have applied similar ideas to the morphing of banner ads (Urban et al. 2014), which are automatically placed on websites and matched to consumers on the basis of their probabilities of segment membership to maximize click-through rates.

Adaptive personalization will grow with the advent of the Internet of Things and natural user interfaces, through which consumers interact with their digital devices through voice, gaze, facial expression, and motion control. As these data become available to marketers at massive scales, they will enable automated attention analysis, which will potentially benefit marketing-mix personalization in numerous ways.

### *Privacy and Security*

As more customer data are collected and personalization advances, privacy and security have become critical issues for big data analytics in marketing. According to a recent survey (Dupre 2015), more than three-quarters of consumers think that online advertisers have more information about them than they are comfortable with, and approximately half of them believe that websites ignore privacy laws. These perceptions are indicative of two realities. First, firms have been collecting data from multiple sources and fusing them to obtain better profiles of their customers. Easy availability of data from government sources (such as census, heath, employment, and telephone metadata, facilitated by the "Open Data Plan" released by the White House in 2013) and decreasing costs of storing and processing data have led to large ROI on such endeavors (Rust, Kannan, and Peng 2002). However, combining data sets has led to the "mosaic effect," yielding information on consumers that should be private but yet is revealed in the integrated data (e.g., people-search website Spokeo exploits such data). Second, privacy laws and security technology have not kept pace with data collection, storage, and processing technologies. This has resulted in an environment in which high-profile security breaches and misuse of private consumer information are prevalent. In the last ten years, more than 5,000 major data breaches have been reported, the majority in the financial industry. According to research by IBM and the Ponemon Institute, the average cost of a data breach approaches $4 million, approximately $150 per stolen record. Examples of recent high-profile data security breaches are those that hit Target, Sony Pictures Entertainment, Home Depot, and Ashley Madison. With cloud storage increasing, data breaches are predicted to become more common.

Two trends are likely to emerge that will change the status quo. First, governments will increasingly enact strict privacy laws to protect their citizens. This will limit how big data and analytics can be used for marketing purposes. The European Union, which already has stricter privacy laws, is considering expanding the so-called "right to be forgotten" to any company that collects personal individual customer data (Dwoskin 2015). Similar but less restrictive laws could soon be enacted in the United States. Goldfarb and Tucker (2010) show that privacy regulation that restricts the use of personal data may make online display ads less effective and imposes a cost especially on younger and smaller online firms that rely on ad revenues (Campbell, Goldfarb, and Tucker 2015). Second, firms are increasingly likely to police themselves. Currently, most companies communicate privacy policies to their customers. Respecting customers' privacy is good business practice and helps the firm build relationships with customers. Research by Tucker (2014) supports this notion. In a field experiment, she shows that when a website gave consumers more control over their personal information, the click-through rate on personalized ads doubled. In comparing the effects of opt-out, opt-in, and tracking ban policies on the display ad industry, Johnson, Lewis, and Nubbemeyer (2015) find that the opt-out policy has the least negative impact on publisher revenues and advertiser surplus. Increasingly, managers are expected to have a better understanding of new technologies and protocols to protect data security. In addition, marketing automation (as in, e.g., adaptive personalization) will prevent human intrusion and give customers greater confidence that their privacy is protected. Importantly, firms will need to ensure that sensitive customer information is distributed across separated systems, that data are anonymized, and that access to customers' private information is restricted within the organization. With security breaches becoming common, there is an emerging view that firms cannot completely render their systems breach safe. In addition to taking measures to protect data, firms should have data-breach response plans in place.

The implication of the aforementioned factors for marketing analytics is that there will be increased emphasis on data minimization and anonymization (see also Verhoef, Kooge, and Walk 2016). Data minimization requires marketers to limit the type and amount of data they collect and retain and dispose of the data they no longer need. Data can be rendered anonymous using procedures such as k-anonymization (each record is indistinguishable from at least $k - 1$ others), removing personally identifiable information, recoding, swapping or randomizing data, or irreversibly encrypting data fields to convert data into a nonhuman readable form. However, although these methods protect privacy, they may not act as a deterrent to data breaches (Miller and Tucker 2011).

As a result of data minimization, less individual-level data may become available for analytics development in academic and applied research, and increasingly more data will become available in aggregated form only. Research in marketing analytics should develop procedures to accommodate minimized and anonymized data without degrading diagnostic and predictive power, and analytical methods that preserve anonymity. For example, the Federal Trade Commission requires data providers such as Experian or Claritas to protect the privacy of individual consumers by aggregating individual-level data at the zip code level. Direct marketers rely on these data but traditionally ignore the anonymized nature of zip code–level information when developing their targeted marketing campaigns. Steenburgh, Ainslie, and Engebretson (2003) show how to take advantage of "massively categorical" zip code data through a hierarchical Bayesian model. The model enables the researcher to combine data from several sources at different levels of aggregation. Furthermore, if models used for predictive analytics are a priori known and have associated sufficient statistics or posterior distributions (e.g., means, variances, cross-products), those can be retained to allow for analysis without loss of information (rather than using the original data). Methods for analyzing aggregate data that accommodate inferences on unobserved consumer heterogeneity may provide solutions in some cases. Missing data imputation methods can be used to obtain consumer-level insights from aggregate data (e.g., Musalem, Bradlow, and Raju 2008) or for data in which fields, variables, or records have been suppressed. Imputation methods are also useful when only a portion of customers opt in to share their information because data augmentation can impute missing data from those customers who choose not to opt in. Further research in this area needs to focus on how customers' privacy can be protected in the use of rich marketing data while maximizing the utility that can be derived from it by developing models and algorithms that can preserve or ensure consumer privacy.

### Synthesis and Future Research Directions

Ongoing developments in the analytics of big data (see Table 1) involve (1) the inclusion of data obtained from external digital data sources with offline data to improve explanations and predictions of the effects of the marketing mix; (2) the attribution of marketing-mix effects via better understanding the simultaneous impact of marketing-mix elements while accommodating their different planning cycles; (3) the characterization of the entire path to purchase across offline and online channels and multiple devices and dynamic allocation of recourses to individual touch points within that path; (4) the assessment of causal effects of marketing control variables through structural representations of consumer behavior, instrumental variables, instrument-free methods, and field experiments; and (5) personalization of the marketing mix in fully automated closed-loop cycles. Future studies should build on these research directions and focus on the following topics and questions (Table 2).

*Big Marketing Data*

1. How can the fusion of data generated within the firm with data generated outside the firm take advantage of metadata on the context of customer interactions? How can this be done in a way that enables real-time analytics and real-time decisions?
2. What new methodologies and technologies will facilitate the integration of "small stats on big data" with "big stats on small data" approaches? What are key trade-offs that need to be made to estimate realistic models that are sufficient approximations?
3. How can field experiments be used to generate big (observational) data to obtain valid estimates of marketing effects quickly enough to enable operational efficiency without delaying marketing processes?
4. How can machine learning methods be combined with econometric methods to facilitate estimation of causal effects from big data at high speeds? What specific conditions determine where these new methods should be designed in terms of the continuum of machine learning to theory-based models?
5. What are viable data-analysis strategies and approaches for diagnostic, predictive, and prescriptive modeling of large-scale unstructured data?
6. How can deep learning and cognitive computing techniques be extended for analyzing and interpreting unstructured marketing data? How can creative elements of the marketing mix be incorporated in predictive and prescriptive techniques?

*Marketing Mix*

1. How can granular online, mobile data be aligned with more aggregate offline data to shed light on the path to purchase and facilitate behavioral targeting? How can metadata of contexts and unstructured data on creatives be incorporated in the analysis of path-to-purchase data?
2. How can ROI modeling more accurately identify and quantify the simultaneous financial impact of online and offline marketing activities?
3. What new techniques and methods can accurately measure the synergy, carryover, and spillover across media and devices using integrated path-to-purchase data?
4. How can attribution across media, channels, and devices account for strategic behavior of consumers and endogeneity in targeting?
5. How can planning cycles for different marketing instruments be incorporated in marketing-mix optimization models?

*Personalization*

1. What content should be personalized, at which level of granularity, and at what frequency? How can content be tailored to individual consumers using individual-level insights and automated campaign management?
2. How can firms derive individual-level insights from big data, using faster and less computationally expensive techniques to give readings of customers' intentions in real time?
3. How can firms personalize the mix of touch points (across channels, devices, and points in the purchase funnel) for customers in closed-loop cycles so that their experience is consistently excellent?
4. What role can cognitive systems, general artificial intelligence, and automated attention analysis systems play in delivering personalized customer experiences?

*Security and Privacy*

1. What techniques can be used to reduce the backlash to intrusion, as more personalization increases the chances that it may backfire?
2. What new methodologies need to be developed to give customers more control in personalizing their own experiences and to enhance the efficacy of data-minimization techniques?
3. How can data, software, and modeling solutions be developed to enhance data security and privacy while maximizing personalized marketing opportunities?

To provide more detailed examples of what such further research may entail, consider point 4 under "Big Marketing Data." In academic and applied research, unstructured data such as videos, texts, and images are used as input for predictive modeling by introducing structure through the derivation of numerical data—bag-of-words methods for textual information, tags and descriptors for images and videos, and so on. However, in addition to requiring context-specific dictionaries and supervised classification, none of these techniques quite captures the complete meaning contained in the unstructured data. For example, word counts in reviews or blogs ignore dependence between words and the syntax and logical sequence of sentences. Research has already used machine learning methods to detect specific languages and to provide meaningful summaries of text. They can thus be used to provide an interpretation of textual data. The Google cloud machine learning solutions for computer vision are able to interpret image content; classify images into categories; and detect text and objects, including faces, flowers, animals, houses, and logos, in images. In addition, the emotional expression of faces in the image can be classified. This means that an interpretation of the image based on meaningful relations between these objects is possible. These methods can be used to analyze and interpret, for example, earned social media content on platforms such as Facebook, Twitter, and Instagram. These interpretations of text and images in online news, product reviews, recommendations, shares, reposts, or social media mentions can be used to understand online conversations around products and services. They can be used to make predictions about their success, provide recommendations to consumers, and customize content of commercial communication on online platforms, including owned social media content, keywords, and targeted advertising (thus touching on points 1 and 2 under "Personalization"). To accomplish this, researchers need to understand how to include rich interpretations of images and text into predictive models. Exactly how the link between deep learning model output and marketing models can be forged, what the interpretations are that result, and whether they render marketing models more effective are topics for further research.

As a second example, take point 3 under "Marketing Mix." Integrated marketing-mix models need to accommodate expenditures on media vehicles within the classes of television, radio, print, outdoor, and owned and paid social media at granular, spatial, and temporal levels and measure the direct and indirect effects of WOM on earned social media, mindset metrics, and sales, accounting for endogeneity of marketing actions and computing ROI. This requires large-scale models with a time-series structure and multitudes of direct and indirect effects. Such models need to be comprehensive and allow for attribution and the quantification of carryovers and spillovers across these media classes and vehicles. They need to accommodate different levels of temporal and spatial granularity and levels of customer aggregation. Furthermore, current studies on attribution modeling have only scratched the surface of information available on customer touches of websites, display ads, search ads, and so on because they code each touch point on a single dimension. Each touch point can be described with multiple variables. For example, a website has an associated collection of metadata describing the design, content, and layout of the website, ad placements, and so on, all of which could affect customers' behavior when they visit the website. Extant marketing literature has tackled some of these issues, albeit in a piecemeal fashion. What is needed is a comprehensive approach to marketing mix and attribution modeling that integrates these various components of the marketing mix and addresses all these issues simultaneously. Data are no longer a limitation for doing so, although data from various sources will need to be combined. Close collaborations between academics and companies are likely needed to ensure availability of data and computational resources. New methods need to be developed, which might combine such techniques as VAR modeling, hierarchical Bayes and choice models, variable selection, and reduction and data fusion. Further research should investigate which data sources and models are suitable.

## Implementing Big Data and Analytics in Organizations

Organizations use analytics in their decision making in all functional areas—not only marketing and sales, but also supply chain management, finance, and human resources. This is exemplified by Wal-Mart, a pioneer in the use of big data analytics for operations, which relies heavily on analytics in human resources and marketing. Many companies aspire to integrate data-driven decisions across different functional areas. While managing big data analytics involves technology, organizational structure, and skilled and trained analysts, the primary precondition for its successful implementation in companies is a culture of evidence-based decision making. This culture is best summed up with a quote widely attributed to W. Edwards Deming: "In God we trust; all others must bring data." In such a culture, company executives acknowledge the need to organize big data analytics and give data/analytics managers responsibility and authority to utilize resources to store and maintain databases; develop and/or acquire software; and build and deploy descriptive, predictive, and normative models (Grossman and Siegel 2014). In those successful companies, big data analytics champions are typically found in the boardroom (e.g., chief financial officer, chief marketing officer), and analytics are used to drive business innovation rather than merely improve operations (Hagen and Khan 2014).

In an organization such as Netflix, in which analytics is fully centralized, initiatives are generated, prioritized, coordinated, and overseen in the boardroom. Despite Netflix's tremendous success, the highly specialized nature of marketing analytics, which varies dramatically across marketing domains, frequently demands a decentralized or hybrid infrastructure. This provides the flexibility needed for rapid experimentation and innovation and is more conducive to a nimble and effective deployment of analytics to a wide variety of marketing problems. A decentralized organization facilitates cross-functional collaboration and cocreation through communication of analysts and marketing managers in the company. It enables the analysts to identify relevant new data sources and opportunities for analytics, and—in conjunction with marketing managers—allows them to tailor their models and algorithms to the specific demands of marketing problems. However, this organizational structure requires deep distributed analytics capabilities and an emphasis on collaboration. AT&T is among the companies that follow this model by hosting data analytics within its business units.

One of the main challenges of decentralization is to achieve a critical mass of analysts that allows for continued development of broad and deep expertise across the organization and flexible and fast response to emerging issues, without excessive overhead or bureaucracy. Therefore, a hybrid organizational model is often effective. Here, a centralized unit is responsible for information technology and software as well as creating and maintaining databases. Marketing analysts can draw on the expertise of such a central unit when needed. Google takes this approach, whereby business units make their own decisions but collaborate with a central unit on selected initiatives. In some cases, especially for smaller companies and ones that are at the beginning of the learning curve, outsourcing of one or more of these centralized functions is a viable and cost-effective option.

Taking the role of the centralized unit one step further are organizations that form an independent big data center of excellence (CoE) within the company, overseen by a chief analytics officer. The marketing department and other units pursue initiatives overseen by the CoE. Amazon and LinkedIn are two firms that employ a CoE, and this model seems to be the one most widely adopted by big data companies. It provides synergies and economies of scale because it facilitates sharing of data and solutions across business units and supports and coordinates their initiatives. A problem of managing marketing budgets is the "silo" effect. Often, in large marketing organizations, investments in each of the marketing instruments (e.g., branding, search marketing, e-mail marketing) are managed by different teams with their own budgets. This can lead to each silo trying to optimize its own spending without taking a more global view. With more focus on integrated marketing communications, multichannel marketing, and influencing the entire path to purchase, the data analytics function best resides within a central unit or CoE, which prevents the silo effect by taking a more global view of marketing budgets with direct reporting to the chief marketing officer.

Even a decentralized or hybrid analytics infrastructure, however, does not preclude the need for data and analytics governance. Analytics governance functions, residing in centralized units, or CoEs, prioritize opportunities, obtain resources, ensure access to data and software, facilitate the deployment of models, develop necessary expertise, ensure accountability, and coordinate team effort. The teams in question include (1) marketing and management functions, which identify and prioritize opportunities and implement data-driven solutions and decisions; (2) analytics engineers, who determine data, software, and analytics needs; organize applications and processes; and document standards and best practices; (3) data science and data management functions, which ensure that data are accurate, up to date, complete, and consistent; and (4) legal and compliance functions, which oversee data security, privacy policies, and compliance. The chief analytics officer may promote the development of repeatable processes and solutions to gain efficiency and economies of scale across decentralized analytics teams.

To summarize, organizations that aim to extract value from big data analytics should have (1) a culture and leaders that recognize the importance of data, analytics, and data-driven decision making; (2) a governance structure that prevents silos and facilitates integrating data and analytics into the organization's overall strategy and processes in such a way that value is generated for the company; and (3) a critical mass of marketing analysts that collectively have sufficiently deep expertise in analytics as well as substantive marketing knowledge. Almost every company currently faces the challenge of hiring the right talent to accomplish this. An ample supply of marketing analysts with a cross-functional skill set; proficiency in technology, data science, and analytics; and up-to-date domain expertise is urgently needed, as are people with management skills and knowledge of business strategy to put together and lead those teams. We reflect on the implications for business education in the final section of this article.

## Conclusion: Implications for Education

This article has reviewed the history of data and analytics; highlighted recent developments in the key domains of marketing mix, personalization, and privacy, and security; and identified potential organizational barriers and opportunities toward successful implementation of analytics of rich marketing data in companies. Table 1 summarizes the state of the field, and Table 2 summarizes future research priorities. In this section, we round out our discourse with a discussion of the implications for the skill set required for analysts.

In the emerging big data environment, marketing analysts will be working increasingly at the interface of statistics/econometrics, computer science, and marketing. Their skill set will need to be both broad and deep. This poses obvious challenges that are compounded by the fact that various subdomains of marketing (e.g., advertising, promotions, product development, branding) have different data and analytics requirements, and one-size-fits-all analytical solutions are neither desirable nor likely to be effective. Analysts therefore need to have sufficiently deep knowledge of marketing modeling techniques for predicting marketing response, marketing-mix optimization, and personalization. They must be well-versed in the application of estimation techniques such as maximum likelihood methods, Bayesian MCMC techniques, and machine learning methods as well as familiar with optimization techniques from OR. Moreover, they need to possess soft skills and cutting-edge substantive knowledge in marketing to ensure that they can communicate to decision makers the capability and limitations of analytical models for specific marketing purposes. This will maximize the support for and impact of their decision recommendations. In many organizations, marketing analysts will fulfill the role of intermediaries between marketing managers and information technology personnel, or between marketing managers and outside suppliers of data and analytics capabilities, for which they need to have sufficient knowledge of both areas. Increasingly, routine marketing processes and decisions are becoming automated. This creates the challenge of determining how to ground these automated decisions in substantive knowledge as well as managerial intuition and oversight. Future marketers need to be well equipped to do that. Finally, the field will be in need of people with management skills and knowledge of business strategy as well as sufficient familiarity with technology and analytics to oversee and manage teams and business units. A recent study by research firm Gartner revealed that business leaders believe that the difficulty of finding talent with these skills is the main barrier toward implementing big data analytics (Levy 2015).

These skill set requirements also present a challenge for educators; few people will be able to develop deep knowledge in all these areas early in their career. In organizations, these skill sets are most often cultivated through on-the-job training and collective team effort. Some students of analytics may specialize and develop deep expertise in substantive marketing, soft skills, and management, such that they can take up management positions and can oversee analysts, negotiate with outside suppliers of analytics, and help formulate problems and interpret and communicate results. At the other end of the spectrum are those who aspire to be marketing analytics engineers or data scientists and work to develop deep knowledge of the technical aspects of the field, including database management, programming, and statistical/econometric modeling. Each of those marketers will have a role to play in analytics teams in organizations. All those working in the field will need to continue updating their knowledge across a broad domain, through conferences and trainings, to stay abreast of the tidal wave of new developments. Companies need to systematically invest in training and education of current employees and hire new ones with an up-to-date skill set to fill specific niches in their teams. Wal-Mart, for example, organizes its own yearly analytics conference with hundreds of participants and uses crowdsourcing to attract new talent.

The training and education of marketing analysts to develop this broad and deep skill set poses a challenge to academia. In many cases, people directly from programs in mathematics, statistics, econometrics, or computer science may not become effective and successful marketing analysts. Instead, next to existing specializations in successful undergraduate and masters of business administration programs at many universities, recently created masters programs in marketing analytics at institutions such as the University of Maryland and University of Rochester focus on developing these multidisciplinary skill sets in students who already have a rigorous training in these basic disciplines. Similar programs are urgently needed and are being developed elsewhere to meet to the increasing demand for marketing analysts worldwide. In addition, our field may need to embrace the model of the mathematics and computer science disciplines to educate doctoral students uniquely for industry functions (Sudhir 2016).

Finally, we emphasize opportunities for cross-fertilization of talent from academia and industry—for example, practitioners can benefit from specialized classes developed by universities, and academics can spend time within companies to be exposed to current problems and data. Such opportunities are becoming increasingly common and will benefit the field significantly in the near future, as big data analytics will continue to challenge and inspire academics and practitioners alike.

## REFERENCES


Adigüzel, Feray and Michel Wedel (2008), "Split Questionnaire Design for Massive Surveys," *Journal of Marketing Research*, 45 (October), 608–17.

Ailawadi, Kusum L., Jie Zhang, Aradhna Krishna, and Michael Kruger (2010), "When Wal-Mart Enters: How Incumbent Retailers React and How this Affects Their Sales Outcomes," *Journal of Marketing Research*, 47 (August), 577–93.

Albers, Sönke (2012), "Optimizable and Implementable Aggregate Response Modeling for Marketing Decision Support," *International Journal of Research in Marketing*, 29 (2), 111–22.

Allenby, Greg, Eric Bradlow, Edward George, Joh Liechty, and Robert McCulloch (2014), "Perspectives on Bayesian Methods and Big Data," *Customer Needs and Solution*, 1 (3), 169–75.

Andrews, Michelle, Xueming Luo, Zeng Fang, and Anindya Ghose (2015), "Mobile Ad Effectiveness: Hyper-Contextual Targeting with Crowdedness," *Marketing Science*, 35 (2), 218–33.

Andrews, Rick L., Andrew Ainslie, and Imran S. Currim (2002), "An Empirical Comparison of Logit Choice Models with Discrete Versus Continuous Representations of Heterogeneity," *Journal of Marketing Research*, 39 (November), 479–87.

Ansari, Asim, Skander Essegaier, and Rajeev Kohli (2000), "Internet Recommendation Systems," *Journal of Marketing Research*, 37 (August), 363–75.

——— and Carl F. Mela (2003), "E-Customization," *Journal of Marketing Research*, 40 (May), 131–45.

Banko, Michele and Eric Brill (2001). "Scaling to Very Very Large Corpora for Natural Language Disambiguation," in *Proceedings of 39th Annual Meeting of the Association for Computational Linguistics*. Toulouse, France: Association for Computational Linguistics, 26–33.

Bartels, Robert (1988), *The History of Marketing Thought*, 3rd ed. Columbus, OH: Publishing Horizons.

Bass, Frank (1969), "A New Product Growth for Model Consumer Durables," *Management Science*, 15 (5), 215–27.

Berger, James O. (1985), *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63 (4), 841–90.

Bijmolt, Tammo H.A., Harald J. van Heerde, Rik G.M. Pieters (2005), "New Empirical Generalizations on the Determinants of Price Elasticity," *Journal of Marketing Research*, 42 (May), 141–56.

Bodapati, Anand (2008), "Recommender Systems with Purchase Data," *Journal of Marketing Research*, 45 (February), 77–93.

Bradlow, Eric T., Bruce G.S. Hardie, and Peter S. Fader (2002), "Bayesian Inference for the Negative Binomial Distribution via Polynomial Expansions," *Journal of Computational and Graphical Statistics*, 11 (1), 189–201.

Braun, Michael and Paul Damien (2015), "Scalable Rejection Sampling for Bayesian Hierarchical Models," *Marketing Science*, 35 (3), 427–44.

——— and Jon McAuliffe (2010), "Variational Inference for Large-Scale Models of Discrete Choice," *Journal of the American Statistical Association*, 105 (489), 324–35.

Brockwell, Anthony E. and Joseph B. Kadane (2005), "Identification of Regeneration Times in MCMC Simulation, with Application to Adaptive Schemes," *Journal of Computational and Graphical Statistics*, 14 (2), 436–58.

Bronnenberg, Bart J., Jean-Pierre Dubé, and Matthew Gentzkow (2012), "The Evolution of Brand Preferences: Evidence from Consumer Migration," *American Economic Review*, 102 (6), 2472–508.

Bult, Jan-Roelf and Tom Wansbeek (1995), "Optimal Selection for Direct Mail," *Marketing Science*, 14 (4), 378–94.

Campbell, James, Avi Goldfarb, and Catherine Tucker (2015), "Privacy Regulation and Market Structure," *Journal of Economics & Management Strategy*, 24 (1), 47–73.

Casher, Jonathan D. (1969), *Marketing and the Computer*. Boston: D.H. Mark Publications.

Chevalier, Judith A. and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43 (August), 345–54.

Chintagunta, Pradeep (2002), "Investigating Category Pricing Behavior in a Retail Chain," *Journal of Marketing Research*, 39 (May), 141–54.

———, Tülin Erdem, Peter E. Rossi, and Michel Wedel (2006), "Structural Modeling in Marketing: A Review and Assessment," *Marketing Science*, 25 (6), 604–16.

Chung, Doug J., Thomas Steenburgh, and K. Sudhir (2014), "Do Bonuses Enhance Sales Productivity? A Dynamic Structural Analysis of Bonus-Based Compensation Plans," *Marketing Science*, 33 (2), 165–87.

Chung, Tuck-Siong, Roland T. Rust, and Michel Wedel (2009), "My Mobile Music: An Adaptive Personalization System for Digital Audio Players," *Marketing Science*, 28 (1), 52–68.

Coombs, Clyde (1950), "Psychological Scaling Without a Unit of Measurement," *Psychological Review*, 57, 148–58.

Danaher, Peter J., Michael S. Smith, Kulan A. Ranasinghe, and Tracey S. Danaher (2015), "Where, When, and How Long: Factors That Influence the Redemption of Mobile Phone Coupons," *Journal of Marketing Research*, 52 (October), 710–25.

Davenport, Thomas H. (2006), "Competing on Analytics," *Harvard Business Review*, (January), 99–107.

DeKimpe, Marnik G. and Dominique M. Hanssens (1995), "The Persistence of Marketing Effects on Sales," *Marketing Science*, 14 (1), 1–21.

DeSarbo, Wayne S. and Vithala Rao (1986), "A Constrained Unfolding Methodology for Product Positioning," *Marketing Science*, 5 (1), 1–19.

Dew, Ryan and Asim Ansari (2015), "A Bayesian Semiparametric Framework for Understanding and Predicting Customer Base Dynamics," working paper, Columbia University.

Dorfman, Robert and Peter O. Steiner (1954), "Optimal Advertising and Optimal Quality," *American Economic Review*, 44 (5), 826–36.

Du, Rex Yuxing and Wagner A. Kamakura (2012), "Quantitative Trendspotting," *Journal of Marketing Research*, 49 (August), 514–36.

Duncan, C.S. (1919), *Commercial Research*. New York: McMillan Co.

Dupre, Elyse (2015), "Privacy and Security Remain Tall Orders for Today's Marketers," *Direct Marketing News*, (January 23), [available at http://www.dmnews.com/infographics/privacy-and-security-remain-tall-orders-for-todays-marketers-infographic/article/393980/].

Dwoskin, Elizabeth (2015), "EU Seeks to Tighten Data Privacy Laws," *The Wall Street Journal*, (March 15), [available at http://blogs.wsj.com/digits/2015/03/10/eu-seeks-to-tighten-data-privacy-laws/].

Ebbes, Peter, Zan Huang, and Arvind Rangaswamy (2015), "Sampling Designs for Recovering Local and Global Characteristics of Social Networks," *International Journal of Research in Marketing*, (available electronically November 19), [DOI: 10.1016/j.ijresmar.2015.09.009].

———, Dominique Papies, and Harald Van Heerde (2011), "The Sense and Non-Sense of Holdout Sample Validation in the Presence of Endogeneity," *Marketing Science*, 30 (6), 1115–22.

———, Michel Wedel, Ulf Böckenholt, and Ton Steerneman (2005), "Solving and Testing for Regressor-Error (In)Dependence When No Instrumental Variables Are Available: With New Evidence for the Effect of Education on Income," *Quantitative Marketing and Economics*, 3 (4), 365–92.

Elrod, Terry (1988), "Choice Map: Inferring a Product-Market Map from Panel Data," *Marketing Science*, 7 (1), 21–40.

Erdem, Tülin and Michael P. Keane (1996), "Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15 (1), 1–21.

Everson, Philip J. and Eric T. Bradlow (2002), "Bayesian Inference for the Beta-Binomial Distribution via Polynomial Expansions," *Journal of Computational and Graphical Statistics*, 11 (1), 202–07.

Fader, Peter and Bruce Hardie (2009), "Probability Models for Customer-Base Analysis," *Journal of Interactive Marketing*, 23 (1), 61–69.

Feit, Eleanor, Pengyuan Wang, Eric T. Bradlow, and Peter S. Fader (2013), "Fusing Aggregate and Disaggregate Data with an Application to Multiplatform Media Consumption," *Journal of Marketing Research*, 50 (June), 348–64.

Ferber, Robert (1949), *Statistical Techniques in Market Research*. New York: McGraw-Hill.

Fischer, Marc, Sönke Albers, Nils Wagner, and Monika Frie (2011), "Practice Prize Winner—Dynamic Marketing Budget Allocation Across Countries, Products, and Marketing Activities," *Marketing Science*, 30 (4), 568–85.

Fong, Nathan M., Zheng Fang, and Xueming Luo (2015), "Geo-Conquesting: Competitive Locational Targeting of Mobile Promotions," *Journal of Marketing Research*, 52 (October), 726–35.

Gatignon, Hubert (1993), "Marketing Mix Models," in *Handbooks in Operations Research and Management Science*, Vol. 5: Marketing, J. Eliashherg and G.L. Lilien, eds. Amsterdam: North, 697–732.

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin (2003), *Bayesian Data Analysis*. London: Chapman and Hall.

Geman, Stuart, Elie Bienenstock, and René Doursat (1992), "Neural Networks and the Bias/Variance Dilemma," *Neural Computation*, 4 (1), 1–58.

Genkin, Alexander, David D. Lewis, and David Madigan (2007), "Large-Scale Bayesian Logistic Regression for Text Categorization," *Technometrics*, 49 (3), 291–304.

Gilula, Zvi, Robert E. McCulloch, and Peter E. Rossi (2006), "A Direct Approach to Data Fusion," *Journal of Marketing Research*, 43 (February), 73–83.

Godes, David and Dina Mayzlin (2004), "Using On-Line Conversations to Study Word-of-Mouth Communication," *Marketing Science*, 23 (4), 545–60.

Goldfarb, Avi and Catherine Tucker (2010), "Privacy Regulation and Online Advertising," *Management Science*, 57 (1), 57–71.

Goldgar, David E. (2001), "17 Major Strengths and Weaknesses of Model-Free Methods," *Advances in Genetics*, 42, 241–51.

Green, Paul E. (1963), "Bayesian Decision Theory in Pricing Strategy," *Journal of Marketing*, 27 (January), 5–14.

——— (1969), "Multidimensional Scaling: An Introduction and Comparison of Non-Metric Unfolding Techniques," *Journal of Marketing Research*, 6 (August), 330–41.

——— and V. Srinivasan (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research*, 5 (2), 103–23.

Grossman, Robert L. and Kevin P. Siegel (2014), "Organizational Models for Big Data and Analytics," *Journal of Organization Design*, 3 (1), 20–25.

Guadagni, Peter M. and John D.C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2 (3), 203–38.

Gupta, Sunil (1988), "Impact of Sales Promotions on When, What, and How Much to Buy," *Journal of Marketing Research*, 25 (November), 342–55.

Hagen, Christian and Khalid Khan (2014), "Building an Analytics-Based Organization," research report, A.T. Kearney, (accessed January 1, 2016), [available at http://www.africa.atkearney.com/documents/10192/6978675/Building+an+Analytics-Based+Organization.pdf/67d6181d-ab7e-40ce-ae4a-7bd92991bf76].

Hanssens, Dominique M. (2014), "Econometric Models," in *The History of Marketing Science*, Russell S. Winer and Scott A. Neslin, eds. Singapore: World Scientific Publishing,99–128.

———, Leonard J. Parsons, and Randall L. Schultz (2001), *Market Response Models: Economic and Time Series Analysis*. Boston: Kluwer Academic Publishers.

———, Koen H. Pauwels, Shuba Srinivasan, Marc Vanhuele, and Gokhan Yildirim (2014), "Consumer Attitude Metrics for Guiding Marketing Mix Decisions," *Marketing Science*, 33 (4), 534–50.

Hartmann, Wesley, Harikesh S. Nair, and Sridhar Narayanan (2011), "Identifying Causal Marketing Mix Effects Using a Regression Discontinuity Design," *Marketing Science*, 30 (6), 1079–97.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (2008), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.

Hauser, John R., Glen L. Urban, Guilherme Liberali, and Michael Braun (2009), "Website Morphing," *Marketing Science*, 28 (2), 202–23.

Hinton, Geoffrey (2007), "Learning Multiple Layers of Representation," *Trends in Cognitive Sciences*, 11 (10), 428–34.

Hui, Sam K., J. Jeffrey Inman, Yanliu Huang, and Jacob Suher (2013), "The Effect of In-Store Travel Distance on Unplanned Spending: Applications to Mobile Promotion Strategies," *Journal of Marketing*, 77 (March), 1–16.

Johnson, Garrett A., Randall A. Lewis, and Elmar I. Nubbemeyer (2015), "Ghost Ads: Improving the Economics of Measuring Ad Effectiveness," Working Paper No. FR 15-21, Simon Business School, University of Rochester [available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2620078].

Joo, Mingyu, Kenneth C. Wilbur, Bo Cowgill, and Yi Zhu (2014), "Television Advertising and Online Search," *Management Science*, 60 (1), 56–73.

Kamakura, Wagner A. and Gary J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26 (November), 379–90.

——— and Michel Wedel (1995), "Life-Style Segmentation with Tailored Interviewing," *Journal of Marketing Research*, 32 (August), 308–17.

——— and ——— (1997), "Statistical Data Fusion for Cross-Tabulation," *Journal of Marketing Research*, 34 (November), 485–98.

Kannan, P.K. and Gordon P. Wright (1991), "Modeling and Testing Structured Markets: A Nested Logit Approach," *Marketing Science*, 10 (1), 58–82.

Khan, Romana, Michael Lewis, and Vishal Singh (2009), "Dynamic Customer Management and the Value of One-to-One Marketing," *Marketing Science*, 28 (6), 1063–79.

Landwehr, Jan R., Aparna A. Labroo, and Andreas Herrmann (2011), "Gut Liking for the Ordinary: Incorporating Design Fluency Improves Automobile Sales Forecasts," *Marketing Science*, 30 (3), 416–29.

Lee, T.Y. and Eric T. Bradlow (2011), "Automatic Construction of Conjoint Attributes and Levels from Online Customer Reviews," *Journal of Marketing Research*, 48 (October), 881–94.

Leeflang, Peter S.H., Dick R. Wittink, Michel Wedel, and Alan Bultez (2000), *Building Models for Marketing Decisions*. Boston: Kluwer Academic Publishers.

Levy, Heather Pemberton (2015), "Marketers Embrace Analytics and Look for Talent," Gartner.com, (September 9), (accessed July 19, 2016), [available at http://www.gartner.com/smarterwithgartner/marketers-embrace-analytics-and-look-for-talent].

Li, Hongshuang (Alice) and P.K. Kannan (2014), "Attributing Conversions in a Multichannel Online Marketing Environment: An Empirical Model and a Field Experiment," *Journal of Marketing Research*, 51 (February), 40–56.

Liaukonyte, Jura, Thales Teixeira, and Kenneth C. Wilbur (2015), "Television Advertising and Online Shopping," *Marketing Science*, 34 (3), 311–30.

Lilien, Gary L. and Arvind Rangaswamy (2006), *Marketing Engineering: Computer-Assisted Analysis and Planning*, CreateSpace Independent Publishing Platform.

Little, John D.C. (1970), "Models and Managers: The Concept of a Decision Calculus," *Management Science*, 16 (8), B-466–485.

——— and Len M. Lodish (1969), "A Media Planning Calculus," *Operations Research*, 17 (1), 1–35.

Liu, Xiao, Alan Montgomery, and Kannan Srinivasan (2015), "Overhaul Overdraft Fees: Creating Pricing and Product Design Strategies with Big Data," working paper, Carnegie Mellon University.

Lodish, Leonard M. (1971), "CALLPLAN: An Interactive Salesman's Call Planning System," *Management Science*, 18, P-25–40.

Louviére, Jordan J. and George Woodworth (1983), "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data," *Journal of Marketing Research*, 20 (November), 350–67.

Luce, R. Duncan and John W. Tukey (1964), "Simultaneous Conjoint Measurement: A New Scale Type of Fundamental Measurement," *Journal of Mathematical Psychology*, 1 (1), 1–27.

Mantrala, Murali K., Prabhakant Sinha, and Andris A. Zoltners (1994), "Structuring a Multiproduct Sales Quota-Bonus Plan for a Heterogeneous Sales Force: A Practical Model-Based Approach," *Marketing Science*, 13 (2), 121–44.

Massy, William F., David B. Montgomery, and Donald G. Morrison (1970), *Stochastic Models of Buying Behavior*. Cambridge, MA: MIT Press.

McFadden, Daniel (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, P. Zarembka, ed. New York: Academic Press, 105–42.

Miller, Amalia and Catherine Tucker (2011), "Encryption and Data Security," *Journal of Policy Analysis and Management*, 30 (3), 534–56.

Mittal, Vikas, Pankaj Kumar, and Michael Tsiros (1999), "Attribute-Level Performance, Satisfaction, and Behavioral Intentions over Time: A Consumption-System Approach," *Journal of Marketing*, 63 (April), 88–101.

Moe, Wendy W. (2003), "Buying, Searching, or Browsing: Differentiating Between Online Shoppers Using In-Store Navigational Clickstream," *Journal of Consumer Psychology*, 13 (1/2), 29–40.

——— and Michael Trusov (2011), "The Value of Social Dynamics in Online Product Ratings Forums," *Journal of Marketing Research*, 48 (June), 444–56.

Montgomery, Alan L., Shibo Li, Kannan Srinivasan, and John C. Liechty (2004), "Modeling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science*, 23 (4), 579–95.

Muller, Frans (2014), "Big Data: Impact and Applications in Grocery Retail," presentation at the 2014 Marketing and Innovation Symposium, Erasmus University, (May 27–28), Rotterdam, Netherlands.

Musalem, Andrés, Eric T. Bradlow, and Jagmohan S. Raju (2008), "Who's Got the Coupon? Estimating Consumer Preferences and Coupon Usage from Aggregate Information," *Journal of Marketing Research*, 45 (December), 715–30.

Naik, Prasad A., Kalyan Raman, and Russell S. Winer (2005), "Planning Marketing-Mix Strategies in the Presence of Interaction Effects," *Marketing Science*, 24 (1), 25–34.

——— and Chi-Ling Tsai (2004), "Isotonic Single-Index Model for High-Dimensional Database Marketing," *Computational Statistics & Data Analysis*, 47 (4), 175–90.

———, Michel Wedel, Lynd Bacon, Anand Bodapati, Eric Bradlow, Wagner Kamakura, et al. (2008), "Challenges and Opportunities in High Dimensional Choice Data Analyses," *Marketing Letters*, 19 (3/4), 201–13.

———, ———, and Wagner Kamakura (2010), "Multi-Index Binary Response Model for Analysis of Large Data," *Journal of Business & Economic Statistics*, 28 (1), 67–81.

Nakanishi, Masao and Lee G. Cooper (1974), "Parameter Estimation for a Multiplicative Competitive Interaction Model: Least Squares Approach," *Journal of Marketing Research*, 11 (August), 303–11.

Nam, Hyoryung and P.K. Kannan (2014), "Informational Value of Social Tagging Networks," *Journal of Marketing*, 78 (July), 21–40.

Neiswanger, W., C. Wang, and E. Xing (2014), "Asymptotically Exact, Embarrassingly Parallel MCMC," in *Proceedings of the 30th International Conference on Conference on Uncertainty in Artificial Intelligence*, [available at http://www.cs.cmu.edu/~epxing/papers/2014/Neiswanger_Wang_Xing_UAI14a.pdf'.

Neslin, Scott A. (2014), "Customer Relationship Management," in *The History of Marketing Science*, Russell S. Winer and Scott A. Neslin, eds. Hackensack, NJ: World Scientific Publishing, 289–318.

Netzer, Oded, Ronen Feldman, Jacob Goldenberg, and Fresco Moshe (2012), "Mine Your Own Business: Market Structure Surveillance Through Text Mining," *Marketing Science*, 31 (3), 521–43.

Nguyen, A., J. Yosinski, and J. Clune (2015), "Deep Neural Networks Are Easily Fooled: High Confidence Predictions for Unrecognizable Images," in *Computer Vision and Pattern Recognition*. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Nitzan, Irit and Barak Libai (2011), "Social Effects on Customer Retention," *Journal of Marketing*, 75 (November), 24–38.

Nixon, H.K. (1924), "Attention and Interest in Advertising," *Archives de Psychologie*, 72 (1), 5–67.

Oravecz, Z., M. Huentelman, and J. Vandekerckhove (2015), "Sequential Bayesian Updating for Big Data," in *Big Data in Cognitive Science: From Methods to Insights*, M. Jones, ed. Sussex, UK: Psychology Press.

Park, Sungho and Sachin Gupta (2012), "Handling Endogenous Regressors by Joint Estimation Using Copulas," *Marketing Science*, 31 (4), 567–86.

Parsons, Leonard J. and Frank M. Bass (1971), "Optimal Advertising Expenditure Implications of a Simultaneous-Equation Regression Analysis," *Operations Research*, 19 (3), 822–31.

Pieters, Rik, Michel Wedel, and Rajeev Batra (2010), "The Stopping Power of Advertising: Measures and Effects of Visual Complexity," *Journal of Marketing*, 74 (September), 48–60.

Prins, Remco and Peter C. Verhoef (2007), "Marketing Communication Drivers of Adoption Timing of a New E-Service Among Existing Customers," *Journal of Marketing*, 71 (April), 169–83.

Raiffa, Howard and Robert Schlaifer (1961), *Applied Statistical Decision Theory*. Boston: Clinton Press.

Rao, Vithala R. (2014), "Conjoint Analysis," in *The History of Marketing Science*, Russell S. Winer and Scott A. Neslin, eds. Singapore: World Scientific Publishing, 47–76.

Reilly, W.J. (1929), *Marketing Investigations*. New York: Ronal Press Company.

Ridgeway, George and David Madigan (2002), "Bayesian Analysis of Massive Datasets via Particle Filters," in *Proceedings of KDD-02, The Eighth International Conference on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery, 5–13.

Risselada, Hans, Peter C. Verhoef, and Tammo H.A. Bijmolt (2014), "Dynamic Effects of Social Influence and Direct Marketing on the Adoption of High-Technology Products," *Journal of Marketing*, 78 (March), 52–68.

Roberts, John, Ujwal Kayande, and Stefan Stremersch (2014), "From Academic Research to Marketing Practice: Exploring the Marketing Science Value Chain," *International Journal of Research in Marketing*, 31 (2), 127–40.

Rossi, Peter E. (2014), "Even the Rich Can Make Themselves Poor: A Critical Examination of the Use of IV Methods in Marketing," *Marketing Science*, 33 (5), 655–72.

——— and Greg M. Allenby (2003), "Bayesian Statistics and Marketing," *Marketing Science*, 22 (3), 304–28.

———, Robert E. McCulloch, and Greg M. Allenby (1996), "The Value of Purchase History Data in Target Marketing," *Marketing Science*, 15 (4), 321–40.

Rust, John (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55 (5), 999–1033.

Rust, Roland T., P.K. Kannan, and Na Peng (2002), "The Customer Economics of Privacy in E-Service," *Journal of the Academy of Marketing Science*, 30 (4), 455–64.

Rutz, Oliver J., Michael Trusov, and Randolph E. Bucklin (2011), "Modeling Indirect Effects of Paid Search Advertising: Which Keywords Lead to More Future Visits?" *Marketing Science*, 30 (4), 646–65.

Schmittlein, David C. and Robert A. Peterson (1994), "Customer Base Analysis: An Industrial Purchase Process Application," *Marketing Science*, 13 (1), 41–67.

Scott, Steven L., Alexander W. Blocker, Fernando V. Bonassi, Hugh A. Chipman, Edward J. George, and Robert E. McCulloch (2013), "Bayes and Big Data: The Consensus Monte Carlo Algorithm," working paper, University of Chicago.

Shapiro, B.T. (2014), "Positive Spillovers and Free Riding in Advertising of Pharmaceuticals: The Case of Antidepressants," working paper, Booth School of Business, University of Chicago.

Shaw, Robert (1987), *Database Marketing*, Gower Publishing Co.

Starch, Daniel (1923), *Principles of Advertising*. Chicago: A.W. Shaw Company.

Steckel, Joel, Russell Winer, Randolph E. Bucklin, Benedict Dellaert, Xavier Drèze, Gerald Häubl, et al. (2005), "Choice in Interactive Environments," *Marketing Letters*, 16 (3/4), 309–20.

Steenburgh, Thomas, Andrew Ainslie, and Peder Hans Engebretson (2003), "Massively Categorical Variables: Revealing the Information in Zip Codes," *Marketing Science*, 22 (1), 40–57.

Stonborough, Thomas H.W. (1942), "Fixed Panels in Consumer Research," *Journal of Marketing*, 7 (April), 129–38.

Suchard, Marc A., Quanli Wang, Cliburn Chan, Jacob Frelinger, Andrew Cron, and Mike West (2010), "Understanding GPU programming for Statistical Computation: Studies in Massively Parallel Massive Mixtures," *Journal of Computational and Graphical Statistics*, 19 (2), 419–38.

Sudhir, K. (2016), "The Exploration-Exploitation Tradeoff and Efficiency in Knowledge Production," *Marketing Science*, 35 (1), 1–9.

Teixeira, Thales, Michel Wedel, and Rik Pieters (2010), "Moment-to-Moment Optimal Branding in TV Commercials: Preventing Avoidance by Pulsing," *Marketing Science*, 29 (5), 783–804.

Telpaz, Ariel, Ryan Webb, and Dino Levy (2015), "Using EEG to Predict Consumers' Future Choices," *Journal of Marketing Research*, 52 (August), 511–29.

Tibbitts, Matthew, Murali Haran, and John C. Liechty (2011), "Parallel Multivariate Slice Sampling," *Statistics and Computing*, 21 (3), 415–30.

Trusov, Michael, Liye Ma, and Zainab Jamal (2016), "Crumbs of the Cookie: User Profiling in Customer-Base Analysis and Behavioral Targeting," *Marketing Science*, 35 (3), 405–26.

Tucker, Catherine (2014), "Social Networks, Personalized Advertising, and Privacy Controls," *Journal of Marketing Research*, 51 (October), 546–62.

Urban, Glen L., Gui Liberali, Erin MacDonald, Robert Bordley, and John Hauser (2014), "Ad Morphing," *Marketing Science*, 33 (1), 27–46.

Varian, Hal R. (2014), "Big Data: New Tricks for Econometrics," *Journal of Economic Perspectives*, 28 (2), 3–27.

Verhoef, Peter C., Edwin Kooge, and Natasha Walk (2016), *Creating Value with Big Data Analytics: Making Smarter Marketing Decisions*. London: Routledge.

Wang, P., Eric Bradlow, and Ed George (2014), "Meta-Analyses Using Information Reweighting: An Application to Online Advertising," *Quantitative Marketing and Economics*, 12 (2), 209–33.

Wedel, Michel and Wayne S. DeSarbo (1995), "A Mixture Likelihood Approach for Generalized Linear Models," *Journal of Classification*, 12 (1), 21–55.

——— and Rik Pieters (2000), "Eye Fixations on Advertisements and Memory for Brands: A Model and Findings," *Marketing Science*, 19 (4), 297–312.

White, Percival (1931), *Market Research Technique*. New York: Harper & Brothers.

Wilson, Melanie A., Edwin S. Iversen, Merlise A. Clyde, Scott C. Schmidler, and Joellen M. Schildkraut (2010), "Bayesian Model Search and Multilevel Inference for SNP Association Studies," *Annals of Applied Statistics*, 4 (3), 1342–64.

Winer, Russell S. and Scott A. Neslin, eds. (2014), *The History of Marketing Science*. Hackensack, NJ: World Scientific Publishing.

Wittink, Dick R., M.J. Addona, William J. Hawkes, and John C. Porter (2011), "SCAN*PRO: The Estimation, Validation and Use of Promotional Effects Based on Scanner Data," in *Liber Amicorum in Honor of Peter S.H. Leeflang*, J.E. Wierenga, P.C. Verhoef, and J.C. Hoekstra, eds. Groningen, The Netherlands: Rijksuniversiteit Groningen, 135–62.

Xiao, Li and Min Ding (2014), "Just the Faces: Exploring the Effects of Facial Features in Print Advertising," *Marketing Science*, 33 (3), 338–52.

Yao, Song and Carl F. Mela (2011), "A Dynamic Model of Sponsored Search Advertising," *Marketing Science*, 30 (3), 447–68.

Ying, Yuan Ping, Fred Feinberg, and Michel Wedel (2006), "Leveraging Missing Ratings to Improve Online Recommendation Systems," *Journal of Marketing Research*, 43 (August), 355–65.

Zhang, Jie and Lakshman Krishnamurthi (2004), "Customizing Promotions in Online Stores," *Marketing Science*, 23 (4), 561–78.

——— and Michel Wedel (2009), "The Effectiveness of Customized Promotions in Online and Offline Stores," *Journal of Marketing Research*, 46 (April), 190–206.