Chapter 18
Retail Analytics

Ramandeep S. Randhawa

1 Introduction

1.1 Background

Retail is one of the largest sectors in today's economy. The global retail sector is estimated to have revenues of USD 28 trillion in 2019 (with approximately USD 5.5 trillion sales in the USA alone). This sector represents 31% of the world's GDP and employs billions of people throughout the globe.[1] A large and growing component of this is e-commerce or e-tail, which includes products and services ordered via the Internet, with sales estimated to be about USD 840 billion in 2014, and expected to grow at a rate of about 20% over the subsequent years.[2] Analytics is gaining increasing prominence in this sector with the retail analytics market size being estimated at over USD 3.52 billion in 2017 and is expected to grow at a CAGR of over 19.7% over the next few years.[3]

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-319-68837-4_18) contains supplementary material, which is available to authorized users.

[1] https://www.businesswire.com/news/home/20160630005551/en/Global-Retail-Industry-Worth-USD-28-Trillion (accessed on Jul 31, 2018).

[2] https://www.atkearney.com/consumer-goods/article?/a/global-retail-e-commerce-keeps-on-clicking (accessed on Mar 1, 2019).

[3] https://www.marketsandmarkets.com/Market-Reports/retail-analytics-market-123460609.html (accessed on Jul 31, 2018).

R. S. Randhawa (✉)
Marshall School of Business, University of Southern California, Los Angeles, CA, USA
e-mail: Ramandeep.Randhawa@marshall.usc.edu

B. Pochiraju, S. Seshadri (eds.), *Essentials of Business Analytics*, International Series in Operations Research & Management Science 264,
https://doi.org/10.1007/978-3-319-68837-4_18

Retail acts as the last stop in the supply chain by selling products directly to customers. Given that retailers are focused on this aspect, collecting data on customer behavior and preferences and incorporating these into business decisions are quite natural. And so, retail has indeed been an early adopter of analytics methodologies and focuses heavily on advancing knowledge in this domain.

1.2 What Is Retail Analytics?

Retail analytics is an umbrella term that comprises various elements which assist with decision-making in the retail business. Typically, this includes data collection and storage (data warehousing), data analysis that involves some statistical or predictive modeling, and decision-making. Traditionally, the analysis of data was limited to monitoring and visualizing some key performance indicators (KPIs) retrospectively.

One may use the term *business intelligence* to refer to the gamut of activities that underlie intelligent business decision-making. However, typically this term is used to refer to the collection and presentation of historical information in an easy-to-understand manner, via reports, dashboards, scorecards, etc. The term *advanced analytics* is typically reserved for when predictive modeling is applied to data via statistical methods or machine learning. Our focus in this chapter will be on the later, *advanced analytics*, methodologies that can significantly assist in the decision-making process in retail.

To understand the role analytics plays in retail, it is useful to break down the business decisions taken in retail into the following categories: consumer, product, workforce, and advertising.

1. *Consumer*: Personalization is a key consumer-level decision that retail firms make. Personalized pricing by offering discounts via coupons to select customers is one such decision. This approach uses data collection via loyalty cards to better understand a customer's purchase patterns and willingness to pay and uses that to offer personalized pricing. Such personalization can also be used as a customer retention strategy. Another example is to offer customers a personalized sales experience: in e-tail settings, this entails offering customers a unique browsing experience by modifying the products displayed and suggestions made based on the customer's historical information.
2. *Product*: Retail product decisions can be broken down into single product and group of product decisions. Single or individual product decisions are mostly inventory decisions: how much stock of the product to order, and when to place the order. At the group level, the decisions are typically related to pricing and assortment planning. That is, what price to set for each product in the group and how to place the products on the store-shelves, keeping in mind the variety of products, the number of each type of product, and location. To make these decisions, predictive modeling is called for to forecast the product

demand and the price-response function, and essentially the decision-maker needs to understand how customer reacts to price changes. A fine understanding of consumer choice is also needed to understand how a customer chooses to buy a certain product from a group of products.

3. *Human resources*: The key decisions here are related to the number of employees needed in the store at various times of the day and how to schedule them. To make these decisions, the overall work to be completed by the employees needs to be estimated. Part of this is a function of other decisions, such as the effort involved in stocking shelves, taking deliveries, changing prices, etc. There is additional work that comes in as a function of the customer volume in the store. This includes answering customer questions and manning checkout counters.
4. *Advertising*: In the advertising sphere, companies deal with the typical decisions of finding the best medium to advertise on (online mediums such as *Google Adwords*, *Facebook*, *Twitter*, and/or traditional mediums such as print and newspaper inserts) and the best products to advertise. This may entail cultivating some "loss-leaders" that are priced low to entice customers into the store, so they may also purchase other items which have a greater margin.

We refer the reader to the survey article by Bradlow et al. (2017). It reviews big data and predictive analytics practices in retailing. They discuss several statistical issues and methods, including Bayesian analysis, which are important in collecting, processing, modeling, and analysis of data. In addition, they emphasize ethical and privacy issues.

1.3 Examples of Retail Analytics in Action

- Analytics has revealed that a great number of customer visits to online stores fail to convert at the last minute, when the customer has the item in their shopping basket but does not go on to confirm the purchase. Theorizing that this was because customers often cannot find their credit or debit cards to confirm the details, Swedish e-commerce platform Klarna moved its clients (such as Vistaprint, Spotify, and 45,000 online stores) onto an invoicing model, where customers can pay after the product is delivered. Sophisticated fraud prevention analytics are used to make sure that the system cannot be manipulated by those with devious intent.
- Trend forecasting algorithms comb social media posts and Web browsing habits to elicit what products may be causing a buzz, and ad-buying data is analyzed to see what marketing departments will be pushing. Brands and marketers engage in "sentiment analysis," using sophisticated machine learning-based algorithms to determine the context when a product is discussed. This data can be used to accurately predict what the top selling products in a category are likely to be.

- Russian retailers have found that the demand for books increases exponentially as the weather gets colder. So retailers such as Ozon.ru increase the number of book recommendations which appear in their customers' feeds as the temperature drops in their local areas.[4]
- The US department store giant, Macy's, recently realized that attracting the right type of customers to its brick-and-mortar stores was essential. Due to its analytics showing up a dearth of the vital millennials demographic group, it recently opened its "One Below" basement[5] at its flagship New York store, offering "selfie walls" and while-you-wait customized 3D-printed smartphone cases. The idea is to attract young customers to the store who will hopefully go on to have enduring lifetime value to the business.
- Amazon has proposed using predictive shipping analytics[6] to ship products to customers before they even click "add to cart." According to a recent trend report by DHL, over the next 5 years, this so-called psychic supply chain will have far reaching effects in nearly all industries, from automotive to consumer goods. It uses big data and advanced predictive algorithms to enhance planning and decision-making.

1.4 Complications in Retail Analytics

There are various complications that arise in retail scenarios that need to be overcome for the successful use of retail analytics. These complications can be classified into (a) those that affect predictive modeling and (b) those that affect decision-making.

Some of the most common issues that affect predictive modeling are demand censoring and inventory inaccuracies (DeHoratius and Raman 2008). Typically, retail firms only have access to sales information, not demand information, and therefore need to account for the fact that when inventory runs out, actual demand is not observed. Ignoring this censoring of information can result in underestimating demand. There is also a nontrivial issue of inventory record inaccuracies that exists in retail stores—the actual number of products in an inventory differs from the number expected as per the firm's IT systems (DeHoratius 2011). Such inaccuracy may be caused by theft, software glitches, etc. This inaccuracy needs to be incorporated into demand estimation because it confounds whether demand is low or appears low due to product shortage. Inaccuracy also affects decision-

[4]https://www.forbes.com/sites/bernardmarr/2015/11/10/big-data-a-game-changer-in-the-retail-sector/#651838599f37 (accessed on Jul 31, 2018).

[5]https://www.bloomberg.com/news/articles/2015-09-25/this-is-macy-s-idea-of-a-millennial-wonderland (accessed on Jul 31, 2018).

[6]https://www.forbes.com/sites/onmarketing/2014/01/28/why-amazons-anticipatory-shipping-is-pure-genius/#178cd4114605 (accessed on Jul 31, 2018).

making by impacting the timing of order placement. Some of the other factors that affect decision-making are constraints on changing prices, physical constraints on assortments, supplier lead times, supplier contracts, and constraints on work-force scheduling. In particular, retail firms deal with many constraints on changing prices. Some of these are manpower constraints: changing assortments requires a reconfiguration of store shelves, and changing prices may involve physically tagging products with the new price (a labor-intensive process). To make it easier to change prices, many stores such as Kohl's are turning to electronic shelf labels as a means of making the price-changing process efficient.[7] There are additional nonphysical constraints that a firm may need to deal with. For instance, in fashion, prices are typically only marked down, and once a price is lowered, it is not increased. There are also limits to how often prices may be changed, for instance, twice a week. There are many supplier-based constraints that need to be considered as well, for instance, lead times on any new orders placed and any terms agreed to in supplier contracts.

In this chapter, our focus will be on use of retail analytics for product-based decision-making. We continue in Sect. 2, by exploring the various means of data collection that are in use by retailers and those that are gaining prominence in recent times. In Sect. 3, we will discuss some key methodologies that are used for such decision-making. In particular, we will discuss various statistics and machine learning methodologies for demand estimation, and how these may be used for pricing, and techniques for modeling consumer choice for assortment optimization. In Sect. 4, we will focus on the many business challenges and opportunities in retail, focusing on both e-tail settings and the growth in retail analytic startups.

2 Data Collection

Retail data can be considered as both structured (spreadsheet with rows and columns) and unstructured (images, videos, and other location-based data). Traditional retail data has been structured and derived mostly from point-of-sale (POS) devices and data supplied by third parties. POS data typically captures sales information, number of items sold, prices, and timestamps of transactions. Combined with inventory record keeping, this data provides a rich trove of information about products sold and, in particular, product baskets (collection of items in the cart) sold. Retailers tend to use loyalty programs to attach customer information to this information, so that customer level sales data can be analyzed. Third-party data typically consists of competitor information, such as prices and product assortments. It also consists of some broad information about the firm's customers, such as their demographics and location.

[7]https://www.wsj.com/articles/now-prices-can-change-from-minute-to-minute-1450057990 (accessed on Jul 31, 2018).

The recent trend is to capture more and more unstructured data. There now exists technology that can help retailers collect information not only about direct customer sales but also about product comparisons, that is, what products were compared by the customer in making decisions. Video cameras coupled with image detection technology can help collect data on customer routes through a store. This video data can also be used to collect employee data (e.g., what tasks are employees doing, how are customers being engaged, and how much time does a customer needing assistance have to wait for the assistance to be provided). Recently, many firms have also employed eye-tracking technology in controlled environments to collect data on how the store appears from a customer's perspective; a major downside of this technology is that it requires the customer to wear specialized eyeglasses.

With the advent of *Internet of Things* (IoT), the potential to collect in-store data has increased. Walmart began using radio-frequency identification (RFID) technology about a decade ago. Initially, the main goal of using this technology was to track inventory in the supply chain. However, increasingly, retailers are finding it beneficial to track in-store inventory. RFID tags are far easier to read than barcodes because they do not require direct line-of-sight scanning. This ease of tracking allows the tags to be used to collect data on the movement of products through the store. For instance, in fashion retail, the retailer can track the items that make their way to the fitting rooms; the combination of items tried can also be tracked, and finally it can easily be detected whether the items were chosen or not. All of this provides a rich set of data to feed into the system for analytics.

Near-field communication (NFC) chips are also being used by retailers to simplify the shopping experience. Most of the current NFC usage is targeted at payments. However, several retailers are also using NFC scanning as a means to provide customers with additional information about the product. This helps collect information about the products a customer is considering. Because NFC readers are not present in all smartphones, some retailers also use Quick Response (QR) codes for their products that customers can typically scan using an app for similar functionality.

Another new method of collecting customer data is via *Bluetooth beacons*. Beacons use Bluetooth Low Energy, a technology built into recent smartphones. The beacons are placed throughout the store and can detect the Bluetooth signal from a customer's smartphone that is in the vicinity. These devices can send information to the smartphone via specialized apps. In this sense, the beacons provide a lot of flexibility for the retailer to engage with and interact with the customer (assuming that the customer has the specialized app). This can be used to push notifications about products, coupons, etc. in real time to the customer. Furthermore, because the customer interacts with the app to utilize this information, the effect of sending the information to the customer can also be tracked immediately. This technology seems to have a lot of potential for personalizing the retail experience for customers, as well as for collecting information from the customer. As per Kline,[8] nearly a quarter of US retailers have implemented such beacons. Macy's and Rite Aid are some of the prominent retailers to complete a rollout of beacons into most of its stores in 2015.

Some of the most exciting potentials for data collection can be seen in the recently launched *Amazon Go* retail store. The store allows customers to simply grab items and go, without needing to formally check out at a counter. The customer only needs to scan an app while entering the store. The use of a large number of video cameras coupled with *deep learning*-based algorithms make this quite plausible. Deep learning is an area of machine learning that has gained considerable attention recently because of its state-of-the-art ability to decipher unstructured data, especially for image recognition; see Chap. 17 on deep learning. In the retail context, the video cameras capture customers and their actions, and the deep learning algorithms decipher what the actions mean: what items are they grabbing from the shelves and if they are putting back any items from their bag. Such an approach would revolutionize customers' retail experience. However, at the same time, it provides the firm with large amounts of data beyond customer routes. It allows the firm to pick up on moments of indecision, products that were compared, especially when one product is replaced by a similar product.

3 Methodologies

We will focus on product-based analytics to support inventory decisions, assortment, and pricing decisions. The key elements of such analytics are to estimate consumer demand for products, include the case of groups of products, and then take decisions by optimizing over the relevant variables.

Some of the fundamental decisions a retailer makes is to decide on the inventory level and set the price for each SKU. Typically this involves forecasting the demand distribution and then optimizing the decision variable based on the retailer's objective. Forecasting demand is a topic that has received a lot of attention and has a long history of methodologies.

3.1 Product-Based Demand Modeling

Typical forecasting methods consider the univariate time series of sales data and use time-series-based methods such as exponential smoothing and ARIMA models; see Chap. 12 on forecasting analytics. These methods typically focus on forecasting

[8]https://www.huffingtonpost.com/kenny-kline/how-bluetooth-beacons-wil_b_8982720.html (accessed on May 10, 2018).

sales and may require uncensoring to be used for decision-making. Recently, there have been advances that utilize statistical and machine learning approaches to deal with greater amounts of data.

As the number of predictors grow, estimating demand becomes statistically complicated because the potential for overfitting increases. Typically, one deals with such a situation by introducing some "regularization." Penalized L_1 regularization is a common, extremely successful methodology developed to deal with high dimensionality as it performs variable selection. Penalized L_1 regression called LASSO (least absolute shrinkage and selection operator) was introduced in Tibshirani (1996) and in the context of the typical least squares linear regression can be understood as follows: suppose the goal is to predict a response variable $y \in \mathrm{R}^n$ using covariates $X \in \mathrm{R}^{n \times p}$, then the LASSO objective is to solve

$$\min_{\beta \in \mathrm{R}^p} ||y - \mathrm{X}\beta||_2 + \lambda||\beta||_1 \tag{18.1}$$

where $||\cdot||_x$ represents the L_x norm of the expression in parentheses. Such a formulation makes the typical least squares estimator biased because of the regularization term; however, by selecting the regularizer appropriately, the variance can be reduced, so that on the whole, the estimator performs better for prediction. The use of the L_1-norm facilitates sparsity and leads to "better" variable selection. This is especially useful in the case of high-dimensional settings in which the number of parameters p may even exceed the data points n. Prior to the introduction of LASSO, L_2-based regularization, also called ridge regression, was a common way to alleviate overfitting. More recently the elastic net has been proposed that uses both the L_1- and L_2-norms as a regularizer. We direct the reader to the open source book[9] by James et al. (2013) for a detailed discussion of these methodologies. A description is also contained in Chap. 7 on regression analysis.

Recently, Ma et al. (2016) used a LASSO-based approach (along with additional feature selection) to estimate SKU-based sales in a high-dimensional setting, in which the covariates included cross-category promotional information, involving an extremely large number of parameters. They found that including cross-category information improves forecast accuracy by 12.6%.

Ferreira et al. (2015) recently used an alternate machine learning-based approach for forecasting demand in an online fashion retail setting using regression trees. Regression trees are a nonparametric method that involves prediction in a hierarchical manner by creating a number of "splits" in the data. For instance, Fig. 18.1 (reproduced from Ferreira et al. 2015) displays a regression tree with two such splits.

Demand is then predicted by answering the questions pertaining to each of the splits: first, whether the price of the product is less than \$100. If not, then the demand is predicted as 30. Otherwise, the following question is asked: whether the relative price of competing styles is less than 0.8 (i.e., is the price of this style

[9] http://www-bcf.usc.edu/~gareth/ISL/ (accessed on Jul 31, 2018).

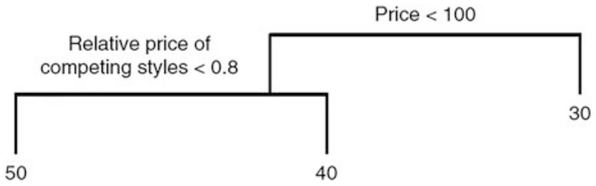

Fig. 18.1 Regression tree example (Reproduced from Ferreira et al. 2015)

less than 80% of the average price of competing styles); if the answer is no, then the demand is predicted as 40. Otherwise, it is predicted as 50. The paper further uses the variance reduction method of bootstrap aggregation or bagging, in which an ensemble of trees is "grown," with each tree trained on a random sampling of the dataset, that is, if the data has N records, then each tree is trained on $m < N$ records randomly sampled from the dataset with replacement. This reduces the interpretability of the model but improves the performance. We refer the reader to the previously cited book on statistical learning and Chap. 16, Machine Learning (Supervised), for details on this methodology. A closely related method is that of random forests, which is similar to bagged trees, except each tree is only allowed to split on a subset of the parameters; this reduces the correlation between the trees and lowers the variance further. Random forests are extremely good out-of-the-box predictors; however, because each tree only uses a subset of parameters for training, its overall interpretability is quite limited.

Recently, (artificial) neural networks (NN) have also been employed for demand forecasting (Au et al. 2008). A neural network is a large group of nodes that are arranged in a layered manner, and the arcs in the network are associated with weights (Sect. 2.1 in Chap. 17 on deep learning). The input to the network is transformed sequentially layer by layer—the input to a layer is used to compute the output of each node in the layer based on the inputs to the node, and this serves as an input to the next layer. In this manner the neural network produces an output for any given input. The weights of the neural network are "trained" typically by gradient descent methods to minimize a loss function that relates to the error between the output and the input. Neural networks can model highly nonlinear dependencies and as such work extremely well in detecting patterns and trends in complex scenarios. Neural networks have been around for a long time; see Chaps. 16 and 17. The well-known logistic regression function can be represented by a single-layer neural network. However, more interesting networks are obtained by creating a large number of layers; hundred-layer neural nets are not uncommon. Deep neural networks are notoriously difficult to train, as they require a lot of data and computational power. With recent advances in data collection and computing, it has become possible to harness the potential of these networks. Initial research demonstrates that NN can be used effectively in the context of predicting fashion demand, and these illustrate the potential for using such methods for demand forecasting in the future. The performance of NNs (and their sophistication) increases as the amount of training data increases. With the spurt in data collection, especially that of unstructured data, NNs provide an exciting potential for demand forecasting.

Turning to decision-making, the application of these statistical and machine learning methodologies generates new challenges and the potential for generating better decisions. For instance, the work of Ferreira et al. (2015) studies the problem of optimizing the prices of the products (at fixed inventory levels). In this case, the split-based demand prediction approach implies that the optimization program becomes an integer program, where, for a fixed set of prices, the decision variable is a binary variable relating to whether a particular <product, price> pair is offered or not. The paper proposes an efficient approximation solution by using a linear programming relaxation. The solution is tested in a field experiment and found to generate a revenue lift of about 10%.

Example—Price and Plan Inventory for a Reseller of Hand Tools:

The sample dataset given in "handtools_reseller.csv" contains information about the demand for products offered by a reseller of refurbished hand tools. The data contains information about the department, the category, the average price of competing products in the category, the MSRP (manufacturer's suggested retail price), the number of competing styles, the total sales events (for all products), and the number of sales events for the product. The data provides the "demand" for all products unlike the article which estimates demand for products that run out of stock. The first five lines of data are shown in Table 18.1.

The regression tree used to train and predict demand included all these variables. It based the fit on a 70–30 split between train and test samples. It grew a full tree and pruned it to match the lowest error on the test set. The details of the approach used are given in Chap. 16 on supervised learning. The final tree looks similar to the one given below (your tree might be slightly different depending on the test and train samples).

In this tree, moving to the left means answering "Yes." A naïve method for setting price and ordering inventory can be based upon this tree. For example, one can focus on introducing a new product into a category and ask what price to charge and how much inventory to order. Suppose the store is considering introducing a new product within a category that already has five competing styles in it. The average price (average) is 125, and the number of sales events (total sale) is 4. The MSRP for the product is also 125. Notice that the tree does not need other information to

Table 18.1 Demand for refurbished hand tools (sample from handtools_reseller.csv)

Product	Department	Category	MSRP	Price	Average price competing	Number of competing styles	Total sales events	Past 12 months sales events	Demand
9728	3	3	417	261	215.4	4	4	3	9
9131	3	2	290	124	133	7	5	1	18
2102	3	1	122	21	50.2	3	1	1	40
1879	1	2	258	84	135.6	3	6	1	38
1515	3	1	133	128	98.6	8	3	2	6

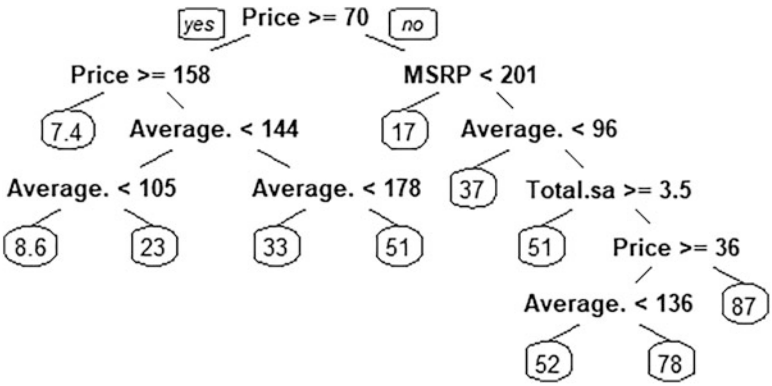

Fig. 18.2 Regression tree output

Table 18.2 Demand prediction model (Tree_Regression_Example.xlsx)

Price	MSRP	Average	Total sale	Price <70	Price >70	Demand	Number of competing styles	Revenue
$157.97	125	125	4	0	23	23	5	$3634.00

Given a price and other variables calculates the revenue
Does not adjust the average if the price changes nor computes impact on other products

predict the demand except the price of the product. We can then create an Excel model shown in Table 18.2.

In this simple model, one can set a price for the product and drop it down the tree to predict the demand using a series of "IF" statements. The demand if price <70 and price >70 are computed separately to minimize the number of "IF" blocks. The demand is the total of the two. Using *Solver*, an add-in to Microsoft Excel, one can maximize the revenue by varying the price. The "optimal" price is $158 and the optimal revenue is $3634. The reader is asked to make several enhancements to the model shown above in the chapter exercises.

The work by Ferreira et al. (2015) takes the typical approach of first estimating demand and then optimizing decisions. From the perspective of decision-making based on available data, the sequential nature of this approach is unnecessary, and one can conceive directly optimizing the decision of interest based on all available data. Indeed, Liyanage and Shanthikumar (2005) prove that formulating the decision as a function of the data and then directly optimizing it can lead to improved performance compared with the sequential approach.

The work by Ban and Rudin (2018) studies such an approach in the context of optimal inventory choice for a single product (at fixed prices). This is the setting of the classical newsvendor problem in which the goal is to select the stock level to minimize shortfall and holding costs (this dates back to Edgeworth 1888 and is a building block for more sophisticated models of stochastic inventory optimization; see Porteus 2002, for more background). The work by Ban and Rudin (2018) considers a feature-based approach in which the demand is assumed to be a function of many features which are observable before the order is placed, such as season and weather. Thus, the newsvendor's goal is to optimize the order quantity based on the observed features, when the demand is uncertain and dependent on these observable features. The paper sets up the optimization problem as a machine learning problem, including regularization as discussed before when dealing with a large number of features, and proves that the problem can be solved using LP, MIP, or QCQP programs. The methodology is applied to the case of nurse staffing in a hospital emergency room and shown to reduce costs by about 24% relative to existing benchmarks.

3.2 Incorporating Consumer Choice in Demand Modeling

The previous discussion focuses on exogenous demand models without getting into specifics on how demand arises. In retail situations, product substitution is abundant wherein when a customer's preferred item is out of stock, they substitute by selecting a different product. (The work reported in Gruen et al. 2002 suggests that such substitution rates can be significant.)

This directly motivates modeling customer preferences over all the products carried by the retailer. One of the workhorse models for such consumer choice modeling is the multinomial logit (MNL); also see Chap. 8 on advanced regression analysis and Chap. 23 on pricing analytics for further applications. The MNL model considers each consumer as selecting a product that maximizes their utility. In particular, the model describes each product-j from a set of numbered products 1,...,N via an average utility v_j so that the utility of a consumer i, from choosing product-j, is then given by $u_{i,j} = v_j + \varepsilon_{ij}$, where ε_{ij} is a zero-mean idiosyncratic noise term. In the MNL, the noise is modeled as a Gumbel distribution,[10] and for this distribution we can define the probability that customer i chooses product-j when given a subset of products A as follows:

$$P\left(Customer\ chooses\ product\ j\right) = \frac{\exp\left(v_j\right)}{1 + \sum_{k \in A} \exp\left(v_k\right)} \tag{18.2}$$

Mixed MNL is a natural extension of this model, in which customer heterogeneity is incorporated. Moreover, utility can be a function of the attributes of the product. In particular, the coefficient β applied to the attributes is not fixed but assumed to differ across customers. The coefficients are typically modeled as arising from a certain distribution with an unknown parameter that is used as part of the estimation process.

[10] An alternative has been proposed in Alptekinoglu and Semple (2016) that considers exponentially distributed noise, and the model is called the exponomial choice (EC) model.

This MNL model has inbuilt an “independence of irrelevant alternatives,” and the model becomes invalid in situations where this effect does not hold (for further details, refer to the book on choice models: Train 2009). To alleviate this, the Nested Logit model has been proposed in which the products are organized into categories in a tree structure, and the customer selects a product by iteratively selecting one of the many categories that the customer needs at each step (i.e., the customer selects the branch of tree). This continues until the customer selects a product (i.e., arrives at a leaf node of the tree).

The MNL model has proved to be extremely valuable in demand estimation when dealing with situations with product substitutability. It has been used extensively in the marketing literature, which uses panel data that includes both household- and store-level data aggregated over time. Such an approach traces back to Guadagni and Little (1983), Chandukala et al. (2008), and Wierenga et al. (2008), who present a detailed overview of choice modeling using panel data.

A small example of inventory planning under static substitution, assortment planning at AbD (inspired by the lecture notes of Dorothee Honhon at UT Dallas):

AbD is a wholesaler of granite blocks that offers a huge assortment of slabs. However, luckily, the slabs are grouped into types of stones, and even within stones, classified based on color and quality. For example, within granite slabs, they offer four grades that are priced differently. These grades come in light to dark shades. There are two modes of supply and two possible assumptions about demand. In the first mode of supply, AbD orders from Brazil after the customer places an order. This method is used only for the highest quality of slabs because the delivery commands a premium price. This method is labeled Make to Order or MTO. For MTO, the demand depends on the assortment offered and is static. There is no carrying cost. However, some customers might not opt to wait that long!

For the rest of the market, AbD uses a make to stock (MTS) approach. In this model, stone is ordered in anticipation of demand. For online sales, AbD has seen that the demand is somewhat predictable within each group. However, because the slabs are ordered in advance, AbD has to carry inventory. Moreover, the customer rarely sees the “stock in hand” and simply has to go with what is offered on the website. AbD has seen that there is little overlap between the variety preferred on the Internet channel and the colors preferred by walk-in customers. Sales also take place to walk in customers. Such in-store sales are highly random. Typically, AbD has seen the variance of the sales in a category to be close to the mean demand and therefore assumes a Poisson distribution for the demand. AbD faces the risk of carrying unsold inventory as well as that of running out of stock. In this case, the customer also sees the inventory. Therefore, the customer (who is typically impatient) might switch to another product within the category or walk out of the store.

For illustration purposes, we consider a sample (Table 18.3) within the dark category of slabs. The name of the slab, the selling price, and the carrying cost are listed. A market survey has identified the customer “gross value” attached to each of the products. The “value” of no purchase is also provided. All figures are in dollars.

Table 18.3 Sample from the dark category of slabs

Name of slab	Price per slab	Value attached by customer	Carrying cost per period
Roman Blue	1440	1750	14.4
Niagara	2323	2500	20
Forever Black	1717	1987	19.87
Violet Black	1744	1800	18

Table 18.4 AbD sells MTO

	Demand, D		5		Stocking cost	175		
	Scale factor, s		100					
	Price (p)	Value (v)	Hold cost	v − p	Exp (v − p)	Prob. purchase	Profit	Less stock cost
Roman Blue	1440	1750	14.4	310	22.19795128	0.4857	\$ 664.49	\$ 489.49
Niagara	2323	2500	20	177	5.870853361	0.1285	\$ 285.58	\$ 110.58
Forever Black	1717	1987	19.87	270	14.87973172	0.3256	\$ 526.71	\$ 351.71
Violet Black	1744	1800	18	56	1.7506725	0.0383	\$ 63.36	\$ (111.64)
						0.9781		
	The scale factor is chosen based on the no purchase probability (here = 0.0219)							
				Expected profit				\$ 840.15

Profit = D * Prob. purchase (Price − Cost − hold cost)
Note that cost = 80% price

For simplicity, assume that the value of no purchase is normalized to zero. The net value to a customer from a purchase of a slab is equal to (Value – Price). Also, assume that each customer will purchase exactly one slab. The cost of a slab is 80% of the price. There is a per unit holding cost on account of interest, storage, breakage, etc. There is a fixed cost of 175 per slab stocked. The average number of customers per period in all examples is 5.

Consider the case when the product is sold as MTO. We use the MNL choice model in this example: In the simplest version of the model the probability that a customer chooses a product i is given by $\pi_i = e^{((v_i - p_i)/s)} / (1 + \sum_j e^{((v_l - p_j)/s)})$; where p_i stands for the price and v_i for the value attached to product-*i*, s is a scale factor, and the sum is taken over all products which are compared with product-*i*. Chapter 23, Pricing Analytics, has further details about this model of consumer choice. Let D be the average demand. Assume the scale factor is 100. The expected profit is then given by

$$D \times \Sigma_j (0.2\ \pi_j p_j - \text{Hold cost}) - \text{number of products stocked} \times \text{stocking cost}$$

The expected profitability from offering all four products in the assortment is computed as shown (the profit is computed after subtracting the stocking cost) in Table 18.4.

The reader is asked to verify these calculations, as well as evaluate whether this is the optimal MTO assortment to offer in the exercises.

Continuing the discussion of consumer choice-based demand modeling, the work in Vulcano et al. (2012) analyzes a model of demand that combines an MNL choice model with nonhomogeneous Poisson model of arrivals. The paper jointly estimates the customer preferences and parameters of the (demand) arrival process using sales data, which includes product availability information. The paper uses an expectation–maximization (EM) algorithm and demonstrates the efficacy of its method using industry datasets, including that from an actual retail chain.

While the MNL is a great parametric class of models, it does have additional shortcoming in modeling product substitution. Some examples of papers that consider alternative models of product substitution are Anupindi et al. (1998), Chong et al. (2001), Bell et al. (2005), Kök and Fisher (2007), and Fisher and Vaidyanathan (2014).

Recently, a Markov chain-based choice model that generalizes the MNL has been proposed in Blanchet et al. (2016). The paper uses state transitions of the Markov chain to model product substitutions and to approximate general models, including a mixture of MNLs and nested logit.

An alternative to these models is the class of nonparametric choice models that have recently been studied in Haensel and Koole (2011), Farias et al. (2013), and Van Ryzin and Vulcano (2017). In these papers, the customer types are defined via a direct ranking of the product types, and when faced with choosing from a set of offered products, the customers go down their rank list and pick the highest-ranked available product (the list includes the no-purchase option, so customers may leave without a purchase). Farias et al. (2013) take a robust optimization approach by solving for customer type distribution that leads to worst-case revenue, while imposing the observed choice data as a constraint. A key result therein is that the demand model obtained is approximately the sparsest choice model. They show a 20% improvement in accuracy over benchmark models for an automobile sales dataset. Haensel and Koole (2011) and Van Ryzin and Vulcano (2017) use EM methods to estimate the unknown parameters. A recent paper (Jagabathula and Vulcano 2017) also proposes a nonparametric approach that focuses on repeated consumer interactions wherein customer preferences may be altered by price or display promotions, and customers exhibit bounded rationality by only considering an unobserved subset of the offered products, referred to as consideration sets. Using a grocery dataset, they show 40% improvement in prediction accuracy over state-of-the-art benchmarks based on variants of the MNL commonly used in current industry practice. Recently, Jagabathula and Rusmevichientong (2016) add to this literature by incorporating pricing by associating customers with a price threshold in addition to the consideration sets.

Turning to decision-making, there is a large collection of papers that have looked at assortment optimization using such choice models. Overall, there are three decisions here: assortment to offer; the inventory levels of each product in assortment; and the price of each item in the assortment. Assortment optimization introduces a complexity because it represents a combinatorial problem that may require enumeration of the different products. One of the pioneering papers by Ryzin and Mahajan (1999) optimizes the inventory and assortment using a MNL-based demand framework. In the paper, a nesting feature is observed, so that the products can be ranked by their value and the optimal assortment only requires considering nested subsets of the ranked products. Empirically, such optimization has been seen to improve a firm's financials significantly, for instance, Kök and Fisher (2007) estimate a 50% increase in profit to a retailer, Fisher and Vaidyanathan (2014) report a sales lift of 3.6% and 5.8% for two different product categories, and Farias et al. (2013) and Jagabathula and Rusmevichientong (2016) estimate about 10% increases in revenue. We direct the reader to Rusmevichientong et al. (2006), Smith et al. (2009), Honhon et al. (2010), and Honhon et al. (2012) for optimized decision-making under variants of rank-based choice models. More recently, there has been a growing interest in optimizing assortments using nonparametric methods. For instance, Bertsimas and Mišic (2015) use a nonparametric choice model related to Farias et al. (2013) but forego the robust approach to directly estimate the choice model by efficiently solving a large-scale linear optimization problem using column generation and then solve the assortment optimization piece based on the solution to a practically tractable mixed integer optimization problem. The previously referenced Jagabathula and Rusmevichientong (2016) solves a joint assortment and pricing problem (which is known to be NP-hard) using an approximation algorithm with a provable performance guarantee based on a DP formulation.

Interestingly, as we move to dynamic assortments, which become more relevant in the context of e-tail, Bernstein et al. (2015) and Golrezaei et al. (2014) solve this problem in a limited inventory setting. The latter, in fact, consider a very general consumer choice model and propose an algorithm that does not require knowledge of customer arrival information.

Example: Assortment over Internet

Consider the previous example and the case when the assortment is offered over the Internet. In this case the assortment is changed every selling season. We model this as MTS problem with *no* substitution. Once the product runs out, the customer who asks for it gets a message that it is out of stock, and s/he walks away with no purchase. We are also told that disposing of unsold product at the end of a period (or selling season) recovers 85% of the cost of the product. This cost is estimated as the cost of the item less the cost of shipping to a discount outlet and holding cost for selling the product at the end of the season. The demand is assumed to be distributed Poisson with mean equal to 5. Though we can calculate the expected profit analytically, to illustrate an alternative method, we will use simulations! For doing this, we generate demand 1000 times. In each simulation, we draw the number of customers according to the Poisson distribution. Then determine which slab is preferred by each of the customers. Given a stocking quantity, it is straightforward but a little tedious to calculate the expected profit. The "Assortment_Examples.xlsx" sheet contains the simulation. The sample summary results when stocking one slab of each type is shown below (data and average profit above and first few rows of simulation below in Tables 18.5 and 18.6).

For example, in the first simulation, three customers arrived. All wanted Roman Blue. The actual sale was for one slab of Roman Blue and one slab each of Niagara,

Table 18.5 AbD sells MTS

	Demand, D		5		Stocking cost	175		
	Scale factor, s		100					
	Price (p)	Value (v)	Stock cost	v-p	Exp (v-p)	Prob. Purchase	Profit	Less stock cost
Roman Blue	1440	1750	14.4	310	22.19795128	0.4857	$ 685.07	$ 510.07
Niagara	2323	2500	20	177	5.870853361	0.1285	$ 284.03	$ 109.03
Forever Black	1717	1987	19.87	270	14.87973172	0.3256	$ 544.66	$ 369.66
Violet Black	1744	1800	18	56	1.7506725	0.0383	$ 52.41	$ (122.59)
						0.9781		
	The scale factor is chosen based on the no purchase probability (here = 0.0219)							
				Expected profit				$ 866.16

Table 18.6 MTS profit using simulation

		Inventory	1	1	1	1	Total
Simulation	Random	Profit	Roman Blue	Niagara	Forever Black	Violet Black	No. of customers
1	0.17144837	$ (1106.08)	3	0	0	0	3
2	0.73233508	$ 195.36	1	3	0	2	6
3	0.98212926	$ 744.80	6	1	2	1	10
4	0.48677441	$ 1.44	3	0	1	1	5
5	0.13994017	$ (556.64)	2	0	1	0	3

Forever Black, and Violet Black had to be salvaged at a loss. The reader is asked to verify the simulation setup in the exercise and then make optimal choices.

AbD sells to walk-in customers or sells MTS with substitution. We can already see how this problem becomes vastly more complicated when there is substitution! In addition to creating arrivals, we have to keep track of the sequence in which the customers arrive and then see if there is stock. If there is stockout, we need to model whether there is substitution from the remaining products or the customer leaves the store empty-handed. Exercise 18.4 gives a simple example to illustrate these ideas.

In many settings, customer preferences are not known, and one may need to learn these while simultaneously optimizing the assortment. Caro and Gallien (2007) and Rusmevichientong et al. (2010) were among the first to study this problem. Caro and Gallien (2007) undertook a Bayesian learning approach in which the underlying primitives have a certain distribution, and the Bayesian approach is used to learn these parameters. On the other hand, Rusmevichientong et al. (2010) used an adaptive learning approach in which such priors are not assumed, and an explore–exploit paradigm is used: in this approach, the decision-maker balances "exploration" to collect relevant data points, with "exploitation" to generate revenue based on the data observed thus far. More recently, the notion of *personalized assortments* is becoming prevalent, especially in e-tail settings, wherein a customer could be shown an assortment of items based on customer-specific information, such as historical data on preferences, demographics, etc. In such a situation, estimating customer choice becomes a high-dimensional estimation problem (dimensionality equals the number of customers times the number of products). A recent paper (Kallus and Udell 2016) considers such a problem and proposes a low-rank mixed multinomial logit choice model in which the customer choice matrix is assumed to have a low-rank latent representation. The paper proposes a nuclear-norm regularized maximum likelihood estimator for learning this model and shows that it can learn the model with few customer interactions. Broadly, with the growing data collection capabilities, we expect further proliferation in such models that estimate individual customer-level choices and use it to make personalized decisions.

4 Business Challenges and Opportunities

4.1 Omni-Channel Retail

The tremendous success of e-commerce has led many retailers to augment their brick-and-mortar stores with an online presence, leading to the advent of multichannel retail. In this approach the retailer has access to multiple channels to engage with and sell to customer. Typically, each of these channels is managed separately. This multichannel approach has been overshadowed by what is commonly referred to as *omni-channel* retail, in which the firm integrates all the channels to provide a seamless experience to customers. A good example of such an approach is the “buy online, pick up in store” (BOPS) approach that has become quite commonplace. This seamless approach inarguably improves the customer experience and overall sales; however, it can lead to unintended outcomes. For instance, Bell et al. (2014) show that such a strategy can reduce online sales and instead lead to an increase in store sales and traffic. In that context, the authors find that additional sales are generated by cross-selling in which the customers who use the BOPS functionality buy additional products in the stores, and further there is a channel effect as well in which online customers may switch to becoming brick-and-mortar customers.

The benefits of the omni-channel approach are even spurring online retailers to foray into physical stores. For instance, recent studies show how, by introducing an offline channel via display showrooms, WarbyParker.com was able to increase both overall and the online channel’s demand (see Bell et al. 2014).

Thus, there is significant value for a retailer to foray into omni-channel. However, while doing so, it is crucial for the firm’s retail analytics to transcend to omni-channel analytics for correct estimation and optimal decision-making.

4.2 Retail Startups

There has been a spurt in retail analytics startups recently. A majority of these companies can be classified as those using technology to aid in data collection and those that are using sophisticated means of analyzing the data itself.

In terms of data collection, there are many startups that cater to the range of retailers both small and large. Some illustrative examples here are Euclid Analytics, which uses in-store Wi-Fi to collect information on customers via their smartphones. The company is able to collect in-store behavior of customers and also data on repeat visits. Collection of Wi-Fi-based information also allows the retailer to track what customers do online while in store. This lets the retailer better understand their customer base, including what products they are researching on their smartphones (showrooming). Another recent startup is Dor, which is targeted at smaller retailers and sells a device that counts the foot traffic in the store. It then provides a dashboard view of the traffic and provides insights to optimize staffing. Startups like Brickstream and Footmarks produce sensors that monitor foot traffic, so associates can react to shoppers in real time. Swirl works with brands like Urban Outfitters, Lord & Taylor, and Timberland to monitor shopper behavior with beacons. Point Inside provides beacons to Target. Startups like Estimote, Shelfbucks, and Bfonics leverage beacons for in-store proximity marketing, such as sending mobile notifications to shoppers about the products they are currently browsing.

Video analytics is another exciting area that startups are getting into. For instance, Brickstream uses video analytics to capture in-depth behavior intelligence. RetailNext is one of the larger startups in this domain and covers a wide gamut of data collection, starting with Wi-Fi-based data collection to more sophisticated methods using video camera feeds. RetailNext also delves into an array of analytics solutions including staffing schedules and some A/B testing. Video analytics helps in heat mapping customer paths and can also be used for loss prevention.

Oak Labs builds interactive touchscreen mirrors that aim to revolutionize the fitting room. The mirror allows customers to explore product recommendations and digitally seek assistance from store associates. The use of technology here enhances the customer experience and collects valuable data at the same time.

PERCH produces interactive digital displays used by brands like Kate Spade and Cole Haan. Blippar focuses on integrating digital and physical domains by using an app that unlocks content upon scanning products. Aila manages interactive in-aisle tablets for stores that provide customers with detailed product information upon scanning a product barcode. These startups focus on improving customer experience by providing more information while also collecting information on how customers choose products that can help the retailer make smarter decisions.

Turning to analytics, some startups such Sku IQ and Orchestro focus on providing a unified view to multichannel firms. Sku IQ provides a unified view of inventory, sales, and customers from all channels. Orchestro focuses on demand estimation by combining data from POS systems, internal ERP, and third-party data into a common view of demand. Orchestro was recently acquired by E2open, a cloud-based supply chain network management systems provider.

Turning to advanced analytics-focused startups, Blue Yonder focuses on analytics to optimize price and inventory replenishment, while Celect focus on using machine learning to optimize product assortments. Finally, Stitch Fix provides an interesting perspective on how personalization for customers may have enormous business potential. Stitch Fix curates fashion apparel and gives recommendations to its members. To do so, it uses machine learning algorithms that make suggestions to human stylists who then use their experience and knowledge to make recommendations to the end customers.

5 Conclusion and Further Reading

This chapter has set out various models and approaches used in the retail industry. The survey is not exhaustive, simply because of changes that take place every day in design, manufacture, and delivery of products and services. Retailing and supply chains are joined together, and any progress in one will lead to changes in the other. The changes do not occur synchronously—due to constant experimentation, opening of new markets, new channels, and proliferation of supply sources—the approach has been opportunistic. The references and the journals that published the papers cited in this chapter are a good starting point for learning more on the subject and staying on top of the developments.

Electronic Supplementary Material

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

- Data 18.1: handtools_reseller.csv
- Data 18.2: Assortment_Examples.xlsx
- Data 18.3: Tree_Regression_Example.xlsx

Exercises

Ex. 18.1 Reseller of Hand Tools:

(a) Replicate the regression tree shown in Fig. 18.2. The details of the procedure can be found in Chap. 16 on supervised learning.
(b) Enhance the model to consider the impact of product price on the average price of competing products.

(c) Enhance the model to not only consider the new product revenue but also the revenue of competing products. To do so, assume the other products have prices 75, 100, 125, 150, and 175, and their MSRPs are equal to their prices.
(d) The store management would like to ensure that the total sales of the new product is at least 25 units. How does this change affect your solution?
(e) The store wishes to order inventory based on the forecast demand. The store manager argues that the demand prediction is just one number! He says, "We need an interval forecast." How would you modify the model to predict an interval (such as (20,26))? What inventory would you order? (Hint: Look at the error associated with the prediction at a node of the regression tree. Can you use this?)

Ex. 18.2 AbD Sells MTO:

(a) Compute the expected profit (shown in the Table 18.4) using the formulae provided in the chapter. You can consult AssortmentExamples.xlsx MTO sheet.
(b) Is this the optimal assortment to offer?
(c) What would change if the stocking cost were to increase to $230 per product stocked?
(d) How would AbD evaluate whether to add a new product to this category? Create your own example and show the analysis.

Ex. 18.3 AbD Sells MTS:

(a) Compute the expected profit (shown in Table 18.5) using the formulae provided in the chapter. Verify the simulation.
(b) Is this the optimal assortment to offer? Can you optimize the expected profit?
(c) What would change if the stocking cost were to increase to $230 per product stocked?
(d) How would AbD evaluate the impact of a 5% increase in price of all products?

Ex. 18.4 AbD Sells to Walk-Ins:

AbD decides to offer only two types of slabs to walk-in customers, Roman Blue and Forever Black. Assume that the choice probabilities are 0.4857 and 0.3256, respectively. If a customer does not find the desired product, s/he will switch to the other product with half the original probability (0.2428 and 0.1628). AbD keeps just one slab of each as inventory. Use the same costs and prices as in the previous exercise. Calculate the expected profit when exactly one customer arrives and also when exactly two customers arrive. (Hint: Enumerate all possible sequences in which a product is demanded.)

References

Alptekinoğlu, A., & Semple, J. H. (2016). The exponomial choice model: A new alternative for assortment and price optimization. *Operations Research, 64*(1), 79–93.

Anupindi, R., Dada, M., & Gupta, S. (1998). Estimation of consumer demand with stock-out based substitution: An application to vending machine products. *Marketing Science, 17*(4), 406–423.

Au, K. F., Choi, T. M., & Yu, Y. (2008). Fashion retail forecasting by evolutionary neural networks. *International Journal of Production Economics, 114*(2), 615–630.

Ban, G.-Y., & Rudin, C. (2018). *The big data newsvendor: Practical insights from machine learning*. Operations Research. Available at SSRN: https://ssrn.com/abstract=2559116. doi.org/10.2139/ssrn.2559116

Bell, D. R., Bonfrer, A., & Chintagunta, P. K. (2005). Recovering stockkeeping-unit-level preferences and response sensitivities from market share models estimated on item aggregates. *Journal of Marketing Research, 42*(2), 169–182.

Bell, D. R., Gallino, S., & Moreno, A. (2014). *Offline showrooms and customer migration in omni-channel retail*. Retrieved August 1, 2018, from https://courses.helsinki.fi/sites/default/files/course-material/4482621/17.3_MIT2014%20Bell.pdf

Bernstein, F., Kök, A. G., & Xie, L. (2015). Dynamic assortment customization with limited inventories. *Manufacturing & Service Operations Management, 17*(4), 538–553.

Bertsimas, D., & Mišic, V. V. (2015). *Data-driven assortment optimization*. Operations Research Center, MIT.

Blanchet, J., Gallego, G., & Goyal, V. (2016). A Markov chain approximation to choice modeling. *Operations Research, 64*(4), 886–905.

Bradlow, E. T., Gangwar, M., Kopalle, P., & Voleti, S. (2017). The role of big data and predictive analytics in retailing. *Journal of Retailing, 93*(1), 79–95.

Caro, F., & Gallien, J. (2007). Dynamic assortment with demand learning for seasonal consumer goods. *Management Science, 53*(2), 276–292.

Chandukala, S. R., Kim, J., Otter, T., Rossi, P. E., & Allenby, G. M. (2008). Choice models in marketing: Economic assumptions, challenges and trends. *Foundations and Trends® in Marketing, 2*(2), 97–184.

Chong, J. K., Ho, T. H., & Tang, C. S. (2001). A modeling framework for category assortment planning. *Manufacturing & Service Operations Management, 3*(3), 191–210.

DeHoratius, N. (2011). Inventory record inaccuracy in retail supply chains. *Wiley encyclopedia of operations research and management science* (pp. 1–14).

DeHoratius, N., & Raman, A. (2008). Inventory record inaccuracy: An empirical analysis. *Management Science, 54*(4), 627–641.

Edgeworth, F. Y. (1888). The mathematical theory of banking. *Journal of the Royal Statistical Society, 51*(1), 113–127.

Farias, V. F., Jagabathula, S., & Shah, D. (2013). A nonparametric approach to modeling choice with limited data. *Management Science, 59*(2), 305–322.

Ferreira, K. J., Lee, B. H. A., & Simchi-Levi, D. (2015). Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & Service Operations Management, 18*(1), 69–88.

Fisher, M., & Vaidyanathan, R. (2014). A demand estimation procedure for retail assortment optimization with results from implementations. *Management Science, 60*(10), 2401–2415.

Golrezaei, N., Nazerzadeh, H., & Rusmevichientong, P. (2014). Real-time optimization of personalized assortments. *Management Science, 60*(6), 1532–1551.

Gruen, T. W., Corsten, D. S., & Bharadwaj, S. (2002). *Retail out-of-stocks: A worldwide examination of extent, causes and consumer responses*. Washington, DC: Grocery Manufacturers of America.

Guadagni, P. M., & Little, J. D. (1983). A logit model of brand choice calibrated on scanner data. *Marketing Science, 2*(3), 203–238.

Haensel, A., & Koole, G. (2011). Estimating unconstrained demand rate functions using customer choice sets. *Journal of Revenue and Pricing Management, 10*(5), 438–454.

Honhon, D., Gaur, V., & Seshadri, S. (2010). Assortment planning and inventory decisions under stockout-based substitution. *Operations Research, 58*(5), 1364–1379.

Honhon, D., Jonnalagedda, S., & Pan, X. A. (2012). Optimal algorithms for assortment selection under ranking-based consumer choice models. *Manufacturing & Service Operations Management, 14*(2), 279–289.

Jagabathula, S., & Rusmevichientong, P. (2016). A nonparametric joint assortment and price choice model. *Management Science, 63*(9), 3128–3145.
Jagabathula, S., & Vulcano, G. (2017). A partial-order-based model to estimate individual preferences using panel data. *Management Science, 64*(4), 1609–1628.
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). New York: Springer.
Kallus, N., & Udell, M. (2016). Dynamic assortment personalization in high dimensions. *arXiv preprint arXiv, 1610*, 05604.
Kök, A. G., & Fisher, M. L. (2007). Demand estimation and assortment optimization under substitution: Methodology and application. *Operations Research, 55*(6), 1001–1021.
Liyanage, L. H., & Shanthikumar, J. G. (2005). A practical inventory control policy using operational statistics. *Operations Research Letters, 33*(4), 341–348.
Ma, S., Fildes, R., & Huang, T. (2016). Demand forecasting with high dimensional data: The case of SKU retail sales forecasting with intra-and inter-category promotional information. *European Journal of Operational Research, 249*(1), 245–257.
Porteus, E. L. (2002). *Foundations of stochastic inventory theory*. Stanford, CA: Stanford University Press.
Rusmevichientong, P., Shen, Z. J. M., & Shmoys, D. B. (2010). Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations Research, 58*(6), 1666–1680.
Rusmevichientong, P., Van Roy, B., & Glynn, P. W. (2006). A nonparametric approach to multiproduct pricing. *Operations Research, 54*(1), 82–98.
Ryzin, G. V., & Mahajan, S. (1999). On the relationship between inventory costs and variety benefits in retail assortments. *Management Science, 45*(11), 1496–1509.
Smith, J. C., Lim, C., & Alptekinoglu, A. (2009). Optimal mixed-integer programming and heuristic methods for a Bilevel Stackelberg product introduction game. *Naval Research Logistics, 56*(8), 714–729.
Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.
Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge: Cambridge University Press. https://eml.berkeley.edu/~train/distant.html.
van Ryzin, G., & Vulcano, G. (2017). An expectation-maximization method to estimate a rank-based choice model of demand. *Operations Research, 65*(2), 396–407.
Vulcano, G., Van Ryzin, G., & Ratliff, R. (2012). Estimating primary demand for substitutable products from sales transaction data. *Operations Research, 60*(2), 313–334.
Wierenga, B., van Bruggen, G. H., & Althuizen, N. A. (2008). Advances in marketing management support systems. In B. Wierenga (Ed.), *Handbook of marketing decision models* (pp. 561–592). Boston, MA: Springer.