International Journal of Research in Marketing 37 (2020) 3–14

Contents lists available at ScienceDirect

IJRM

International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

Retailing and retailing research in the age of big data analytics☆

Marnik G. Dekimpe

[a] Department of Marketing, School of Economics and Management, Tilburg University, Warandelaan 2, 5000LE Tilburg, The Netherlands
[b] Department of Marketing, Faculty of Economics and Business, KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium

ARTICLE INFO

Article history:
First received on July 25, 2019
and was under review for 1 month.
Available online 28 September 2019

Guest Editor: Gary L. Lilien

Keywords:
Retailing
Big data analytics

ABSTRACT

As a research domain, the retail sector has always had many appealing features, such as its size, its multi-faceted and dynamic nature, the possibility for researchers to exploit their own domain knowledge, and an extensive coverage by business analysts. In addition, the above-average availability of good-quality data has historically been an additional selling point to empirical researchers. The paper considers to what extent the latter still holds, and explores a number of additional opportunities and challenges that emerge from the ongoing big data revolution. This is done from five perspectives: retail managers, retailing researchers, public-policy makers, investors, and retailing educators.

1. Introduction

The papers that have appeared in IJRM's invited "EMAC Distinguished Scholar" series have covered a wide range of topics. A common theme in most of these contributions is that the authors have taken the opportunity to reflect not only on the *past*, but also (and even more so) on the *future* of the field. In so doing, they have focused either on the marketing discipline as a whole (see, e.g., the contributions in the current issue by Don Lehmann (2019) and Roland Rust (2019)) or on the sub-field they are most closely associated with (see, e.g., Wierenga (2011) on managerial decision making, or Lilien (2016) on the B2B knowledge gap).

In terms of my own scientific career, my first professional publication was on the survival rate of retail stores (Dekimpe & Morrison, 1991), and throughout the subsequent years/decades, most of my work has continued to focus on retailing-related issues. In this article, I will first reflect on why I feel the retailing sector is a fascinating and fertile ground for managerially relevant academic research. I will argue how the above-average availability of good-quality data has historically been a key selling point for empirical researchers. After that, I will consider to what extent this still holds, and explore a number of additional opportunities and challenges that emerge from the ongoing, and rapidly accelerating, big data revolution. I will do so from five perspectives: retail managers, retailing researchers, as well as public-policy makers, investors, and retailing educators.

The retailing sector is an ideal setting for such a discussion, as analysts like to present the sector as a poster child for all the benefits that one can envision from the use of big data. Industry reports have proclaimed that retailing is "one of the hottest markets for big data analytics" (Ingram Micro, 2018), that "big data is especially promising and differentiating for retailers" (IBM-

☆ This paper emerged from talks given on the occasion of my 2016 European Marketing Academy Distinguished Marketing Scholar Award and my honorary doctorate at the University of Hamburg in 2018. I have benefitted greatly from discussions throughout the years with Inge Geyskens, Els Gijsbrechts, Katrijn Gielens, Mike Hanssens, John Roberts, Jan-Benedict Steenkamp and Harald van Heerde on various topics covered in the paper. I am also very grateful to all my co-authors on previous retail projects, and to AiMark for the generous data access over the years, both before and during the big data era.
E-mail address: M.G.Dekimpe@uvt.nl.

https://doi.org/10.1016/j.ijresmar.2019.09.001
0167-8116/

Analytics, 2018), or that big data will be “a complete game changer in the retail sector” (Forbes, 2015). A 2011 McKinsey report put forward that big data have the potential to increase retailers' operating margins with up to 60%, and cause a sector-wide annual productivity gain of up to 1% in the next five years (McKinsey, 2011). Similar enthusiastic endorsements have appeared in the academic literature. Grewal, Roggeveen, and Nordfält (2017), for example, argue how new technologies along with big data/predictive analytics will cause a quantum leap in retailers' understanding of the shopping process.

At the same time, concerns are increasingly raised that firms (retailers) that invest heavily in big data might well face negative returns on those investments (Arthur, 2013; Verhoef, Kooge, & Walk, 2016). In a 2016 PlanetRetail study, almost half of the surveyed retailers felt that data proliferation had already reached near unsustainable levels (PlanetRetail, 2016), and five years after their initial study, McKinsey computed that less than 40% of the presumed potential had actually materialized. Moreover, much of the already realized added value appears to have gone to consumers, with very limited change in the retailer's own EBITA (McKinsey, 2016). In a similar vein, concerns are increasingly raised in the academic community that the abundance of big data will lead to an unhealthy bias in the type of problems being studied (Houston, 2016, 2019; McAlister, 2016).

These observations suggest that retailing, for both practitioners and academics, is at the center of a storm of big data opportunities and challenges, which calls for more work on how to derive more value from big data.

2. Retailing: an attractive research domain

As a research domain, the retailing sector has always had many appealing features, such as (i) its size, (ii) its multi-faceted and dynamic nature, (iii) the possibility for researchers to exploit their personal domain knowledge, (iv) an extensive (not necessarily consistent) coverage by business analysts, and (v) good data availability. While the retail sector is not the only sector with (some of) these features, their combined presence makes it a very fertile ecosystem to study.

2.1. Size of the sector

Retailing is big business, and plays a central role in all countries' economy. The retailing industry is one of the largest and most diversified industries in the world, and several retailers rank among the biggest corporations in the world (see, for example, Table 18.1 in Geyskens, 2018). At the same time, the retailing sector still contains a huge number of small “mom-and-pop” operations (Bronnenberg & Ellickson, 2015), making the sector essential for the livelihood of many families.

In Europe, even when limiting oneself to grocery retailing, sales forecasts reach 2289 billion euros by 2022 (IGD, 2018), with millions of people employed in the sector. Given its size and ubiquitous presence, retailing research has the potential to benefit/influence many stakeholders.

2.2. Multi-faceted and dynamic

The retail sector involves multiple parties, including consumers, manufacturers (suppliers), retail chains, investors and policy makers. Depending on the stakeholder or the dyadic relationship one focuses on, key phenomena can be studied from different angles.[1] For example, private-label (PL) growth has been studied from a consumer (Ailawadi, Neslin, & Gedenk, 2011), manufacturer (Gielens, 2012), retailer (Ailawadi, Pauwels, & Steenkamp, 2008), and public-policy (Institute of European and Comparative Law, 2015) point of view, and has been linked not only to a shift in power from the manufacturer to the retailer (Farris & Ailawadi, 1992), but also to reduced innovation activities in the sector (Ezrachi & Bernitz, 2009). In addition, the phenomenon has been studied both at the level of an individual retail chain (ter Braak, Dekimpe, & Geyskens, 2013), at the national level (Lamey, Deleersnyder, Dekimpe, & Steenkamp, 2007), and in a global context (Steenkamp & Geyskens, 2014).

Moreover, the sector is in a constant flux, as already recognized in early discussions on the Wheel-of-Retailing concept. This non-stationarity can be exploited either by focusing on gradual evolutions (Lamey et al. (2007), for example, link the over-time growth in private-label share to business-cycle fluctuations) or on discrete, often highly publicized, events. Examples of the latter are price wars (van Heerde, Gijsbrechts, & Pauwels, 2008), product-harm crises (Cleeren, van Heerde, & Dekimpe, 2013), and conflict delistings (Van der Maelen, Breugelmans, & Cleeren, 2018). On top of that, the digital transformation has initiated a large number of further changes to the retail value chain, as recently reviewed in Reinartz, Wiegand, and Imschloss (2019).

Because of this multi-faceted and highly dynamic nature, it has been argued (Reinartz, Dellaert, Manfred, Kumar, & Varadarajan, 2011) that several of the challenges faced by the sector become more daunting than the ones faced in some other sectors, such as automobiles or computers. This, of course, makes the sector especially appealing from a research perspective.

2.3. Domain knowledge

In our personal lives, we are all consumers, we visit stores on a regular basis, and have considerable experience making off- and/or online purchases. As discussed in Lilien (2016), this is a very different situation from the one faced by B2B researchers. A key advantage of this personal experience is that we can observe discrepancies between what we teach in our retailing courses and what we observe in the marketplace. Retailing textbooks, for example, posit that convenience goods should be distributed as

[1] Because of the diverse nature of the industry, it is not possible to provide examples on all sub-sectors. In the remainder of the article, I will (in line with the nature of most of my research), focus primarily on grocery retailing, even though the substantive ideas apply to other retailers as well.

intensively as possible. Gielens, Gijsbrechts, and Dekimpe (2014), however, observed an increasing number of exclusivity agreements, where popular grocery products (produced and owned by national-brand manufacturers and clearly positioned as such) are for sale at just one retail chain. Following that observation, they developed a framework to identify potential gains and losses to both the retailer and the manufacturer involved in such an agreement, along with ways to develop compensating mechanisms for the losing party. Similarly, at some point, a leading tissue manufacturer proudly advertised how it was putting the same effort in the production of PL products as in its own (national) brands, which contradicts textbook discussions on the intense competition between retail chains and national-brand manufacturers, and why such production agreements (if they would exist at all) should best be kept secret. This apparent contradiction inspired ter Braak, Deleersnyder, Geyskens, and Dekimpe (2014) to study the underlying reasons (and benefits to be expected) of PL production by leading national-brand manufacturers.

More generally, easily observable practices often seem to leapfrog academic thinking, which can offer inspiration for new conceptual and empirical work (I refer to Dekimpe & Geyskens, 2019 for a more in-depth discussion on this issue).

2.4. Coverage by business analysts

Apart from personal experiences and observations, retail researchers also benefit from the extensive (and not necessarily consistent) coverage of the sector by business analysts. For example, Edge Retail Insight (formerly PlanetRetail RNG) tracks over 2000 leading retailers worldwide, and publishes both daily news updates and summary reports on numerous retail trends. Similar reports are published on a regular basis by, among others, Euromonitor, AC Nielsen, AT Kearney and IRI.[2] These reports not only contain a wealth of information, they also regularly posit opposing expectations, which offers a natural opening to position one's own research. For example, the analysts of Rabobank (2013, p. 1) predict that within 15 years “the main Asian grocery markets, India and China, will have closed in on the PL share currently seen in Europe,” while the analysts of Euromonitor (2014, p. 6) expect PLs to “only make gradual, if any, inroads into emerging markets in the years to come”. Clearly, such discrepancies between different “expert opinions” are an open invitation for more academic work.

2.5. Above-average data availability

Retailing problems can, as many other marketing problems, be studied using a variety of data sources. For example, to study the determinants of store choice, both data collected through conjoint experiments and through surveys have been used. However, a key distinguishing feature of the retail sector is its long history of scanner-data availability (Inmann & Nikolova, 2017; Wedel & Kannan, 2016), which has given a strong impetus to econometric research on the sector.

Apart from store-level scanner data, also household panel data have been collected for many years by research agencies such as GfK and Kantar Worldpanel. In addition, several research institutes have made data access easier to academic researchers. For example, over the years, AiMark has provided consumer-panel data for more than 50 academic publications, and currently offers data access (covering countries ranging from Europe to Asia, as well as North and South America) to over 100 academic researchers. In Asia, SMU's Retail Centre of Excellence serves as a linking pin between academic research and business practice in the region. Apart from helping with data access, such research institutes also facilitate the early identification of managerially relevant research problems.

Importantly, this long history of panel-data availability has also provided retailers with a headstart to deal with the recent big data explosion, even though some have argued that the presence of previous (legacy) data systems may actually be a liability in moving to a more flexible architecture needed to take full advantage of the big data opportunities (McKinsey, 2016).

3. Retailing and big data

Retailing is almost by definition a big data industry. At the macro level, thousands of stores sell hundreds of thousands of SKUs to millions of customers through billions of transactions. Walmart, for example, operates more than 11,000 stores in more than 25 countries, serves more than 35 million customers on a daily basis, and sells around 140,000 items in most of its supercenters. At a micro level, individual consumers have become walking data generators that leave a data trail each time they make a purchase with their credit card, use a loyalty card, send a text message, or search the web (Muller, 2014). Kroger, for example, uses the information on the shopping behavior of its 40+ million cardholders to send a large fraction of them personalized coupons on a regular basis.

Given the sheer *volume* of data that retailers can collect on their customers' (on- and/or offline) transactions, data-driven customization has become a real possibility. Moreover, by augmenting these data with information on the status of inventories throughout the supply chain, location-specific weather data, a multitude of social-media metrics, and/or sensor data, the data at a retailer's disposal typically exhibit considerable *variety*, and include both very structured and highly unstructured data. Given the nature of the business, a timely response to incoming data is essential (which is often referred to as the *velocity* dimension of big data). For example, to take full advantage of predictive analytics to know which SKUs sell more on days with certain weather conditions, in a certain type of store, at a specific location, real-time (or near real-time) adjustments are crucial to avoid stock-outs. Accuracy of both data input and subsequent analysis (*veracity*) is essential as well, given that pricing errors

[2] As a case in point, the rise of the hard-discount format has recently been documented and analyzed in, among others, PlanetRetail (2018), Bain and Company (2017) and BCG (2017).

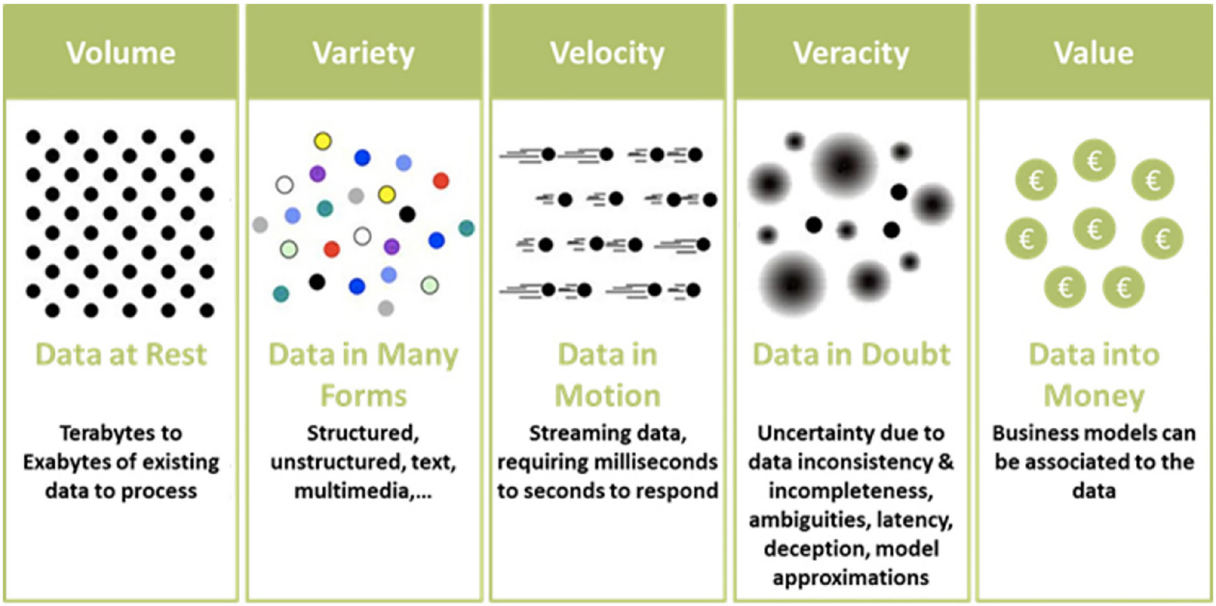

Fig. 1. The 5 Vs of big data.
Figure taken from Heremans (2018).

(Goodstein, 1994) or less appropriate recommendations (Bradlow, Gangwar, Kopalle, & Voleti, 2017) have been found to quickly lead to customer irritation and attrition. Finally, given the tiny margins retailers (and especially grocery retailers) tend to operate on, an adequate assessment on the monetary *value* implications of any data investments is essential. As such, all 5 "Vs" typically brought to the fore in big data discussions, and summarized in Fig. 1, are highly relevant in a retail setting.

Still, there appears to be considerable discrepancy between the intrinsic value *potential* of big data to the sector on the one hand, and the *actual* ease of capturing those benefits on the other hand. Even though the sector is characterized by quite a number of data-rich chains that are also on the forefront of big data analytics, most of the players are much smaller, and thus have fewer resources to gather and/or analyze the data, and to take full advantage of the big data opportunities. But also bigger retailers often miss a full understanding of the potential benefits of big data analytics, and are either not willing to invest at a level that would be commensurate to those benefits (Germann, Lilien, Fiedler, & Kraus, 2014), or struggle to gain actionable customer insights from the increasing amount of available data (Leeflang, Verhoef, Dahlström, & Freundt, 2014).

4. Big data: opportunities and challenges to retail managers

Big data offer retailers numerous new opportunities to create value. While various reports (see, for example, McAfee & Brynjolfsson, 2012; Datameer, 2015) have emphasized the obvious potential for improved micro-targeting, opportunities are present throughout the entire value chain. The aforementioned McKinsey (2011) report identified in this respect 16 big data levers that retail managers can benefit from, which they organized (p. 67) along five functional domains: (i) *marketing* (cross-selling, location-based marketing, in-store behavior analysis, customer micro-segmentation, sentiment analysis and enhancing the multi-channel consumer experience), (ii) *merchandising* (assortment optimization, pricing optimization, and placement & design optimization), (iii) *operations* (performance transparency and labor inputs optimization), (iv) *supply chain* (inventory management, distribution and logistics optimization, and informing of supplier negotiations), and (v) *new business models* (price comparison services and web-based markets).

Still, to maximally take advantage of (even a subset of) these opportunities, managers should avoid a number of pitfalls, and adequately deal with a number of challenges. Specifically:

4.1. Look beyond the hype

The business press is filled with anecdotes and case studies that are meant to demonstrate the benefits of being big data driven. However, not only is there a bias toward reporting on the more successful applications (where Tesco, Walmart and Amazon are, not surprisingly, often featured), the same stories are also recurring (with some variations) in multiple reports/blogs (which may create a false impression of success frequency). Moreover, some of these reports are written by technology firms that have a vested interest in the wider adoption of their own software and/or hardware solutions (see also Weinberg, Davis, & Berger, 2013 for a discussion on this concern). As such, the reported success stories may not be representative, may not be fully objective, and may not be de facto applicable to other (perhaps smaller or less sophisticated) retailers. By thinking "too big" too soon, retailers may become disappointed and abandon their data-driven initiatives prematurely. There is a clear need for academic research that (i) *objectively* documents (ii) *for what retailers*, (ii) *in what settings*, and (iv) *for what type of decisions*, there is demonstrable value to being (big) data driven.

4.2. Avoid over-confidence and over-reliance on big data analytics

Managers may be inclined to automatically assume that "numbers don't lie", that bigger data are by definition better data, and become more susceptible to apophenia (seeing patterns where none actually exist) as the sample size on which some correlations are computed or parameters estimated, increases (Boyd & Crawford, 2012). Even though a data set may contain millions of observations, this does not necessarily mean that the data is free from errors, nor that the data set is representative. Also, when combining multiple data sets, which each may contain some errors or biases, the problem may become magnified. As a corollary, managers should not lose sight of the value of "small data",[3] nor automatically dismiss their gut feeling when the latter is not aligned with the outcome of some extensive number crunching (where modeling choices from the computer scientists/marketing analysts with less experience with the business may well have driven some of the findings). Finally, managers may have a tendency to keep investing in larger quantities of the same (type of) data, while other relevant pieces of information remain missing. For example, Dekimpe and Geyskens (2019) deplore how most data providers still focus almost exclusively on measuring the advertising spending of national-brand manufacturers, without doing the same for category-specific and/or tier-specific spending by retailers to support their private labels. Also, in almost all instances, having a 360 degree picture of the customer, which covers all possible touchpoints, remains an illusion.

4.3. Don't put your customers' trust on the line

Retailers will have to critically evaluate how far they want to go in "exploiting" the data they have collected on their customers. Indeed, just because certain data are accessible, does not make it ethical to use them (Boyd & Crawford, 2012). Customer trust, which is essential to a retailer's success, takes time to build up, but can be lost very quickly if customers feel their data have been used without their permission, or have not been adequately protected. To complicate matters, customers may have fewer problems receiving a recommendation from a retailer they personally know than when receiving the same suggestion through an automated recommendation system. Similarly, they may have no problem receiving recommendations for some product categories, but may feel that their privacy is intruded when receiving a similar suggestion for other categories. More research along the lines of Longoni, Nonezzi, and Morewedge (2019) study on patients' resistance to medical Artificial Intelligence is needed on how and when certain customers appreciate, or even expect, personalized advice.

4.4. Not every problem requires big data

Big data are especially useful for more tactical and frequently recurring decisions such as the fine-tuning of the promotional calendar, shelf-space allocation, or the development of targeted offerings to loyalty-card holders, but may be much less useful for strategic decisions (see also Allenby, Bradlow, George, Liechty, & McCulloch, 2014 for a similar observation). And what to do with strategic decisions for which (almost) no data are available? If big data thinking becomes the norm, there is a real danger that tactical thinking will take the fore over more strategic thinking in retail managers' time allocation. More research is also needed on how to proceed from micro-analytics, which may be highly relevant for category managers, to macro-analytics in support of the retail organization's top management (Sorescu, 2017). Finally, the abundance of secondary data may lead one to abandon primary data collection efforts, even though many important organizational issues faced by retail managers can, as also emphasized by Moorman (2016), best (if not only) be addressed with targeted survey data. As a general rule, the size of the data should fit the research question; and in some cases, small may still be best.

4.5. Big data capabilities need to be developed

Just having (access to) big data is not a sufficient condition to automatically derive value from these data. Verhoef et al. (2016) discuss how one should invest in big data capabilities along four dimensions: (i) people, (ii) systems (especially in terms of the integration of multiple data sources, which can be a real challenge for both multichannel and internationally operating retailers), (iii) processes, and (iv) organization. Unfortunately, the retail sector has been found to lag behind other sectors in their willingness to do so (Germann et al., 2014), which again underscores the need for objective measures, across a broad set of implementations, of the business surplus created by retailers' big data investments. More research is also needed (see also Sorescu, Frambach, Singh, Rangaswamy, & Bridges, 2011 or Sorescu, 2017) in terms of what business-process adjustments are most likely to facilitate not only *value creation* (for the retailer's customers), but also *value appropriation* (for the retailer and/or its partners) of big data investments. As indicated before, many retailers fall short on the latter.

Relatedly, more research is called for on whether intelligence departments should be staff departments that serve the retailers' marketing and sales department, or whether it would be better to integrate this function within the marketing/sales department?

[3] Fournier and Rietveld (2014) even make a case for a qualitative (almost anthropologist) analysis of social listening data to complement the commonly used econometric/statistical analyses that focus on aggregate summary statistics derived from many entries.

Also, should one develop data-analytic capabilities in house, or outsource to outside specialists? How can one ensure optimal communication channels between domain experts (e.g. category managers) and data scientists? Is there a special role or need for a Chief Data Officer (CDO) at the C-level (Brown, 2014; Redman, 2014)? If so, from what point onwards and with what responsibilities?

5. Big data in retailing: opportunities and challenges for retail researchers

The big data revolution has been a real blessing to retail researchers. First of all, each big data lever identified in Section 4 has already been the subject of multiple research projects. To give just a few examples, Larson, Bradlow, and Fader (2005) were able to analyze the path taken by individual shoppers in a grocery store using RFID tags on their shopping carts, while Grewal, Ahlbom, and Beitelspacher (2018) studied, using eye-tracking technology, the link between in-store mobile phone use and customer shopping behavior. Rooderkerk, Van Heerde, and Bijmolt (2013) developed a scalable optimization method to optimize real-life large-scale assortments at the individual store level, while Gielens et al. (2014) analyzed the purchase records of thousands of customers across six retail chains across 50+ product categories to provide a decision aid in the supplier-manufacturer negotiations on exclusivity contracts. Further examples of how retailers have successfully used (big) data about their customers and their business operations to improve their value exploration and exploitation can be found in Fisher and Raman (2010, 2018).

None of these studies would have been possible without the presence of and access to some unique big data sets. The advent of extensive new data sources has clearly allowed retail researchers to study new problems. As described in Bradlow et al. (2017), big data not only refer to more rows of data (on more consumers, on more time periods, on more SKUs, …), but also (and from a research point of view perhaps more importantly) to "new columns" and "better columns". Having better information on new, substantively relevant, variables naturally opens up a multitude of research possibilities. Using once more the earlier big data lever classification, Table 1 lists per topic an exemplary (and by no means exhaustive) list of pertinent research questions that have emerged through recent data advances.

The fact that comparable data sets have become available across multiple entities (which could be individual stores, retail chains, countries, or SKUs/brands/categories within those stores/chains or countries) has enabled the derivation of a very rich set of retail-related empirical generalizations (EGs). A large number of the EGs compiled in Hanssens (2015) have a clear retail focus, such as the generalizations on, respectively, the relationship between PL shares and the business cycle (p. 23), the size of PL margins (p. 113), the effectiveness of within-store vs. between-store semantic price cues (p. 115), or the relationship between PLs and story loyalty (p. 119), to name just a few.

Moreover, the increased data availability has allowed researchers to empirically verify a much richer set of boundary conditions. For example, several studies (see, for example, Pauwels, Srinivasan, & Franses, 2007; Srinivasan, Pauwels, Hanssens, & Dekimpe, 2004) have investigated under what conditions price promotions are more or less likely to be beneficial to retailers and/or manufacturers. Given that contextual factors "set the boundaries of generalizability and as such constitute the range of a theory" (Whetten, 1989, p. 492), big-data-enabled analyses have contributed significantly to the development and refinement of retail theories.

Still, the big data revolution also brings along a new set of challenges to retail researchers:

5.1. More difficult balance between technical complexity and managerial relevance

The massive size and high dimensionality of many secondary data bases have brought along a new set of computational and statistical problems, such as scalability, noise accumulation, spurious correlations, and incidental endogeneity. I will not elaborate on these technical complications, given that excellent reviews on these issues have recently appeared in, among others, Ansari and Li (2018), Fan, Han, and Liu (2014), Ng (2017) and Wedel and Kannan (2016). However, I do want to warn once more[4] for the imminent danger that because of the increased methodological sophistication that is required, less attention is likely to be paid to the managerial question that should be asked. An adequate treatment of these technical/statistical issues is undoubtedly important, but foremost, retail researchers should "keep their eye on the problem." I fully agree with Wendy Moe, who states: "Data is data. I think all data is just data. … it doesn't matter if it's big data or small or medium-sized data. *The key is still knowing what the research objective is, from an academic standpoint and from a practitioner's standpoint*" (Weinberg et al., 2013, p. 190, italics added). The bigger the data, however, the more difficult it may be to keep that lesson in mind, and the danger that the retail problems studied will be "guided more by the cool data sets we can find or the advanced methodological techniques we can employ" (Houston, 2016, p. 559) looms larger than ever.

5.2. More difficult to "listen to the data"

Laurent (2013) urges us to let the data speak, and perhaps especially so beyond predefined scripts. Even though there is an (applaudable) tendency to provide model-free evidence prior to bringing in the heavy modeling artillery, the exploding size, especially across columns and data sources, makes it harder to "truly listen" to the data. As eloquently put by John Roberts[5]: "The days of the data whisperer seem to be over. S/he has been replaced by guys and gals with big statistical sledge hammers who

[4] See, among others, Houston (2016, 2019) or McAlister (2016) for a similar assessment.
[5] Personal communication, March 2019.

Table 1
An illustrative research agenda organized by substantive big data lever.

Marketing

- *Cross-selling*
  - How to optimize in-store promotions to better link complementary items and design optimal promotional bundles?
  - How to better exploit in-store adjacencies, both offline and online?
- *Location-based marketing*
  - How to optimize location-based targeting? Who to best target at what distance of the store with what type of message?
  - How to blend in-store solutions (like beacons) with mobile technology to deliver real-time location-based offers and rewards? How to best balance in-store engagement with the shopper and creating a feeling of privacy invasion?
  - Can regional retail flows be inferred from social-media data?
- *In-store behavior analysis*
  - How to best track in-store traffic patterns, either through shopping-card transponders or by passively monitoring the location of mobile phones? How to best respond to short-term fluctuations in such traffic patterns?
  - How to use video analysis and facial recognition to create better in-store consumer experiences?
- *Customer micro-segmentation*
  - Which customers prefer a personalized recommendation, for what type of purchase (e.g. bulk versus fill-in) for what categories, and in what format?
  - How to leverage big data to understand how shoppers behave online? How to use those insights to reduce shopping-cart abandonment?
  - How to create more retailer-owned touchpoints to enhance customer loyalty? How to work together with social-media platforms and other intermediaries to get better insights on, and greater consistency among, the various touchpoints?
- *Sentiment analysis*
  - How to react quickly to, rather than report on, trends, risks, threats and opportunities?
  - How to facilitate, engage with, and reward shoppers for sharing positive shopping experiences?
  - How can retailers become part of shoppers' conversations on social media?
- *Enhancing the multichannel consumer experience*
  - Should one empower the digitally connected customer even further?
  - How to harness the power of mobile technologies to enhance the connected shopper's cross-channel experience?
  - How to optimally blend digital technologies with consumers' physical retail experiences?

Merchandising

- *Assortment optimization*
  - How to balance cost (production) efficiency with increased demands for product personalization?
  - How to best create assortment differentiation? Through a diverse own-label offering, through product exclusives, or through limited edition ranges?
  - How to rethink the concept of impulse purchasing for convenience-driven customers?
  - How to improve new-product success rates? What is the added value of attribute-based approaches?
- *Pricing optimization*
  - How to develop dynamic best-response pricing algorithms that take into account consumer choice behavior, competitors' actions and supply parameters? How to accommodate local weather conditions?
- *Placement and design optimization*
  - What brands gain most from preferential locations, leading brands or smaller one?
  - What is the impact of surrounding brands on the success of a focal brand? Does this differ between online and offline placements?
  - How to win the digital-shelf (third-shelf) battle?

Operations

- *Performance transparency*
  - Is it advisable to give the shopper a level of price transparency that involves a break-down of the costs associated with making the product? Should this only be done for premium products to motivate the higher price? And can this help increase consumer acceptance of higher-priced organic and/or free-trade products?
  - How to best identify under-performing stores, and how to "optimize" the store liquidation process?
- *Labor inputs optimization*
  - How to best empower employees? What is the added value of mobile and wearable technologies?
  - How to improve employee acceptance of technological in-store assistant systems?

Supply chain

- *Inventory management*
  - How to obtain (and capitalize on) a real-time view of inventory across the entire supply chain, and across all sales channels (online as well as offline)?
  - How to optimize and evaluate backroom automation?
- *Distribution and logistics optimization*
  - How to cost-efficiently shorten customer order fulfillment time?
  - How to best tackle the last-mile problem? How to balance shopper expectations and feasibility?
  - Should one enter into partnerships with other, even competing, retailers? How about leveraging mom-and-pop stores in emerging markets?
  - How to link home automation to e-commerce?
  - How to identify/deal with serial returners?
- *Informing supplier negotiations*
  - How to enhance collaboration between (PL) product development and supply-chain teams to quickly adapt to consumer trends and changing shopper preferences?

(continued on next page)

Table 1. (*continued.*)

- What data can (should) best be shared in order to create win-win situation? How to ensure consumer privacy in such situations?

New business models

- *Price comparison services*
  - How can retailers that cannot compete on price best differentiate themselves in a price-transparent world?
  - How can retailers that do compete on price obtain optimal visibility on price comparison sites?
  - How to account for cross-location differences in price sensitivity in one's clientele?
- *Web-based markets*
  - What are the implications of D2C brand websites and D2C marketplaces on incumbent retailers? What are the best defense mechanisms?

first smash the data to smithereens and then examine the results to see what is left." More work is needed, when confronted with hundreds of columns of data, often derived from initially unstructured data sources, on how to provide the retail manager and/or academic reader with a genuine feel of the most relevant model-free information.

An often-ignored issue in this respect is the representativeness of the data: merely having a large number of rows does not make the sample unbiased. An excellent discussion in the context of Twitter data is provided in Boyd and Crawford (2012). Researchers have a tendency to describe their smaller data sets in great precision (think, for example, of how survey researchers take great effort to convince the reader that there is no non-response bias, or that common-method biases are negligible). Similar care is taken when describing the data collection procedure in experimental settings. Descriptions of big data collection efforts, in contrast, often appear to be done with a very thick brush, even though these data sets may also be unreliable or not representative. Similarly, concerns about measurement comparability seem to become less prominent in big data cross-national studies than they are in cross-national survey research.

5.3. How much should one still "listen to the theory"?

Retailing research has a long tradition, and a rich set of theories. Still, when confronted with an abundance of data, there may well be a tendency to ignore well-established theoretical insights, and to blindly "throw in" all possible variables, in the hope of finding some significant effects. Some authors (see, for example, Anderson, 2008) have even proclaimed that data abundance has made theory obsolete. In sharp contrast to this opinion, both Bradlow et al. (2017) and Allenby et al. (2014) have made a compelling case of how theory remains essential to guide researchers "where to look in the data".

Still, one cannot ignore that in some cases, retailing theory may not be sufficiently developed to formulate formal hypotheses for all potentially relevant effects, especially when they involve (higher-order) interactions. Both Lamey, Deleersnyder, Steenkamp, and Dekimpe (2018) and Steenkamp and Geyskens (2014) have used this argument to test in an exploratory way a larger number of interactions, after which the significant ones were retained, and a post-hoc explanation provided. The discussion on what should come first, Theory or Empirics, is not new, however, and precedes the big data discussion (see, for example, Bass, 1995 on the merits of both ETET and TETE, or Alba, 2011 on the defense of bumbling). The emergence of very large data sets with many columns, along with the increasing popularity of data-mining techniques, has brought the discussion once more to the fore. I think there is indeed a shift to be expected (in relative-occurrence terms) from more theory development-focused research to more data-driven research (see also Varadarajan, 2019 for a similar observation). Still, I am convinced that "theory will never become obsolete". To managers, black-box predictions are notoriously dangerous when doing policy (what-if) simulations without knowing the underlying mechanisms at work, for which theory building is essential (Coveney, Dougherty, & Highfield, 2016). To academic researchers, theories will remain crucial to reconcile (or contrast) new empirical findings with what is already known. If not, the field would rapidly become a haphazard accumulation of disparate facts, and loose much of its appeal.

Having large troves of data should therefore not stand in the way of solid theory development and construct definitions/ operationalizations, even though this may well be tempting to some. On the contrary, I remain convinced that "*some of the major challenges of tomorrow's retail researchers will remain conceptual, not methodological.*" (Dekimpe & Geyskens, 2019, p. 9, italics added). Building on Coveney et al. (2016), one could even argue that big data will need even bigger theories.

5.4. How to remain relevant to practice?

As more and more retail giants develop their own analytic departments, they may feel less of a need to share their data for academic research purposes. In certain disciplines, it has already been argued that academic researchers should no longer bother with research that industry can "presumably" do better (Conover, 2011; see also the discussion in Boyd & Crawford, 2012). Alternatively, retailers may only make a small (not necessarily representative) sample of their data available for academic research, and do so only to a select set of institutions/researchers. This might limit the type of academic research questions being asked (or answers published) by fear of losing one's data access. Moreover, limiting the pool of researchers that can study a certain problem (by denying them the necessary data access) has been shown to eventually hinder the discipline as a whole (see in this respect also Stremersch & Verhoef, 2005).

Assuming that retailers remain open to work with academics, the question will be to what discipline they will turn. Retailing research has a long tradition of being open to new methods and working with other disciplines, most notably economics and psychology. However, the big data revolution may well necessitate new skills (like large data statistics, machine learning, text analytics, audio analytics, video analytics, …) that are less widespread among "traditional" retail researchers. Because of that, there is the danger not only of being leapfrogged by practice, but also of being leapfrogged by other disciplines, such as information management or computer science. Gandomi and Haider (2015) distinguish five stages in typical big data processes, three data-management stages (acquisition and recording; extraction, cleaning and annotation; integration, aggregation and representation) and two analytic stages (modeling and analytics; interpretation). The two analytic stages refer to the derivation of business intelligence from big data, which is where marketing researchers in general, and retailing researchers specifically, should be in a unique position to remain valuable to the retail community. However, rather than seeing this as a competitive game, there may be more value in the collaboration with these other disciplines, not only to identify the managerially most relevant research questions, but also to jointly develop methods that are better and more robust.

Importantly, not only the collaboration with computer and information science may receive an impetus because of the big data revolution, there are also numerous, thus far less explored, joint research opportunities with the operations-management field. Indeed, while much big data research in retailing has focused on how to offer better (e.g., more personalized) solutions to consumers, or how to enhance the in-store shopping experience, a remaining challenge (with potentially large ROI implications) is how to better manage the retailers' back end, and fully streamline their inventory management and supply chain.

6. Big data in retailing: implications for other stakeholders

The big data evolution will not only affect retail managers and researchers, but also several other stakeholders, such as public-policy makers, investors, and retail (marketing) educators.

6.1. Implications for policy makers

The big data evolution in the retail sector is likely to have substantial consumer-welfare implications, bringing the potential role of public-policy makers to the fore. This applies not only to the ongoing privacy discussion, but also to potential implications on retail concentration.

6.1.1. Privacy issues

The big data revolution has led to a heightened focus on consumer privacy. A review on some of the more recent regulation changes in both Europe and the US is available in Palmatier and Martin (2019). As more and more information becomes digitized and movable across organizational boundaries, issues related to privacy, security, intellectual property and liability will become more prominent (McKinsey, 2011). Moreover, given that retailers increasingly operate on an international, and even global, scale, they are often confronted with cross-national regulatory variations (e.g., the European Union Data Protection directive is considerably more stringent than US based regulations, which affects for example what retailers can do with the loyalty-card information they have collected in the various markets that they operate in). Importantly, little is known about how, apart from government regulations, self-regulating practices and consumers' privacy preferences vary cross-nationally (Martin, Borah, & Palmatier, 2017).

6.1.2. Reducing the data divide

The potential involvement of policy makers is not necessarily limited to the privacy debate. They may also play a role in bridging the increasing divide between the haves and the have-nots. If data-driven retailers indeed have a sizeable competitive advantage, the retail landscape may become much more concentrated, with smaller businesses that do not have the means to invest in big data capabilities driven out of business. To avoid such a scenario, some (often local) governments have started experiments where smaller firms are provided inexpensive (or even free) data access and training. Donnelly and Simons (2014a, 2014b), for example, describe how government funding was used to provide seven small supplier firms in the Northern Ireland region with loyalty-card information from supermarket giant Tesco. The question whether, from a consumer-welfare perspective, large data-rich supermarket chains should be stimulated, or even forced, to make some of their data available to (local) suppliers and/or smaller competitors is an intriguing one, and in need of more research.

6.2. Implications for investors

Prior research had established that data-driven decision making (Brynjolfsson, Hitt, & Kim, 2011) and the use of customer analytics (Germann et al., 2014) can positively affect firm (retailer) performance. In a recent cross-industry study, Huang, Wang, and Huang (2018) show how also big data implementations are positively related to improvements in financial performance (like ROA, ROE, profit margins and collection efficiency) and market value (Tobin's Q), even though the obtained effects were not found to be significantly higher for first movers.

However, more work is needed not only to assess whether this also applies to retail firms, but also to identify relevant contingency factors. Are certain implementations more valuable to some retail formats, or in certain retail environments, than others? Moreover, much of the value evaluations appear to be based on (perhaps overly optimistic) expectations of future growth. As discussed in Gielens and Steenkamp (2019), the foreseeable revenue growth (either organically or due to specific big data investments) is much higher for e-commerce retailers than for stationary retailing. Given that more and more retailers move to a hybrid model, large fluctuations in market value can be expected that differ in sign depending on the direction of the business-model change. A better understanding is needed on how big data related business-model adaptations affect the retailer's brand equity, financial performance and firm value.

6.3. Implications for retail educators

A shortage of personnel with the skills necessary to take full advantage of the big data opportunities is a constraint that affects many retail organizations. This shortage applies not only to the more technically-oriented personnel needed to develop and maintain the necessary hardware and software tools, and to personnel with the necessary statistical, econometric and/or machine-learning skills, but also to "data-savvy managers and analysts who have the skills to be effective consumers of big data insights — *i.e. capable of posing the right questions, interpreting and challenging the results, and making appropriate decisions*" (McKinsey, 2011, p. 103, italics added).

In response to this clear shortage (which, admittedly, not only affects retail firms), many universities have recently launched Master of Science in Marketing Analytics, or Master in Data Science, programs. While there is a clear need for such programs, the question becomes whether retail-specific courses receive sufficient attention in those programs, or whether they are dominated by more technical/statistical/computational topics. Put differently, it is currently unclear which skills (computational or retail substantive) should be valued the most, but several of these newly-developed curricula appear to favor the former. A critical issue is how students can best be educated so that they are equally comfortable with algorithms and data analyses as with domain-specific (i.e., retailer-related) insights (see also Boyd & Crawford, 2012). Similar concerns (albeit at the broader marketing-strategy level) have been raised in the context of the course load in doctoral programs (see, for example, Houston, 2016). In addition, the possibilities that business schools can offer their students to be exposed to real-life big data (e.g., through internships) is likely to vary considerably, which can reinforce the divide between well- and less-resourced universities.

7. Conclusion

As argued in the foregoing discussion, the big data revolution is a driving force (antecedent) of many changes, with clear implications, opportunities as well as challenges, for both retail managers and retail researchers. However, it is also important to keep in mind that the big data revolution in itself is largely the consequence (some would even say, a side product) of more fundamental technological and digital revolutions. These have not only resulted in more and better data (the focus of the current essay), but have also transformed the entire retailing value chain in major ways. Using the typology of Eckhardt et al. (2019), they have had a major impact on the very foundations of the retailing domain by affecting (i) its *institutions* (such as consumers, brand manufacturers, and institutional retailers), (ii) its *processes* (e.g., the consumers' shopping experience and engagement), and (iii) the ways of *value creation* and *value appropriation* in the sector. Both Gielens and Steenkamp (2019) and Reinartz et al. (2019) discuss in great detail how digital (dis)intermediation has increasingly blurred the distinction not only been offline and online retailers, but also between manufacturers and retailers, and between suppliers and consumers. Consumers' needs have also evolved because of the digital transformation, in that they increasingly expect both personalized and frictionless shopping (GlobalData, 2018), while different parties not only compete for the initiation of the product purchase, but also try to create and appropriate value during the subsequent product-usage stage (Reinartz et al., 2019). Sorescu et al. (2011), among others, discuss how big data have spurred a wide variety of innovations in Retail Business Models.

While practitioners appear to be ahead of the academic field in their experimentation with all kinds of new business concepts (Dekimpe & Geyskens, 2019), there is a dire need (see also Sorescu, 2017) for systematic academic research on their economic viability (as many of them seem to rely primarily on still very uncertain future growth expectations). To some extent, it may be possible to (try to) do so within the confines of existing theories, relying on mainstream assumptions, and using well-established methods. For example, Gielens and Steenkamp (2019) wonder to what extent the opening of brick-and-mortar stores by pure e-commerce players such as Amazon and Alibaba is just another manifestation of the venerable Wheel of Retailing. On the other hand, one can wonder to what extent some of the new business models that are up-and-coming in practice (like digital platforms, or various manifestations of the sharing economy) should not be seen as "new wine in old caskets", but rather as manifestations of what Yadav (2018) calls novel "emerging phenomena," which require a broadening of the supporting retail theories, with new definitions of constructs, new conceptualizations, and a substantial rethinking of prevailing nomological networks (see also Moorman, van Heerde, Moreau, & Palmatier, 2019). Sorescu et al. (2011, p. S5), for example, argue that one should no longer (or not necessarily) use the traditional characterization of retailers as mere merchant intermediaries that buy from suppliers and sell to customers, but as orchestrators of increasingly complex ecosystems in which value is created and delivered to customers, and, subsequently, appropriated by the retailer and its business partners. They argue how this requires a new look at the interdependencies between *retailing formats* (the structures for sequencing and organizing various retailer activities into coherent processes to fulfill the customer experience), the *retailing activities* themselves (such as acquiring, stocking, displaying and exchanging goods), and *retailing governance mechanisms* (which are needed to motivate all relevant actors to carry out their role in fulfilling the customer experience, such as stimuli for co-creation). Given that all parties increasingly take on multiple roles, the once-predominant (fairly linear) sequence and role division no longer applies in many settings.

Importantly, one of the key impediments to timely research on such emerging phenomena has traditionally been the scarcity of data (Yadav, 2018, p, 363). I am confident that the big data revolution will not only increase the size and scope of more conventional data files, but also accelerate the availability of data on emerging retail phenomena, which should ultimately allow for a more efficient progress of the discipline.

In spite of all the changes brought along, directly or indirectly, through the big data (and related) revolution(s), certain things have remained the same. Foremost, consumer continue to buy the products that they feel satisfy their needs the best, using the retail channel that best meets their needs. Similarly, we can expect the Holy Grail for retailers to remain the same, i.e. "to ensure true continuity of consistent and delightful customer experiences" (Gielens & Gijsbrechts, 2018, p. 9; PlanetRetail, 2017). In a similar vein, also the Holy Grail for retailing researchers should ultimately remain the same, i.e. to develop and disseminate actionable knowledge about real-world questions that are relevant to managers, consumers, policy makers and/or other societal stakeholders.

References

Ailawadi, K. L., Neslin, S. A., & Gedenk, K. (2011). Pursuing the value-conscious consumer: Store brands versus national brand promotions. *Journal of Marketing Research*, *65*(1), 71–89.

Ailawadi, K. L., Pauwels, K., & Steenkamp, J. -B. E. M. (2008). Private-label use and store loyalty. *Journal of Marketing*, *72*(4), 19–30.

Alba, J. W. (2011). In defense of bumbling. *Journal of Consumer Research*, *38*(6), 981–987.

Allenby, G. M., Bradlow, E. T., George, E. I., Liechty, J., & McCulloch, R. E. (2014). Perspectives on Bayesian methods and big data. *Customer Needs and Solutions*, *1*(3), 169–175.

Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. *Wired Magazine*, *23*(June).

Ansari, A., & Li, Y. (2018). Big data analytics. In N. Mizik, & D. Hanssens (Eds.), *Handbook of marketing analytics: Methods and applications in marketing management, public policy, and litigation support* (pp. 280–302).

Arthur, L. (2013). *Big data marketing: Engage your customers more effectively and drive value.* John Wiley & Sons.

Bain & Company (2017). Getting ready to battle grocery's hard discounters. available at https://www.bain.com/insights.

Bass, F. M. (1995). Empirical generalizations and marketing science: A personal view. *Marketing Science*, *14*(3), G6–G19.

BCG (2017). How discounters are remaking the grocery industry. available at https://www.bcg.com/publications/2017, Accessed date: 30 May 2019.

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*(5), 662–679.

ter Braak, A., Dekimpe, M. G., & Geyskens, I. (2013). Retailer private-label margins: The role of supplier and quality-tier differentiation. *Journal of Marketing*, *77*(4), 86–103.

ter Braak, A., Deleersnyder, B., Geyskens, I., & Dekimpe, M. G. (2014). Does private-label production by national-brand manufacturers create discounter goodwill? *International Journal of Research in Marketing*, *30*(4), 343–357.

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.

Bronnenberg, B. J., & Ellickson, P. B. (2015). Adolescence and the path to maturity in global retail. *Journal of Economic Perspectives*, *29*(4), 113–134.

Brown, B. (2014). Can your C-suite handle big data? *Harvard Business Review Insight Center Report*, 12–13.

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). *Strength in numbers; how does data-driven decision making affect firm performance?* (Available at SSRN 1819486).

Cleeren, K., van Heerde, H. J., & Dekimpe, M. G. (2013). Rising from the ashes: How brands and categories can overcome product-harm crises. *Journal of Marketing*, *77*(2), 58–77.

Conover, M. (2011). Jimmy Lin, complexity and social networks blog. available at http://www.iq.harvard.edu/blog/netgov/2011/07/the-international_conference_o.html.

Coveney, P. V., Dougherty, E. R., & Highfield, R. R. (2016). Big data need big theory too. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2080), 20160153.

Datameer (2015). Five big data use cases for retail. ebook available at https://www.datameer.com.

Dekimpe, M. G., & Geyskens, I. (2019). Retailing research in rapidly changing times: On the danger of being leapfrogged by practice. *Journal of Retailing*, *95*(1), 6–9.

Dekimpe, M. G., & Morrison, D. G. (1991). A modeling framework for analyzing retail store durations. *Journal of Retailing*, *67*(1), 68–92.

Donnelly, C., & Simons, G. (2014a). Small businesses need big data too. *Harvard Business Review Insight Center Report*, 35–36.

Donnelly, C., & Simons, G. (2014b). Is there hope for small firms, the have-nots in the world of big data? *Harvard Business Review Insight Center Report*, 45–46.

Eckhardt, G. M., Houston, M. B., Jiang, B., Lamberton, C., Rindfleisch, A., & Zervas, G. (2019). Marketing in the sharing economy. *Journal of Marketing*, *83*(5), 5–27.

Euromonitor (2014). The new face of private label: Global market trends to 2018. available at http://go.euromonitor.com/new-face-of-private-label-global-market-trends-2018-strategy-briefing.html, Accessed date: 17 October 2017.

Ezrachi, A., & Bernitz, U. (2009). *Private labels, brands, and competition policy.* Oxford University Press.

Fan, J., Han, F., & Liu, H. (2014). Challenges of big data analysis. *National Science Review*, *1*(2), 293–314.

Farris, P. W., & Ailawadi, K. L. (1992). Retail power: Monster or mouse? *Journal of Retailing*, *68*(4), 351–369.

Fisher, M., & Raman, A. (2010). *The new science of retailing: How analytics are transforming supply chains and improving performance.* Harvard Business School Press.

Fisher, M., & Raman, A. (2018). Using data and big data in retailing. *Production and Operations Management*, *27*(9), 1665–1669.

Forbes (2015). Big data: A game changers in the retail sector. available at www.forbes.com/sites/bernardmarr.

Fournier, S., & Rietveld, B. (2014). To understand consumer data, think like an anthropologist. *Harvard Business Review Insight Center Report*, *24-25*.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144.

Germann, F., Lilien, G. L., Fiedler, L., & Kraus, M. (2014). Do retailers benefit from deploying customer analytics? *Journal of Retailing*, *90*(4), 587–593.

Geyskens, I. (2018). Retailing power in the grocery industry. In K. Gielens, & E. Gijsbrechts (Eds.), *Handbook of research on retailing* (pp. 399–410). Edward Elgar Publishing.

Gielens, K. (2012). New products: The antidote to private label growth? *Journal of Marketing Research*, *49*(3), 408–423.

Gielens, K., & Gijsbrechts, E. (2018). Introduction to the handbook of research on retailing. In K. Gielens, & E. Gijsbrechts (Eds.), *Handbook of research on retailing* (pp. 1–13). Edward Elgar Publishing.

Gielens, K., Gijsbrechts, E., & Dekimpe, M. G. (2014). Gains and losses of exclusivity in grocery retailing. *International Journal of Research in Marketing*, *31*(3), 239–252.

Gielens, K., & Steenkamp, J. -B. E. M. (2019). Branding in the era of digital (dis)intermediation. *International Journal of Research in Marketing* (Forthcoming).

GlobalData (2018). Frictionless shopping in retail. www.globaldata.com.

Goodstein, R. C. (1994). UPC scanner pricing systems: Are they accurate? *Journal of Marketing*, *58*(2), 20–30.

Grewal, D., Ahlbom, C. -P., & Beitelspacher, L. (2018). In-store mobile phone use and customer shopping behavior: Evidence from the field. *Journal of Marketing*, *82*(4), 102–126.

Grewal, D., Roggeveen, A. L., & Nordfält, J. (2017). The future of retailing. *Journal of Retailing*, *93*(1), 1–6.

Hanssens, D. M. (Ed.). (2015). *Empirical generalizations about marketing impact.* Marketing Science Institute.

van Heerde, H. J., Gijsbrechts, E., & Pauwels, K. (2008). Winners and losers in a major price war. *Journal of Marketing Research*, *45*(5), 499–518.

Heremans, S. (2018). *Market research and big data, presentation at Tilburg University and KU Leuven.*

Houston, M. B. (2016). Is 'strategy' a dirty word? *Journal of the Academy of Marketing Science*, *44*(5), 557–561.

Houston, M. B. (2019). Four facets of rigor. *Journal of the Academy of Marketing Science*, *47*(4), 570–573.

Huang, C. -K., Wang, T., & Huang, T. -Y. (2018). Initial evidence on the impact of big data implementation in firm performance. *Information Systems Frontiers*.

IBM-Analytics (2018). Analytics: The real-world use of big data in retail. available at https://www-935.ibm.com.

IGD (2018). European grocery retail market to be worth €2,289 billion by 2022. available at https://www.igd.com.

Ingram Micro (2018). 6 big data use cases in retail. available at http://www.ingrammicroadvisor.com.
Inmann, J. J., & Nikolova, H. (2017). Shopper-facing retail technology: A retailer adoption decision framework incorporating shopper attitudes and privacy concerns. *Journal of Retailing*, *93*(1), 7–28.
Institute of European and Comparative Law (2015). Trends in retail competition: Private labels, brands and competition policy. *Report on the eleventh annual symposium on competition amongst retailers and suppliers* (pp. 1–44). University of Oxford.
Lamey, L., Deleersnyder, B., Dekimpe, M. G., & Steenkamp, J. -B. E. M. (2007). How business cycles contribute to private-label success: Evidence from the United States and Europe. *Journal of Marketing*, *71*(1), 1–15.
Lamey, L., Deleersnyder, B., Steenkamp, J. B., & Dekimpe, M. G. (2018). New product success in the consumer packaged goods industry: A shopper marketing approach. *International Journal of Research in Marketing*, *35*(3), 432–452.
Larson, J. S., Bradlow, E. T., & Fader, P. S. (2005). An exploratory look at supermarket shopping paths. *International Journal of Research in Marketing*, *22*(4), 395–414.
Laurent, G. (2013). EMAC distinguished marketing scholar 2012: Respect the data! *International Journal of Research in Marketing*, *30*(4), 323–334.
Leeflang, P. S. H., Verhoef, P. C., Dahlström, P., & Freundt, T. (2014). Challenges and solutions for marketing in a digital era. *European Management Journal*, *32*(1), 1–12.
Lehmann, D. R. (2019). The evolving world of research in marketing and the blending of theory and data. *International Journal of Research in Marketing* (Forthcoming).
Lilien, G. L. (2016). The B2B knowledge gap. *International Journal of Research in Marketing*, *33*(3), 543–556.
Longoni, C., Nonezzi, A., & Morewedge, C. K. (2019). Resistance of medical artificial intelligence. *Journal of Consumer Research* (Forthcoming).
Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data privacy: Effects on customer and firm performance. *Journal of Marketing*, *81*(1), 36–58.
McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*(Oct), 1–9.
McAlister, L. (2016). Rigor versus method imperialism. *Journal of the Academy of Marketing Science*, *44*(5), 565–567.
McKinsey (2011). *Big data: The next frontier for innovation, competition, and productivity.* McKinsey Global Institute (Report).
McKinsey (2016). *The age of analytics: Competing in a data-driven world.* McKinsey Global Institute (Report).
Moorman, C. (2016). Celebrating marketing's dirty word. *Journal of the Academy of Marketing Science*, *44*(5), 562–564.
Moorman, C., van Heerde, H. J., Moreau, C. P., & Palmatier, R. W. (2019). Challenging the boundaries of marketing. *Journal of Marketing*, *83*(5), 1–4.
Muller, F. (2014). Big data: Implications and applications in grocery retail. *AMA-ECMI-EMAC marketing & innovation symposium*. Erasmus University Rotterdam.
Ng, S. (2017). *Opportunities and challenges: Lessons from analyzing terabytes of scanner data.* National Bureau of Economic Research (No. w23673).
Palmatier, R. W., & Martin, K. D. (2019). *The intelligent marketer's guide to data privacy: The impact of big data on customer trust.* Springer.
Pauwels, K., Srinivasan, S., & Franses, P. H. (2007). When do price thresholds matter in retail categories? *Marketing Science*, *26*(1), 83–100.
PlanetRetail (2016). *The connected retailer: Empowering the digital consumer.* (PlanetRetail.net Report).
PlanetRetail (2017). *Brick, clicks & marketplaces: A 21st century roadmap for digital transformation.* (PlanetRetail.net Report).
PlanetRetail (2018). Winning with global discount. available at http://www.planetretail.net, Accessed date: 7 January 2019.
Rabobank (2013). Private label goes to Asia. available at https://www.rabobank.com/en/press/search/2013/20130812_Rabobank_Report-Private_Label_Goes_to_Asia_Share_in_India_and_China_expected_to_hit_25-30_by_2030.html, Accessed date: 17 October 2017.
Redman, T. C. (2014). Are you ready for a chief data officer? *Harvard Business Review Insight Center Report*, 14–15.
Reinartz, W., Dellaert, B., Manfred, K., Kumar, V., & Varadarajan, R. (2011). Retailing innovations in a globalizing retail market environment. *Journal of Retailing*, *87*, S53–S66.
Reinartz, W., Wiegand, N., & Imschloss, M. (2019). The impact of digital transformation on the retailing value chain. *International Journal of Research in Marketing* (Forthcoming).
Rooderkerk, R. P., Van Heerde, H. J., & Bijmolt, T. H. A. (2013). Optimizing retail assortments. *Marketing Science*, *32*(5), 699–715.
Rust, R. (2019). The future of marketing. *International Journal of Research in Marketing* (Forthcoming).
Sorescu, A. (2017). Data-driven business model innovation. *Journal of Product Innovation Management*, *34*(5), 691–696.
Sorescu, A., Frambach, R. T., Singh, J., Rangaswamy, A., & Bridges, C. (2011). Innovations in retail business models. *Journal of Retailing*, *87*, S3–S16.
Srinivasan, S., Pauwels, K., Hanssens, D. M., & Dekimpe, M. G. (2004). Do promotions benefit manufacturers, retailers, or both? *Management Science*, *50*(5), 617–629.
Steenkamp, J. -B. E. M., & Geyskens, I. (2014). Manufacturer and retailer strategies to impact store brand share: Global integration, local adaptation, and worldwide learning. *Marketing Science*, *33*(1), 6–26.
Stremersch, S., & Verhoef, P. C. (2005). Globalization of authorship in the marketing discipline: Does it help or hinder the field? *Marketing Science*, *24*(4), 585–594.
Van der Maelen, S., Breugelmans, E., & Cleeren, K. (2018). The clash of the titans: On retailer and manufacturer vulnerability in conflict delistings. *Journal of Marketing*, *81*(1), 118–135.
Varadarajan, R. (2019). Theoretical underpinnings of research in strategic marketing: A commentary. *Journal of the Academy of Marketing Science*, *47*(1), 30–36.
Verhoef, P. C., Kooge, E., & Walk, N. (2016). *Creating value with big data analytics: Making smarter marketing decisions.* Routledge.
Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, *80*(6), 97–121.
Weinberg, B. D., Davis, L., & Berger, P. D. (2013). Perspectives on big data. *Journal of Marketing Analytics*, *1*(4), 187–201.
Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, *14*(4), 490–495.
Wierenga, B. (2011). Managerial decision making in marketing: The next research frontier. *International Journal of Research in Marketing*, *28*(2), 89–101.
Yadav, M. S. (2018). Making emerging phenomena a research priority. *Journal of the Academy of Marketing Science*, *46*(3), 361–365.