

Article

# The Research Topics on E-Grocery: Trends and Existing Gaps

Juan Carlos Martín <sup>1,\*</sup>, Francesca Pagliara <sup>2</sup> and Concepción Román <sup>1</sup>

<sup>1</sup> Department of Applied Economics, Faculty of Economics, University of Las Palmas de Gran Canaria, 35017 Las Palmas de Gran Canaria, Spain; concepcion.roman@ulpgc.es

<sup>2</sup> Department of Civil, Architectural and Environmental Engineering, University of Naples Federico II, 80125 Naples, Italy; fpagliar@unina.it

\* Correspondence: jcarlos.martin@ulpgc.es; Tel.: +34-928-458189

Received: 10 December 2018; Accepted: 4 January 2019; Published: 10 January 2019



**Abstract:** The aim of this paper is to analyse through a systematic literature review the main topics studied on a relatively new concept: e-grocery. This study examines the existing literature through the use of some interactive Scopus applications with the goal to provide: (1) a better understanding of the main topics that have been analysed up to now; (2) the principal methodologies that have been used; and (3) the challenges that still exist as well as literature gaps. First, we provide a summary of what has been previously researched in the field, analysing the trends regarding time, sources, authors, countries and document types. Then, tokenization, and a natural language process (NLP) that applies the Latent Dirichlet Allocation (LDA) method, is used to extract the main topics in the field. We obtain that the research topics in the discipline have always been linked with technological changes, innovation and diffusion, and environmental effects have been a cornerstone since the origin of the field. In comparison with other fields in transport economics, we notice that experimental design and choice modelling are not very popular. As a step towards future research, we propose an integrative framework for e-grocery consumer behaviour that starts from the Engel–Kollat–Blackwell (EKB) model.

**Keywords:** E-grocery; Systematic literature review; Natural Language Process; Latent Dirichlet Allocation; Engel–Kollat–Blackwell (EKB) model

## 1. Introduction

Humans need to eat in order to survive, so eating is an essential activity that is carried out several times per day [1]. Thus, it is not strange that research on e-grocery has recently received more attention in many disciplines [2,3]. E-grocery is now gaining momentum because it promises relative convenience to the customer, given that customers can obtain the groceries they need from the comfort of their homes or offices and at a convenient time. Mkansi et al. [3] find that, contrary to expectations, the correlation between e-grocery and buying in stores is not perfect for the 10 biggest markets. However, despite the growing interest in e-grocery research, there is a need for an analysis that shows how the research has evolved in the past. Thus, the aim of this paper is to present a bibliometric analysis of e-grocery/e-retailing that examines the main topics and methods and offers a future research agenda.

A bibliometric analysis determines the academic–intellectual structure and knowledge basis for a predetermined field [4–6]. The analysis can also be used to find the most prolific institutions, scholars, geographical regions and journals, among other things [7,8]. This is the first bibliometric analysis of e-grocery/e-retailing according to our knowledge. Thus, it is also necessary to evaluate the data, the research topics and the methods, highlighting especially those studies that use choice modelling.

Thus, the aim of this paper is threefold: (1) to identify the evolution of e-grocery/e-retailing research with a leading focus on authors, institutions and geographical areas; (2) to determine the main studied topics by applying the Latent Dirichlet Allocation (LDA) method [9,10]; and (3) to analyse the main methods used in the field to extract important insights into those papers that use choice modelling and to suggest a future research agenda.

Having this in mind, a bibliometric analysis is used as the main quantitative approach to address these issues. Pritchard [11] introduces the bibliometric analysis as an effective method that provides a synthesis of the scientific production and research agenda for some specific field. There are a number of databases that can be used for the analysis, being the Science Citation Index (SCI) from the Web of Science and Scopus two of the most popular. Different academic fields have been analysed using this approach, covering different features, such as yearly trends, authors, institutions, journals, categories and countries.

The results provide important insights to both academics and practitioners regarding the existing and future research in e-grocery. The originality of the study resides in its being the first application of a systematic literature review that has been performed with the help of the LDA method as a way to extract the main topics that have been analysed by the literature in the field of e-grocery. An important contribution of the study is to provide an extensive and detailed characterisation of the studied topics. Our study concludes that experimental design and choice modelling needs to be better integrated into different existing theories from transport, economics, marketing, the environment and psychology.

The rest of the paper is organised as follows. First, a brief literature review on e-grocery/e-retailing is provided. Second, the bibliometric analysis's methods are detailed and explained. Next, the findings are presented and discussed. Finally, important insights and conclusions are obtained to provide a tentative future research agenda.

## 2. Literature Review

Kolesar and Galbraith [12] contend that e-retail encompasses three main activities: (1) a product search activity that provides detailed information on the products under evaluation, which is usually referred to as a product-evaluation or information-gathering (IG) facility; (2) an online purchase function that facilitates consumer interaction by reducing the transaction costs; and (3) a product delivery capability that facilitates the final product's distribution to consumers.

Darley et al. [13] present an overall review to understand to what extent the current marketing and consumer behaviour body of literature can be transferred to the analysis of online consumer behaviour and preferences. The paper uses the model proposed by Engel, Kollat and Blackwell [14,15] as the analytical framework to synthesise the findings from the literature.

The Engel–Kollat–Blackwell (EKB) model proposes five core stages of the decision-making process, as follows: (1) problem recognition; (2) search; (3) alternative evaluation purchase; (4) choice; and (5) outcomes.

The authors contend that the underlying alternative evaluation depends on internal and external factors. The internal factors are categorised into three different classes: (1) cognitive or beliefs; (2) affective or attitudes; and (3) cognitive or intentions. On the other hand, the external or environmental factors are subdivided into four categories: (1) individual differences or characteristics such as motives, values, lifestyle and personality; (2) socio-cultural factors, such as culture, social class, reference groups and family; (3) situational and economic factors; and (4) online attributes, such as a Web site's quality, a Web site's interface, Web site satisfaction and Web site experience (p. 95).

The model is finally extended with the consequences of the decision-making process, in which the relevance finally focuses on the satisfaction/dissatisfaction construct. The inherent model recognises the complexity and the multidimensionality of online consumer behaviour, and that—besides being a comprehensive approach—other interactions, antecedents and consequences can be finally envisaged and included. In this respect, the authors highlight that empirically tested constructs and relationships are preferred to conceptual approaches. Two key constructs are found: (1) personal satisfaction and

loyalty; and (2) trust, security and company reputation. These two constructs are clearly related to other three core constructs that have gained significant attention from both scholars and practitioners: purchase, repurchase and product return [16]. Returning products after purchase is important in terms of the volume of the returned goods and the costs of reverse logistics [17].

The authors perform a content analysis of 52 studies, and find that the most common research method is the survey. They find very disappointing the fact that only 31 per cent of the research is based on an experimental design. They criticise the survey design by the limitations imposed by the researchers' bounded rationality, and by the use of student cohorts (46%) as the population of study. There are some areas that will be important in the near future, such as online auctions, bidding behaviour, environmental effects and jobs substitution effects. The authors conclude that the research agenda could benefit from exploring the differences in consumers' preferences between brick-and-mortar and online outlets (brick to click consumer behaviour). In this respect, online consumers tend to be very heterogeneous in different contexts. For example, Rohm and Swaminathan [18] find four stereotypes of e-grocery grocery shoppers: (1) convenience shoppers; (2) variety seekers; (3) balanced buyers; and (4) store-oriented shoppers.

Nguyen et al. [16] emphasise that the smartphone era has created a lot of opportunities for retailers to increase online sales. Firms face multiple challenges in order to facilitate online consumer orders, such as on-time and efficient transport delivery, accurate inventory management and efficient warehouses [19,20]. Online order fulfillment (e-fulfillment) is a critical factor that affects online consumer behaviour, especially brand loyalty [21]. From the firms' perspective, it is important to understand which features of e-fulfillment service provision affect consumer behaviour in order to develop successful e-retailing strategies that improve the supply chain management (SCM). SCM is mainly based on three main pillars: (1) management of the flow from raw materials to final products that are delivered to consumers; (2) being strategically oriented towards cooperation between multiple agents responsible for managing this flow; and (3) being consumer-oriented [22].

Based on a meta-synthesis approach, three e-fulfillment topics have been found: (1) inventory management; (2) last-mile delivery; and (3) returns management. Regarding the first topic, Jing and Lewis [23] find that stock-outs in e-groceries do not reduce the probability of a purchase in the short run in that there are enough new consumers because they are more tolerant to this situation than frequent consumers. Last-mile delivery (LMD) is considered to be the final trip of the SCM, in which products are delivered to consumers. There are three main last-mile delivery methods: (1) attended home delivery; (2) unattended delivery; and (3) manned or unmanned pickup points. The last-mile delivery strategy employed by e-groceries affects the e-grocery service quality [24]. Finally, the returns volume depends on the returns policies, and Bower and Maxham [25] show that free returns increase the return volume of goods, while fee-based returns decrease the volume. Nevertheless, the same policies point into the opposite direction with respect to repurchase and consumption [26]. Similarly, Thirumalai and Sinha [27] find that e-fulfillment is also a moderator of consumer satisfaction, having more impact in e-groceries than in e-apparel and e-electronics. The authors make evident that e-fulfillment depends on consumer expectations, and these vary with respect to product type, being higher for electronics than for groceries. The authors conclude that e-fulfillment strategies should be adapted to take into account the different product types.

LMD is becoming a critical tool for market differentiation, and retailers are innovating with a myriad of delivery options, such as buy-online-pickup-in-store, autonomous delivery solutions, lockers and free delivery upon minimum purchase levels [28]. LMD needs to offer more convenience and flexibility to consumers. Pan et al. [3] find that same-day and on-demand delivery services are gaining traction for e-groceries and highlight that a high rate of failed deliveries significantly increases the logistic costs, especially for perishable food. Lim et al. [29] analyse the key driver variables that characterise different LMD models. They exclude some papers that deal with some specific issues, such as public policy, urban traffic regulations, logistics infrastructure, urban sustainability and environment, from their systematic literature review.

Nguyen et al. [16] find that half of the studies on e-retailing do not take into account the product category in the analysis, and that groceries receive the most attention when different goods are under analysis. It has already been noted that the type of product has an effect on e-fulfillment; however, some authors find that this is not the case for repurchase intentions [30]. The authors find that repurchase does not depend on the types of products under analysis, such as books, clothing and shoes and electronic equipment.

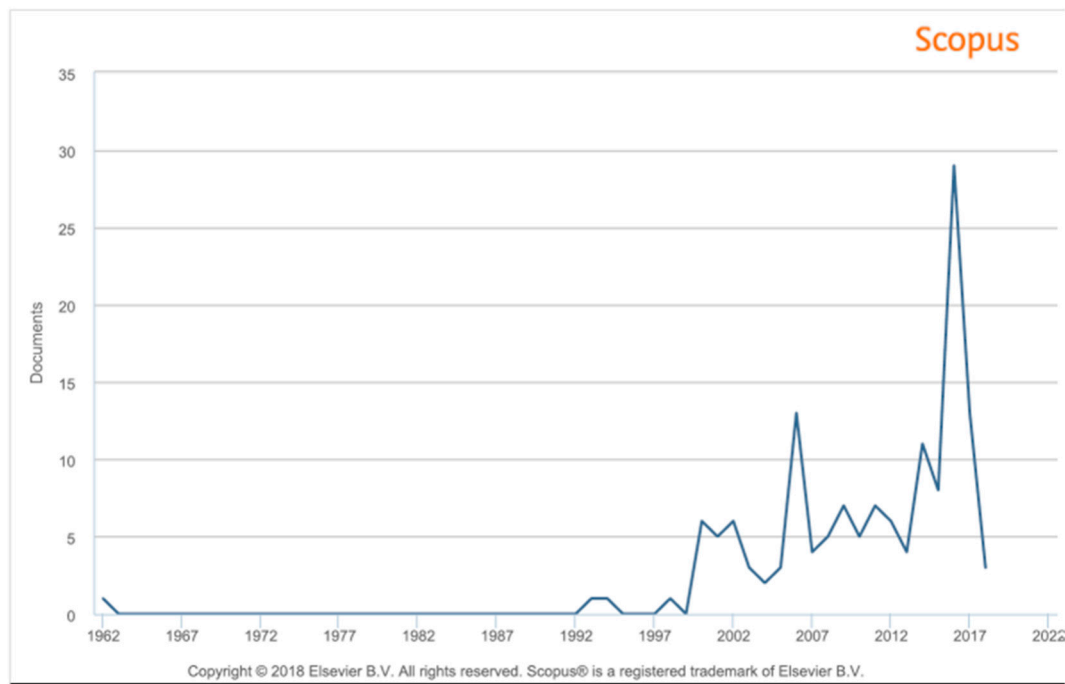
### 3. Bibliometric Analysis

To consolidate the state of the art of e-grocery/e-retailing research, a database is extracted from Scopus, where each observation contains some structured information, and it is based on a published Scopus document. The following main fields were included for each observation: author, country, title, year of publication, source, abstract and document type. The data extracted from Scopus include a total of 144 observations that contain the terms e\*grocer\* or e\*retail\* in the abstract, the title or the keywords. The Scopus database is downloaded using the csv export format so it can be analysed using specialised bibliometric software. The Scopus application program interface (API) allows researchers to conduct various bibliometric analyses to analyse some fields. All of the observations have been published before 14 April 2018, and only sources in English were considered in the analysis.

Figure 1 presents the annual number of published documents from the extracted database, and it can be seen that research starts in 1962. Kirchner [31] investigates possible faking of the Edwards Personal Preference Schedule in personal selection (EPPS). EPPS is a well-known personality scale. The author compares the measures obtained for sale applicants (retail and industrial) and already-employed salesmen in retail and industrial firms. The results show that retail applicants tend to score significantly higher on Orderliness, Intraception, and Dominance scales and lower on the Heterosexuality scale than Retail salesmen. No significant differences are found, however, between Industrial applicants and salesmen.

Online consumer behaviour goes hand-in-hand with the history of the internet. It is true that the first steps of the internet's development as we know it today were developed as basic ideas in the 1950s. The U.S. Department of Defense's Advanced Research Projects Agency has been a key player in such development. The program grew in the 1960s and 1970s, becoming a network of computers that transmitted information by "packet switching". The network of computers was, from the start, an open, diffused and multi-platform network that, since the 1990s, has been expanded globally from the United States, creating an impressive innovation disruption that is still not finished [32].

Figure 1 shows an increasing trend for the total number of documents published since 2000 until the peak obtained for the year 2017. Surprisingly, for the period between 1962 and 1999, there are only four documents in the database. A closed examination of the abstracts shows that these documents should be discarded from the database. The abstracts of the documents contain, e.g., retail as an explanation for some of the issues analysed in the paper; however, the topic does not really belong to e-retailing or e-groceries. Thus, in our database, it is not until the year 2000, when Kolesar and Galbraith contended that there has been an explosion in the number of retail Web sites since 1995. The authors argue that e-retail supply shares a few common elements, specifically: (1) a product search facility; (2) an online purchase function; and (3) a product delivery capability. The authors conclude that the Internet is a poor service delivery medium, as it lacks the capacity for direct personal interaction that brick-and-mortar retailers provide.



**Figure 1.** The total number of documents.

The abstracts were then pre-processed using two basic procedures: tokenization and stop words removal. Tokenization is a lexical analysis tool that splits the abstract into words, phrases and other meaningful elements called tokens. In this study, each abstract was finally broken up into a vector of basic tokens. Stop words are words that do not contribute to the meanings of the text and are usually filtered out before a natural language process (NLP) tool is applied.

Since the vector space represented by all the linguistic tokens is not meaningful, we used an NLP tool to extract the main topics by applying the Latent Dirichlet Allocation (LDA) method [9,10]. The LDA method uses Bayesian statistics to analyse the co-occurrence of existing tokens in the different documents to infer an apparent hidden structure. Mathematically, a set of training documents is used to obtain two main outputs: (1) a list of topics associated with the research on e-grocery/retailing and the respective weights; and (2) the posterior distribution of weight values that gives the probability that each of the documents contains some specific topic. Thus, it is possible to analyse the distribution of the papers according to a selection of topics and a further systematic literature review can be performed using the whole document for a group of documents. Thus, the evolution of, and changes in, a particular topic can be analysed.

#### 4. Results

Figure 2 shows the most productive journals together with the distribution and the dynamic dimension of the publications. There are 108 different outlets where the documents have been published, indicating that the dispersion of the documents is really high. This is also shown in the figure, as there are only five publications with three or more documents in the analysed period: the Business Process Management Journal (3); the International Journal of Applied Engineering Research (4); the International Journal of Retail and Distribution Management (5); the International Journal of Retail & Distribution Management (6); and Lecture Notes in Computer Science (including the subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (4). It can be seen that the four documents published in the International Journal of Applied Engineering Research were all published in 2016. A journal category analysis shows that the number of documents published in Business, Computer Science and Engineering is higher than that in the rest of the categories (Figure 3).

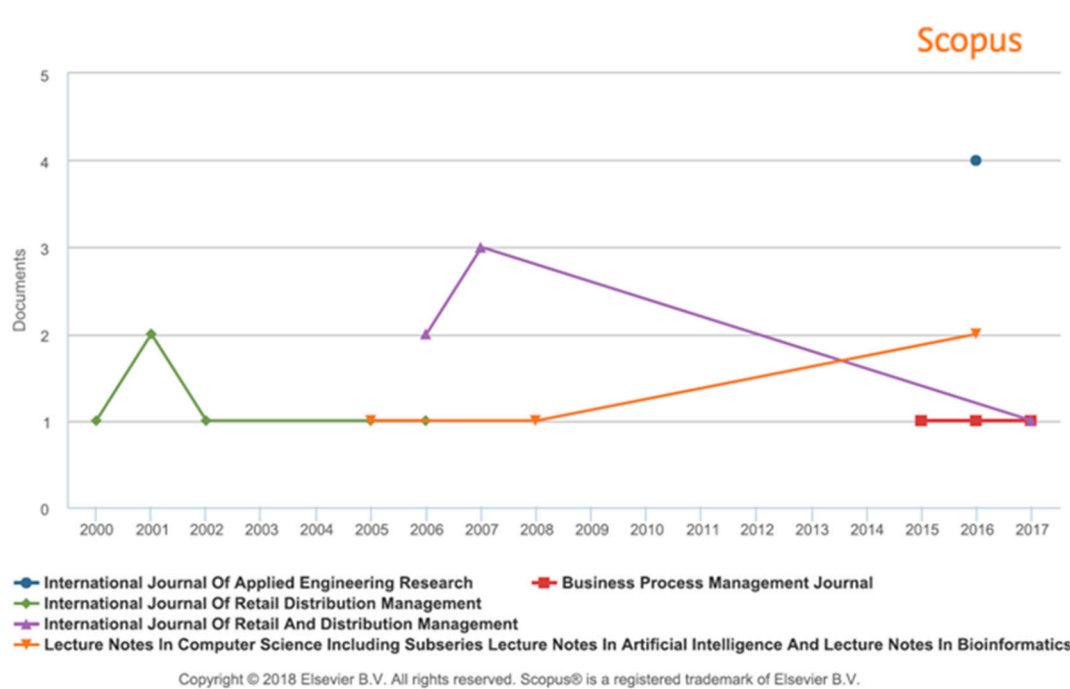


Figure 2. The most productive journals.

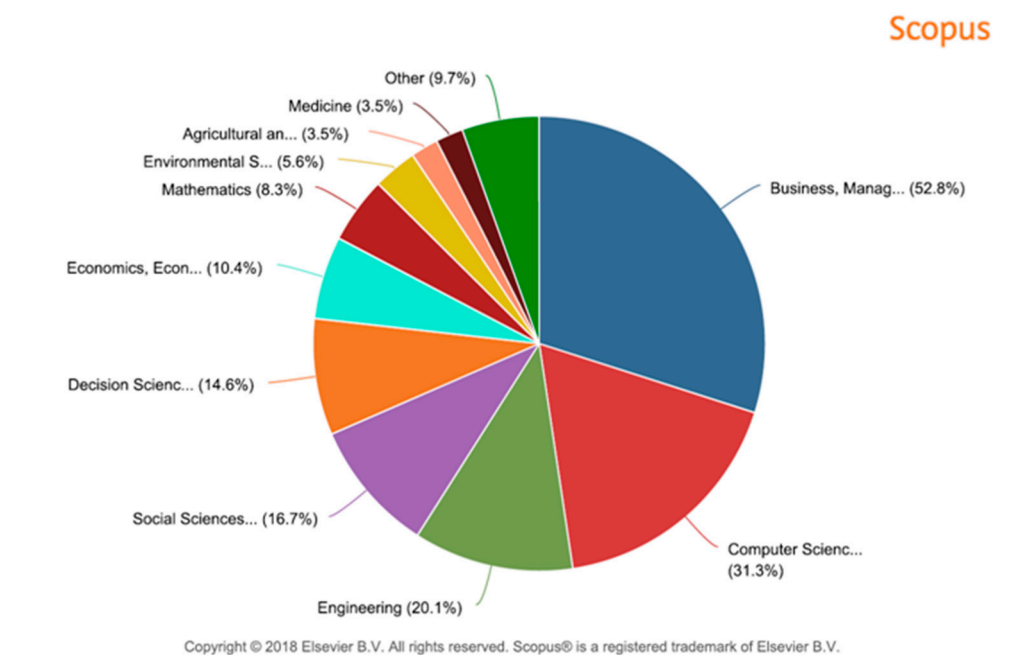
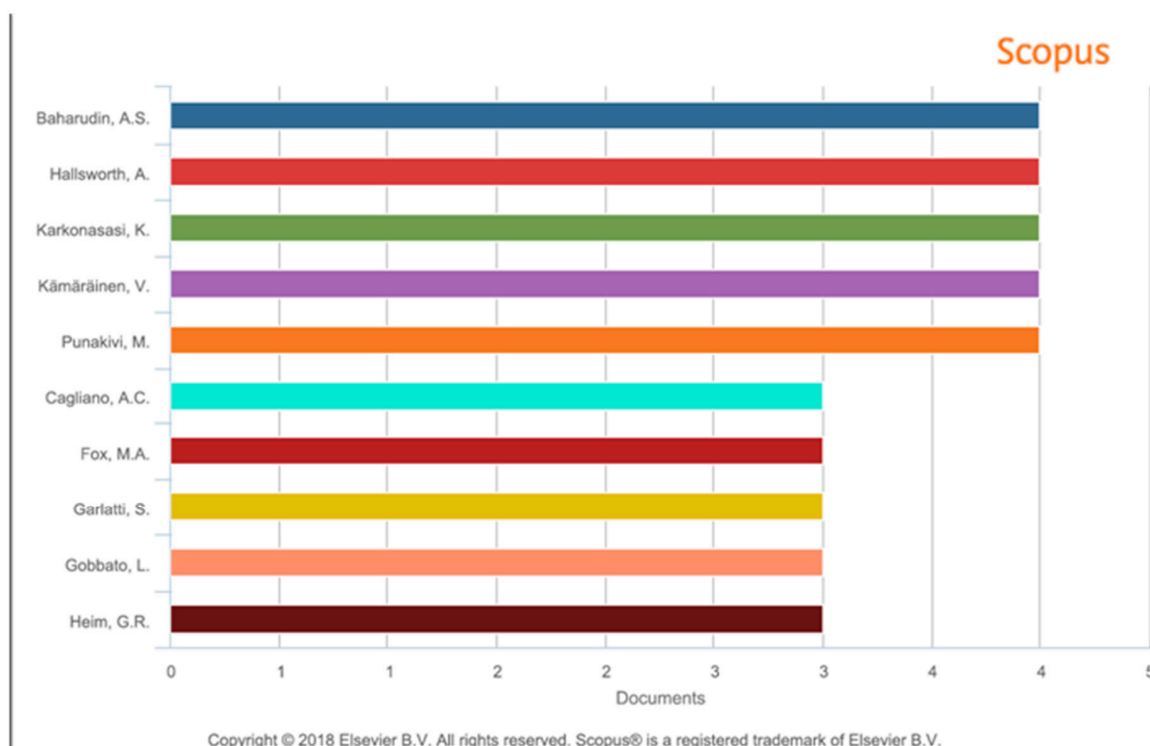


Figure 3. The category analysis.

Figure 4 analyses the most prolific authors in the field. Of the published articles, it can be seen that there are two groups of authors who have published four and three documents, respectively. The first group is formed by the following authors: Baharudin, Hallsworth, Karkonasasi, Kämäräinen and Punakivi. On the other hand, collaboration was limited to two, three or four authors per paper. Papers with five or more authors collaborating do not exist. The pattern of collaboration in publication is presented in four different groups: (1) a single author from a single institution and country: 15 papers that represent 10 per cent of the total; (2) two authors: 49 papers that represent 35 percent of the total and play a determinant role, indicating that this type of collaboration is still the most favourable way to



conduct research related to the field under study; (3) three authors: 34 papers that represent 24 percent of the documents; and (4) four authors: 37 papers that represent 26 percent of the total. There are also five documents in which the authors were not available.



**Figure 4.** The most prolific authors.

The e-groceries/retailing research from various universities and institutions around the world is presented in Figure 5. There are some universities that teach transport, engineering, business and economic programs that favoured the publication of articles during the period under analysis, and it is clear that the pattern of domination in publication of these universities, as indicated by the published research, is highly correlated with the most prolific authors. The most prolific institution was Aalto University, which published eight articles, followed by University Sains Malasia (4), and the University of Surrey (4).

Figure 6 presents the most prolific countries by the number of publications related to e-groceries/retailing that appear in Scopus. Being one of the major industrialised nations globally, it is not strange that the U.S. took the first place with 32 publications, followed by the U.K. (24), India (19) and China (14). Although the publication pattern of European universities is high, it can be observed that there some impotent gaps that appear, especially in Germany and Spain.

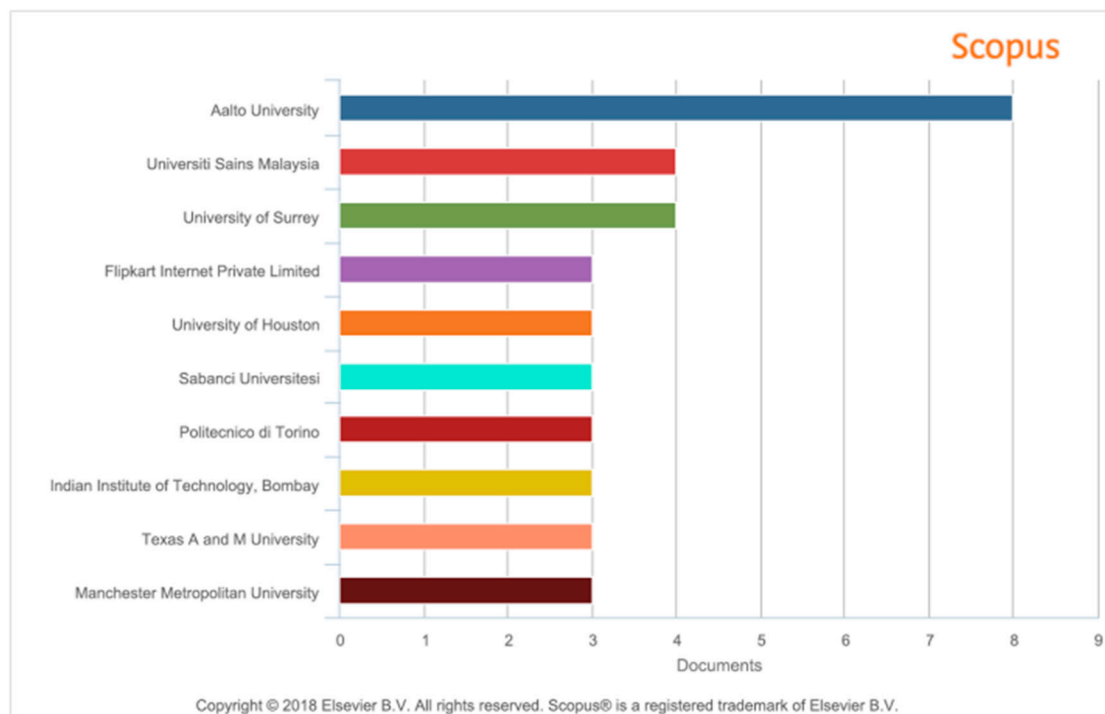


Figure 5. The most prolific universities.

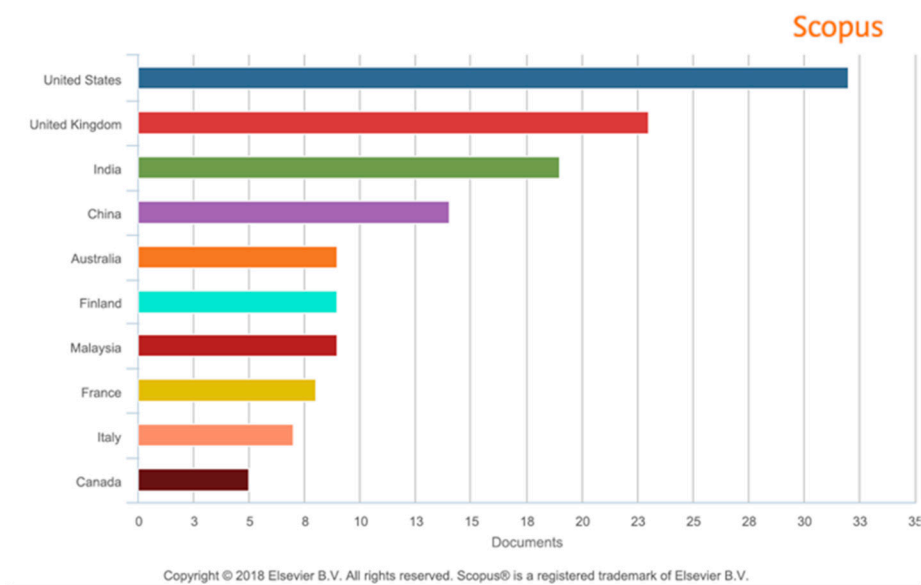


Figure 6. The most prolific countries.

a. The Main Topics of the Documents

Up to now, the analysis has been based on the Scopus api, and now a content analysis of the abstracts will help us to further extract the main topics of e-grocery/retail research. LDA uses an unsupervised Bayesian learning algorithm that effectively captures context-specific topics without making any assumption about the distribution of the words used in the abstracts, and where the grammatical attributes of language are considered to split the words or tokens so that some basic articles or prepositions are not extracted as main keywords in the topic. Thus, LDA can complete many steps of the textual analysis with little human intervention, even labelling dimensions, and is becoming a very powerful tool for dealing with a large amount of unstructured information. Thus,



20 topics were extracted from the abstracts of the 140 documents under analysis, which enables us to perform a more reliable analysis of the method presented in the second section.

This research utilises a recent tool, developed in Python, that is based on the fields of machine learning and natural language processing to effectively extract the topics of e-groceries/retailing research obtained from a large corpus of text data formed by the abstracts extracted from Scopus. A topic model can be defined as a type of probability model for discovering the abstract “topics” that occur in a collection of documents. LDA is the most common method for topic extraction and is a generalisation of probabilistic latent semantic indexing (PLSI) [9,33]. Briefly, LDA assumes that the words in each abstract are independently drawn from a mixture of documents. Each document contains words taken from the vocabulary, a topic-word distribution. Topics can potentially be shared by all of the documents used in the analysis. Thus, every abstract will have its own mixing proportion of topics. LDA is then used to discover the underlying topics from massive volumes of unstructured text data. The topics then can be further explored to analyse the previous research regarding the five stages of the EKB model, the internal factors or the external dimensions. It is also possible to analyse other important constructs, such as satisfaction, or other consequences, such as loyalty or word-of-mouth.

LDA extracts the topics of e-groceries/retails research, the importance of the topics, and the keywords included in each topic based on the pre-processed abstracts dataset. A topic is considered as a latent construct distributed over a vocabulary of words that researchers use to describe the contents of the document [34]. The abstract is formed by a sequence of  $N$  words, which is referred to as a “document” in the literature,  $w = (w_1, w_2, \dots, w_N)$ , whilst  $M$  documents form a corpus,  $D = \{w_1, w_2, \dots, w_M\}$ . Then, researchers assume that there are  $T$  number of topics across the corpus, comprising all the  $M$  documents. LDA generates a probabilistic model of a corpus, and documents are represented as random mixtures over  $T$  latent topics, where each topic is characterised by a distribution of words. Blei [35] contends that LDA is one of the most powerful tools to select topics. The topics, the keywords, and the respective distributions are obtained through Gibbs sampling and the maximum likelihood estimate method [10,36]. Thus, LDA extracts the final topics that researchers have studied and the probability that each topic occurs in each document.

We apply LDA to extract 20 topics of e-groceries/retailing research using the Scopus abstracts as the corpus. Each topic is formed by 10 words, in which the relative weight is decreasing in relation to the order of the keywords shown in Table 1. It is outside the scope of the current paper, but the first purpose of the Table usually consists in naming each of the topics. Thus, by way of example, a tentative name for the first three topics would be “innovation”, “diffusion” and “customer service”. Naming is normally based on the first words that appear in the list, as these have relatively more weight in each of the topics.

Table 1 is also used to form a dictionary that can be further used to perform a network cluster analysis. The keywords of the topics are to some extent related to the main variables of the conceptual model EKB and its extensions. In particular, some keywords are related to consumers’ perceptions and evaluations, such as emotions, and attitudes. Other keywords are related to external factors, such as women, work, family, web, tax and incentives. Other variables are more related to the consequences of the outcomes, such as word of mouth and satisfaction. Finally, some keywords can be used to extend the developed EKB model, such as innovation and diffusion. Innovative processes are usually copied when these determine qualified standards, implying that consumers perceive some alternative experience to be superior, which gives those firms a competitive advantage. On the other hand, it is recognised that there is still an important gap regarding the second conceptual model that deals with e-fulfillment and the obtained constructs: (1) inventory management; (2) last-mile delivery; and (3) returns management.

**Table 1.** The main topics of the documents.

Topic	Keywords
1	scale; innovativeness; companies; easy; adoption; discounts; players; retail; self; penetration
2	rapid; diffusion; results; logistics; proposes; grocery; consumers; competitive; satisfaction; consumer
3	service; grocery; customer; services; retail; new; business; research; virtual; based
4	grocery; delivery; home; service; business; concepts; online; shopping; costs; reception
5	service; quality; research; american; intentions; behavior; evolution; local; reducing; establish
6	wom; years; dimensions; want; website; understand; costs; developed; usa; effectiveness
7	preferences; work; incentives; vs; management; contingency; schedules; working; family; sample
8	efficacy; tax; consists; intention; self; single; novel; solution; significant
9	advertising; service; profit; retailers; make; price; retailer; online; strategy; effect
10	consequences; learning; online; market; consumer; grocers; services; publications; levels; measures
11	retail; industrial; sales; learning; study; situation; significant; model; results; reserved
12	benchmarking; commerce; chain; market; industry; supply; existing; new; firm; growth
13	local; model; conditions; competitive; research; used; retail; women; sales; undergone
14	personalisation; questionnaire; used; assess; sectors; successful; preferences; large; 2006; shows
15	prices; methods; model; task; performance; different; form; database; seeks; demand
16	mental; spending; smart; retail; health; illness; model; married; individuals; traffic
17	loyalty; wom; data; retail; research; using; website; sensory; web; support
18	quality; information; customer; website; retail; satisfaction; service; attitude; applications; identifies
19	retail; model; retailing; agents; results; like; samples; food; study; different
20	features; emotions; used; fish; emotion; study; online; shopping; recognition; effective

#### b. Existing Gaps in E-Groceries/Retailing Research

In this section, a number of existing gaps are commented on according to the following topics that have not been sufficiently studied: (1) taxes; (2) comparative; (3) smart; (4) preferences; and (5) comparative analysis.

Regarding the role of taxes, Ellison and Ellison [37] examined the sales of four different types of memory modules: 128 MB PC100, 128 MB PC133, 256 MB PC100 and 256 MB PC133.9. Price data were obtained by downloading the first (or first and second) screens from Pricewatch's memory price lists on an hourly basis. Pricewatch represents a price search engine where consumers can choose product categories and be given a list of participating e-retailers that sell products in that category sorted by the price. Data on memory modules sales were used to explore aspects of e-retail demand with the objective of examining how the time-invariant variables, i.e., sales tax rates and shipping times, had an impact on consumer demand. Measures of the total number of orders received from each state were built, and regressions (made up of 51 observations) were proposed to look at whether e-retailers sold more in states with high sales taxes than in states with low sales taxes. Discrete choice techniques were used to examine the hourly sales data. A strong relationship was highlighted between e-retail sales to a given state and sales tax rates applying to purchases from offline retailers. This result suggested a substantial online offline substitution and the importance of tax avoidance motives. Another interesting result was that consumers preferred purchasing from firms in nearby states and had a separate preference for buying from in-state firms.

In the paper by van Loon et al. [38], it was shown that online retailing could lower the environmental impacts of shopping under given circumstances. Indeed, the study encouraged consumers to reduce complementary shopping trips and maximise the number of items per delivery. A framework was developed to take into consideration all of the relevant environmental factors relating to retail/ecommerce activities. For example, variables dealing with consumer shopping behaviour, such as basket size, transport mode, trip length and trip frequency, were considered. This framework was created with the aim of building a Life Cycle Analysis model, which was applied to different online retail methods for fast-moving consumer goods in the U.K. The main results were that, within the "last mile" link to the home, the nature of the consumer's behaviour in terms of travel, choice of e-fulfillment method and basket size were critical factors in determining the environmental

sustainability of e-commerce. The nature and routing of van deliveries, the amount and type of packaging used, and the energy efficiency of shop and e-fulfillment centre operations were also contributors to climate change potential.

Regarding smart applications, Othman et al. [39] developed a prototype for an e-groceries system based on an android platform suitable for most mobile devices. A new way of grocery shopping was introduced for Malaysians, and the new system was named MyMart e-Groceries. Users could select many types of groceries in a simple mobile user interface and paid for the selected choice, which would be subsequently delivered to their door steps. According to the system, the administrator would first log in, and then navigate between the menu options manage inventory, manage user accounts, view orders and view transaction history. The administrator could manage products in the inventory by adding, deleting and updating the list of products. Moreover, he/she could upload a promotional brochure through the system and view all of the orders' details. On the other hand, a client was to register before using the application, and then he/she could log in using the registered username and password. The client was able to view a product's catalogue as well as its expiration date and the ingredients of the product as well. The client could add products in the desired quantity to a shopping cart and view the products in the shopping cart. Clients were able to choose the method of payment, and were able to proceed to check out and complete their transaction. This system was similar to web-based grocery shopping using Radio Frequency Identification (RFID) and Near Field Communication (NFC).

In terms of preferences, Rottem-Mindali and Salomon [40] asserted that the proportion of virtual shopping was still significantly smaller than that of traditional shopping. An understanding of this difference could be obtained by modelling consumers' choices. Different shopping activities were examined by modelling both the purchase and delivery choices. Data were collected in the Tel Aviv metropolitan area and included 510 valid questionnaires. Delivery services could act as a major limiting factor for the purchase of products using online stores. Excluding large electric appliances, the need for delivery by others might reduce significantly the probability of purchasing products at online stores. In general, the delivery mode preferred by consumers was self-delivery. This preference was found to be the same for all of the product classes. It is interesting to highlight that, for perishable products, a long distance to the store reduced the probability of using delivery by others, while for large electric appliances the probability increased. The location of the IG was identified as extremely important when deciding on the purchase mode. The more complex a product was, the more important the IG became for the consumer.

Finally, a comparative analysis was carried out by Harris et al. [41]. The authors developed a typology of grocery shoppers on the basis of the advantages and disadvantages of shopping both online and in store for consumers buying groceries using both channels. A total of 871 U.K. shoppers, purchasing groceries online and offline, represented the sample. The survey was represented by items measuring the perceived advantages and disadvantages of grocery shopping online, and items related to the perceived advantages and disadvantages of grocery shopping in traditional supermarkets. The advantages of online shopping are the convenience, the quickness, the possibility of trying new products and the modernity. The disadvantages of online shopping are that deliveries can be late or may not arrive, products can be missing from an order and internet shopping is not secure, too slow or even too complicated. The advantages of supermarket grocery shopping include the possibility of obtaining better prices or new ideas. The disadvantages are that shopping at a supermarket can be time-consuming, and supermarkets can be crowded or too big. A cluster analysis was chosen to develop typologies of online and offline grocery shoppers. The main results showed that the choice of whether to shop online or in store may be driven not by the perceived advantages of one channel versus the other, but by the desire to avoid the greater disadvantages of the alternative. These perceptions differ between consumer groups.

## 5. Conclusions and Discussion

This article provided a critical overview of how e-groceries/retails research has evolved in the last 20 years. The study analysed 144 Scopus documents, and used some keywords as the extraction strategy. The bibliometric analysis has addressed dynamically the trends in the number of publications; the most prolific authors, institutions and countries; and the topics of the research. Three categories were found to be more relevant than others: (1) Business; (2) Computer Science; and (3) Engineering. As one of the first studies in the area, the research findings offer important insights that can be taken into account in further studies.

According to the findings of the study, the number of articles on e-groceries/retailing has increased yearly from 2000 to 2017 (Figure 1). Among the analysed documents, it was seen that the journals with more publications were: the Business Process Management Journal (3); the International Journal of Applied Engineering Research (4); the International Journal of Retail and Distribution Management (5); the International Journal of Retail & Distribution Management (6); and Lecture Notes in Computer Science (including the subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (4) (Figure 2). Although the number of research topics in the discipline have greatly increased along with technological changes, innovation and diffusion, environmental effects have been a cornerstone since the discipline's origin. The emissions depend on the type of fuel and the distance driven by the vans in the home delivery of goods [42,43]. The net effect depends on the degree of substitution between the traditional shopping trips and e-shopping. Cairns [44] contends that vehicle km can indeed be reduced by up to 70% if there is a perfect substitution. However, Mokhtarian [45] argued that the complete substitution is not likely, as groceries form only a part of the goods bought. Van Loon et al. [38] concluded that it is not easy to calculate the real environmental impact of e-commerce, as there are multiple interactions that need to be addressed, such as shopping frequency, transport modes and relative distances between shops and consumers.

Many undergraduate and graduate-level programs have a connection with e-groceries/retailing research; for example, marketing, computer science and logistics. Marketing is indebted to many disciplines, such as sociology, psychology or neuroscience, for concepts, data and methods. Marketing is related to consumer behaviour. Marketing in e-groceries mainly deals with issues that help researchers to understand consumer shopping behaviour. In this respect, many studies compare the advantages and disadvantages between e-commerce and physical retail [40]. Computer science is also important, since the internet, tablets and smartphones reduce the transaction costs between customers and producers. Introducing e-groceries on mobile platforms would make the market more flexible than traditional channels [39]. Finally, logistics is related to one of the main key drivers that make e-groceries more competitive, such as convenient and fast delivery. Verhoef and Langerak [46] noted that inconveniences due to missing deliveries or waiting too much for deliveries can offset the perceived advantages of e-grocery shopping in comparison to traditional in-store shopping. Thus, e-groceries research may need an adequate scientific background that is derived from a number of disciplines. An experimental design can also produce more cross-collaboration opportunities between institutes and researchers.

By chance, it was concluded that, after the initial extraction step, the selected documents need to be manually scanned by at least two authors in order to minimise the possible errors associated with machine learning processes. Thus, a manual scanning procedure that uses at least the abstract needs to be developed in order to facilitate a consensual agreement on the extracted articles. Thus, the validity and the reliability of the database can be better ensured.

This study inspires several areas of future research based on gaps identified in our study. We notice that experimental design and choice modelling is practically nonexistent in the field. As a step towards future research, we propose an integrative framework for e-grocery consumer behaviour that starts from the EKB model. Future research needs to integrate and to extend the outputs with different consequents, taking into account that different existing theories from transport, economics, marketing, the environment and psychology need to be integrated.

In this respect, there is also a number of models that can be developed in the field. Examples are: (1) service engineering (SE), in which sustainable production and consumption should be highly prioritised [47]; (2) product service systems (PSS), a methodology proposed by Fargnoli et al. [48], which is mainly based on the analysis of the needs of customers and considers the quality deployment (QFD) method and the simulation of life-cycle scenarios; (3) the Functional Engineered Product–Service System (FEPSS) proposed by Harber and Fargnoli [49], where a unified service structure is designed following a multistep procedure that contemplates different stakeholders’ needs, activities and other elements; and (4) social norms and beliefs (SNB), which can be affected by the role of social media. Firms pilot different approaches to promote products and services that affect not only purchase behaviour but prescription and recommendation, which ultimately mold Utilitarian (Convenience and Product Selection) and Hedonic (Idea and Adventure) motivations [50,51].

We acknowledge that our research has limitations. First, our systematic review is only focused on e-groceries/e-retailing, and does not include any additional topics or dimensions, so it has mainly an exploratory nature. Another important limitation is that the study’s results originate from a single database (Scopus). For this reason, the authors highlight the need to complement the obtained results with those obtained from other databases, such as Clarivate Analytics, Web of Science or Scimago. Thus, it would be possible to analyse whether the results are robust. We have omitted any additional dimensions, such as consumer behaviour or e-fulfilment. The advent of mobile channels, tablets and social media is leading both retailers and consumers to interact in an omni-channel world [52]. Future research may, therefore, focus explicitly on the interaction of all of the existing channels. Second, similar to any other systematic review, our search is based on a limited set of keywords. Nevertheless, to the best of our knowledge, this study is the first systematic literature review on e-groceries/e-retailing. The insights obtained need further empirical investigation; however, some ideas can be used as a good starting point to develop sensible integrative models based on an experimental design and choice modelling.

**Author Contributions:** J.C.M. acted as the coordinator of this study, collected the material and carried out the analyses on the reviewed studies. F.P. reviewed the studies collected from the perspective of the existing gaps. C.R. reviewed the studies collected from the perspective of the applied models. All authors contributed substantially to the writing and the editing of the paper.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Capaldi, E.D. *Why We Eat What We Eat: The Psychology of Eating*; American Psychological Association: Washington, DC, USA, 1996.
2. Mkansi, M.; Eresia-Eke, C.; Emmanuel-Ebikake, O. E-grocery challenges and remedies: Global market leaders perspective. *Cogent Bus. Manag.* **2018**, *5*, 1459338. [[CrossRef](#)]
3. Pan, S.; Giannikas, V.; Han, Y.; Grover-Silva, E.; Qiao, B. Using customer-related data to enhance e-grocery home delivery. *Ind. Manag. Data Syst.* **2017**, *117*, 1917–1933. [[CrossRef](#)]
4. Tan, J.; Fu, H.Z.; Ho, Y.S. A bibliometric analysis of research on proteomics in Science Citation Index Expanded. *Scientometrics* **2014**, *98*, 1473–1490. [[CrossRef](#)]
5. Wang, L.; Zhao, L.; Mao, G.; Zuo, J.; Du, H. Way to accomplish low carbon development transformation: A bibliometric analysis during 1995–2014. *Renew. Sustain. Energy Rev.* **2017**, *68*, 57–69. [[CrossRef](#)]
6. Zyoud, S.H.; Fuchs-Hanusch, D. A bibliometric-based survey on AHP and TOPSIS techniques. *Expert Syst. Appl.* **2017**, *78*, 158–181. [[CrossRef](#)]
7. Nunkoo, R.; Hall, C.M.; Ladsawut, J. Gender and choice of methodology in tourism social science research. *Ann. Tour. Res.* **2017**, *63*, 207–210. [[CrossRef](#)]
8. Hall, C.M. Publish and perish? Bibliometric analysis, journal ranking and the assessment of research quality in tourism. *Tour. Manag.* **2011**, *32*, 16–27. [[CrossRef](#)]
9. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.



10. Griffiths, T.L.; Steyvers, M. Finding scientific topics. *Proc. Natl. Acad. Sci. USA* **2004**, *101* (Suppl. 1), 5228–5235. [[CrossRef](#)]
11. Pritchard, A. Statistical bibliography or bibliometrics. *J. Doc.* **1969**, *25*, 348–349.
12. Kolesar, M.B.; Galbraith, R.W. A services-marketing perspective on e-retailing: Implications for e-retailers and directions for further research. *Internet Res.* **2000**, *10*, 424–438. [[CrossRef](#)]
13. Darley, W.K.; Blankson, C.; Luethge, D.J. Toward an integrated framework for online consumer behavior and decision making process: A review. *Psychol. Mark.* **2010**, *27*, 94–116. [[CrossRef](#)]
14. Engel, J.F.; Kollat, D.T.; Blackwell, R.D. *Consumer Behavior*, 3rd ed.; Dryden: Hinsdale, IL, USA, 1978.
15. Engel, J.F.; Blackwell, R.D.; Miniard, P.W. *Consumer Behavior*, 5th ed.; Dryden: Hinsdale, IL, USA, 1986.
16. Nguyen, D.H.; de Leeuw, S.; Dullaert, W.E. Consumer behaviour and order fulfilment in online retailing: A systematic review. *Int. J. Manag. Rev.* **2018**, *20*, 255–276. [[CrossRef](#)]
17. De Leeuw, S.; Minguela-Rata, B.; Sabet, E.; Boter, J.; Sigurðardóttir, R. Trade-offs in managing commercial consumer returns for online apparel retail. *Int. J. Oper. Prod. Manag.* **2016**, *36*, 710–731. [[CrossRef](#)]
18. Rohm, A.J.; Swaminathan, V. A typology of online shoppers based on shopping motivations. *J. Bus. Res.* **2004**, *57*, 748–757. [[CrossRef](#)]
19. Agatz, N.; Campbell, A.; Fleischmann, M.; Savelsbergh, M. Time slot management in attended home delivery. *Transp. Sci.* **2011**, *45*, 435–449. [[CrossRef](#)]
20. Vareias, A.D.; Repoussis, P.P.; Tarantilis, C.D. Assessing customer service reliability in route planning with self-imposed time windows and stochastic travel times. *Transp. Sci.* **2018**, forthcoming. [[CrossRef](#)]
21. Maltz, A.B.; Rabinovich, E.; Sinha, R. Logistics: The key to e-retail success. *Supply Chain Manag. Rev.* **2004**, *8*, 48–54.
22. Mentzer, J.T.; Dewitt, W.; Keebler, J.S.; Min, S.; Nix, N.W.; Smith, C.D.; Zacharia, Z.G. Defining supply chain management. *J. Bus. Logist.* **2001**, *22*, 1–25. [[CrossRef](#)]
23. Jing, X.; Lewis, M. Stockouts in online retailing. *J. Mark. Res.* **2011**, *48*, 342–354. [[CrossRef](#)]
24. Hult, G.T.M.; Boyer, K.K.; Ketchen, D.J., Jr. Quality, operational logistics strategy, and repurchase intentions: A profile deviation analysis. *J. Bus. Logist.* **2007**, *28*, 105–132. [[CrossRef](#)]
25. Bower, A.B.; Maxham, J.G. Return shipping policies of online retailers: Normative assumptions and the long-term consequence of fee and free returns. *J. Mark.* **2012**, *76*, 110–124. [[CrossRef](#)]
26. Lantz, B.; Hjort, K. Real e-customer behavioural responses to free delivery and free returns. *Electron. Commer. Res.* **2013**, *13*, 183–198. [[CrossRef](#)]
27. Thirumalai, S.; Sinha, K.K. Customer satisfaction with order fulfillment in retail supply chains: Implications of product type in electronic B2C transactions. *J. Oper. Manag.* **2005**, *23*, 291–303. [[CrossRef](#)]
28. Lim, S.F.W.T.; Rabinovich, E.; Rogers, D.S.; Laseter, T.M. Last-mile supply network distribution in omnichannel retailing: A configuration-based typology. *Found. Trends Technol. Inf. Oper. Manag.* **2016**, *10*, 1–87. [[CrossRef](#)]
29. Lim, S.F.W.T.; Jin, X.; Srari, J.S. Consumer-driven e-commerce: A literature review, design framework, and research agenda on last-mile logistics models. *Int. J. Phys. Distrib. Logist. Manag.* **2018**, *48*, 308–332. [[CrossRef](#)]
30. Dadzie, K.Q.; Chelariu, C.; Winston, E. Customer service in the Internet-enabled logistics supply chain: Website design antecedents and loyalty effects. *J. Bus. Logist.* **2005**, *26*, 53–78. [[CrossRef](#)]
31. Kirchner, W.K. “Real-life” faking on the Edwards Personal Preference Schedule by sales applicants. *J. Appl. Psychol.* **1962**, *46*, 128–130. [[CrossRef](#)]
32. Cohen-Almagor, R. Internet history. In *Moral, Ethical, and Social Dilemmas in the Age of Technology: Theories and Practice*; IGI Global: Hershey, PA, USA, 2013; pp. 19–39.
33. Hofmann, T. Probabilistic latent semantic indexing. In Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Berkeley, CA, USA, 15–19 August 1999; pp. 55–57.
34. Tirunillai, S.; Tellis, G.J. Mining marketing meaning from online chatter: Analysis of big data using latent dirichlet allocation. *J. Mark. Res.* **2014**, *51*, 463–479. [[CrossRef](#)]
35. Blei, D.M. Probabilistic topic models. *Commun. ACM* **2012**, *55*, 77–84. [[CrossRef](#)]
36. Asuncion, A.; Welling, M.; Smyth, P.; Teh, Y.W. On smoothing and inference for topic models. In Proceedings of the UAI, Montreal, QC, Canada, 18–21 June 2009; pp. 27–34.
37. Ellison, G.; Ellison, S.F. Tax sensitivity and home state preferences in internet purchasing. *Am. Econ. J. Econ. Policy* **2009**, *1*, 53–71. [[CrossRef](#)]



38. Van Loon, P.; Deketele, L.; Dewaele, J.; McKinnon, A.; Rutherford, C. A comparative analysis of carbon emissions from online retailing of fast moving consumer goods. *J. Clean. Prod.* **2015**, *106*, 478–486. [[CrossRef](#)]
39. Othman, M.F.B.; Amir, N.B.; Mansu, M.B.; Zeki, A.M. Developing Mobile E-Groceries Model. In Proceedings of the 2015 4th International Conference on Advanced Computer Science Applications and Technologies (ACSAT), Kuala Lumpur, Malaysia, 8–10 December 2015; IEEE: New York, NY, USA, 2015; pp. 30–34.
40. Rotem-Mindali, O.C.; Salomon, I. Modeling consumers' purchase and delivery choices in the face of the information age. *Environ. Plan. B Plan. Des.* **2009**, *36*, 245–261. [[CrossRef](#)]
41. Harris, P.; Dall'Olmo Riley, F.; Riley, D.; Hand, C. Online and store patronage: A typology of grocery shoppers. *Int. J. Retail Distrib. Manag.* **2017**, *45*, 419–445. [[CrossRef](#)]
42. Punakivi, M.; Saranen, J. Identifying the success factors in e-grocery home delivery. *Int. J. Retail Distrib. Manag.* **2001**, *29*, 156–163. [[CrossRef](#)]
43. Siikavirta, H.; Punakivi, M.; Kärkkäinen, M.; Linnanen, L. Effects of E-Commerce on Greenhouse Gas Emissions: A Case Study of Grocery Home Delivery in Finland. *J. Ind. Ecol.* **2002**, *6*, 83–97. [[CrossRef](#)]
44. Cairns, S. Delivering supermarket shopping: More or less traffic? *Transp. Rev.* **2005**, *25*, 51–84. [[CrossRef](#)]
45. Mokhtarian, P.L. A conceptual analysis of the transportation impacts of B2C e-commerce. *Transportation* **2004**, *31*, 257–284. [[CrossRef](#)]
46. Verhoef, P.C.; Langerak, F. Possible determinants of consumers' adoption of electronic grocery shopping in the Netherlands. *J. Retail. Consum. Serv.* **2001**, *8*, 275–285. [[CrossRef](#)]
47. Sakao, T.; Shimomura, Y. Service Engineering: A novel engineering discipline for producers to increase value combining service and product. *J. Clean. Prod.* **2007**, *16*, 590–604. [[CrossRef](#)]
48. Fargnoli, M.; Costantino, F.; Di Gravio, G.; Tronci, M. Product service-systems implementation: A customized framework to enhance sustainability and customer satisfaction. *J. Clean. Prod.* **2018**, *188*, 387–401. [[CrossRef](#)]
49. Haber, N.; Fargnoli, M. Design for product-service systems: A procedure to enhance functional integration of product-service offerings. *Int. J. Prod. Dev.* **2017**, *22*, 135–164. [[CrossRef](#)]
50. Mikalef, P.; Giannakos, M.; Pateli, A. Shopping and Word-of-Mouth Intentions on Social Media. *J. Theor. Appl. Electron. Commer. Res.* **2013**, *8*, 17–34. [[CrossRef](#)]
51. Mikalef, P.; Giannakos, M.N.; Pappas, I.O. Designing social commerce platforms based on consumers' intentions. *Behav. Inf. Technol.* **2017**, *36*, 1308–1327. [[CrossRef](#)]
52. Verhoef, P.C.; Kannan, P.; Inman, J.J. From multichannel retailing to omni-channel retailing: Introduction to the special issue on multi-channel retailing. *J. Retail.* **2015**, *91*, 174–181. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).