Proceedings Paper:
Meriton, R and Graham, G orcid.org/0000-0002-9908-4974 (2016) Big data and supply chain management: A marriage of convenience? In: Proceedings. 20th Cambridge International Manufacturing Symposium, 29-30 Sep 2016, Cambridge, UK. University of Cambridge, Institute for Manufacturing . ISBN 9781902546766

Big Data and Supply Chain Management: A Marriage of Convenience?

[1]Royston Meriton and Gary Graham
Leeds University Business School
United Kingdom

Abstract

Big data is the new "guy about town." Indeed, the buzz about Big Data and business intelligence (BI) as drivers of business information data collection and analysis continues to build steam. But it seems not everyone is taking notice. Whilst scholars in main are excited about the "fields of possibilities" big data and related analytics offer, in terms of optimising firm capabilities, supply chain scholars have been surprisingly quiet. In this work we hope to break this silence and we achieve this through a comprehensive survey of the literature with the aim of exposing the dynamics of big data analytics in the supply chain context. Our findings suggest that the benefits of a big data driven supply chain are many on the proviso that organisations can overcome their own myopic understanding of this socio-technical phenomenon. However, this is not to suggest a one-size fits all approach, our findings also reveal that adopting a big data strategy in the supply chain is a strategic decision and as such, given the idiosyncrasies of industries, firms should leverage these technologies in congruence with their core capabilities. Strategic fit between a firm core competences and its big data strategy creates causal ambiguity which can in turn lead to sustainable competitive advantage.

Keywords: big data, big data analytics, supply chain, big data analytics-enabled supply chain, capabilities, resilience, performance

1. Introduction

Big data has the potential to revolutionize the art of supply chain management (Mehmood et al., 2016). Reporting on the economic impact of Big Data (Manyika et al., 2011), the McKinsey Global Institute has identified value levers along the whole manufacturing value chain with the ability to lead to high productivity levels. Application of big data and advanced analytics to R&D and product development has been estimated to reduce costs by 20-50 per cent, and a big data-enabled supply chain optimization was estimated as yielding a 2-3 percentage point profit margin improvement (Baily et al., 2013). However, in spite of its widely reported strategic impacts, there is a paucity of empirical research exploring the influence of big data in operations and supply chain management (SCM).

Fosso Wamba et al. (2015: 235) define big data as: "a holistic approach to manage, process and analyse the 5Vs (volume, velocity, variety, veracity, and value) in order to create actionable insights for sustained value delivery, measuring performance and establishing competitive advantages." Big data analytics (BDA), on the other hand, is the process of using analysis algorithms running on powerful supporting platforms to uncover potentials concealed in big data, such as hidden patterns or unknown correlations (Hu et al., 2014). Hence, as Russom (2011) observes, BDA is really about two things - big data and analytics - plus how the two have teamed up to create one of the most profound trends in business intelligence (BI) today. According to Fosso Wamba and Akter (2015) BDA is expected to take supply chain (SC) transformation to a level of transformation never before achieved. This could be because BDA represents a critical source of meaningful information that may help SC stakeholders gain improved insights that can be used for competitive advantage. For example, Fawcett and Waller (2014) expose how Amazon is making use of the predictive capability of BDA to predict and ship what a customer wants before the customer places their order. This has been termed anticipatory shipping (Lee, 2016). In turn, this can improve the customer experience and thus a source of competitive advantage. However, whilst supply chain professionals are devising new and innovative ways to extract value from BDA-enabled activities, there is a dearth of empirical academic work published in the last five years, particularly in the business and management subject area. This may be because the area is still an emerging field in an early developmental stage.

The aim of this work is therefore to examine the academic literature with a view of mapping the common terrain where BDA and SCM intersect in a bid to advance knowledge on the emerging dynamics in this space. We are thus guided by the following research questions:

- What is the potential of BDA in the SCM to improve performance?

- Are there any barriers to adopting BDA in SC, and if so what are they?

We conducted a detailed systematic survey of the literature in the traditional manner using keywords, complemented by visualisations in VOS MAX and cluster analysis of the most common terms from the articles extracted.

2. Methodology

In this study we build on the previous review conducted by Fosso Wamba and Akter (2015) which was based on a hybrid approach. This approach consisted of a search using the following keywords: "Big data" AND "supply chain" within the SCOPUS database. Scopus, officially named SciVerse Scopus, was introduced by Elsevier in November 2004 to the information market (Aghaei Chadegani et al., 2013). It is the largest database existing on the market for multidisciplinary scientific literatures (Bornmann et al., 2009) and has often been used as a valuable information repository for supply chain scholars (e.g., Fahimnia et al., 2015). Scopus covers 27 million abstracts, 230 million references and 200 million web pages (Bar-Ilan, 2008). It contains 20,500 peer-reviewed journals from 5,000 publishers, together with 1200 Open Access journals, over 600 Trade Publications, 500 Conference Proceedings and 360 book series from all areas of science (Aghaei Chadegani et al., 2013).

The search was conducted on Monday 19 September 2016, and our search returned 139 titles registered in the Scopus database. These included 59 journal articles of which 12 were in press. Conference papers were in the majority with 75 entries, the balance being made up of reviews, one book chapter and a short survey. We subjected the articles from the result of the Scopus search to cluster analysis to help provide a holistic picture of the most common words/themes and the relationship between them (see details in the appendix). The visualisations of these are illustrated in Figure 1 whereas the result of the cluster analysis is provided in Table 4 in the appendix.

3. Result

This section presents and discusses the findings of big-data-related SC publications identified in SCOPUS, we also provide a summary of the content analysis for the articles only in Table 5 in the appendix. In Table 1 the distribution of publications by the year of publication is presented. The trend suggests an increasing interest in the subject area with 2015 accounting for almost 50% of all the publications (46%). This upward intake in trend is also apparent with the articles, but it remains to be seen whether 2016 will be as successful a year for the subject area as 2015.

Table 1: Distribution of publications by year

Year	Number of Publication	Percentage	Number of Articles (incl. in press)	Percentage
2012	2	1	1	2
2013	9	7	4	7
2014	28	20	9	15
2015	64	46	28	47
2016	36	26	17	29
Total	139	100	59	100

The distribution of publications by subject area is presented in Table 2. The top 4 subject areas accounted for 77% of all the publications and these include: (1) Computer Science (29%); (2) Engineering (24%); (3) Business, Management and Accounting (13%) and; (4) Decision Sciences (11%). It is interesting to note that the same four subject areas occupied the top spots in terms of most number of publications in Fosso Wamba and Akter (2015) review, however in a slightly different order with “Business, Management and Accounting” holding the top spot (25.8%). These findings indicate that the dynamics between supply chain and big data is being discussed from a more technological perspective rather than in a business and management sense, which is the traditional home to supply chain management. This may be attributed to the nascent nature of BDA which still commands a high cost for adoption.

Table 2: Distribution of publications by subject area

Subject Area	Number of publications	Percentage
Computer Science	85	29
Engineering	69	24
Business, Management and Accounting	38	13
Decision Sciences	31	11
Mathematics	21	7
Social Sciences	12	4
Economics, Econometrics and Finance	6	2
Materials Science	5	2
Energy	4	1
Chemical Engineering	3	1
Environmental Science	3	1
Physics and Astronomy	3	1
Agricultural and Biological Sciences	2	1
Arts and Humanities	2	1
Medicine	2	1
Multidisciplinary	2	1
Chemistry	1	0
Nursing	1	0
Psychology	1	0
*Total	291	100

Note: *Some articles are counted more than once because they cover more than one subject area.

Table 3 shows the classification of publications by country. So far, most published articles on big data in the supply chain context come from China (24.2%), with the US following closely behind (22%). These figures are contrasted with the findings in Fosso Wamba and Akter (2015), who reported that the US authors held a significantly larger share of the publication spoils (63.2%) compared to the Chinese authors (21.1%). It seems that in just over one year we have seen a 'change of the guard,' with the rest of the world, spearheading by China, also wanting to flex their muscles in this space.

Table 3: Distribution of publications by country

Country/Territory	Number of publications	Percentage
China	41	24
United States	38	22
Germany	15	9
United Kingdom	9	5
India	7	4
Taiwan	5	3
Netherlands	4	2
Spain	4	2
Finland	3	2
France	3	2
Ireland	3	2
Singapore	3	2
South Korea	3	2
Australia	2	1
Brazil	2	1
Denmark	2	1
Hong Kong	2	1
Italy	2	1
Switzerland	2	1
Belgium	1	1
Canada	1	1
Czech Republic	1	1
Greece	1	1
Japan	1	1
Malaysia	1	1
Norway	1	1
Pakistan	1	1
Russian Federation	1	1
South Africa	1	1
Sweden	1	1
Undefined	12	7
*Total	173	100

Note: *Some articles are counted more than once because their authors come from different countries

Fosso Wamba and Akter (2015: 4) noted that "the impact of BDA in emerging economies and less developed countries should be part of future research directions." Indeed, the advice appears to have been heeded with authors from countries such as Brazil, India, Russia, South

Africa, and Pakistan beginning to make their voices heard, albeit respectively contributing to 1% of the total amount of publications. It is thus not surprising that the most prolific authors researching in this space originate from the US and China respectively. Both S.E. Fawcett and M.A. Waller, have 4 articles each on BDA-related SC topics to their names, followed closely by Li, P, Liu, Y.Q., and Luo, H, each recognised for 3 publications

4. Findings

We elected to retain the entries classed as 'journal articles' in the database for a detailed content analysis. Following further screening of the abstracts, a total of 35 articles were found to meet our criteria for inclusion for detailed analysis. The main reason articles were excluded from the final list related to whether or not the full text was available online. From the sample of 59 journal articles, 17 were dropped because the full text document could not be located or downloaded, 2 were excluded as they were not scholarly articles and a further 5 were deemed not relevant to the study. The analysis of the 35 articles captured attributes of the studies such as methodological approaches and the industries. Table 5 in the appendix provides a summary of our findings whereas the graphics in Figure 1 provide a snap shot of the most significant relationships. In the sections that follow we draw on the analyses in Tables 4 and 5 and the visuals in Figure 1 to discuss the headlines of our findings paying particular attention to emerging trends, the most common of these include method, opportunities and barriers.

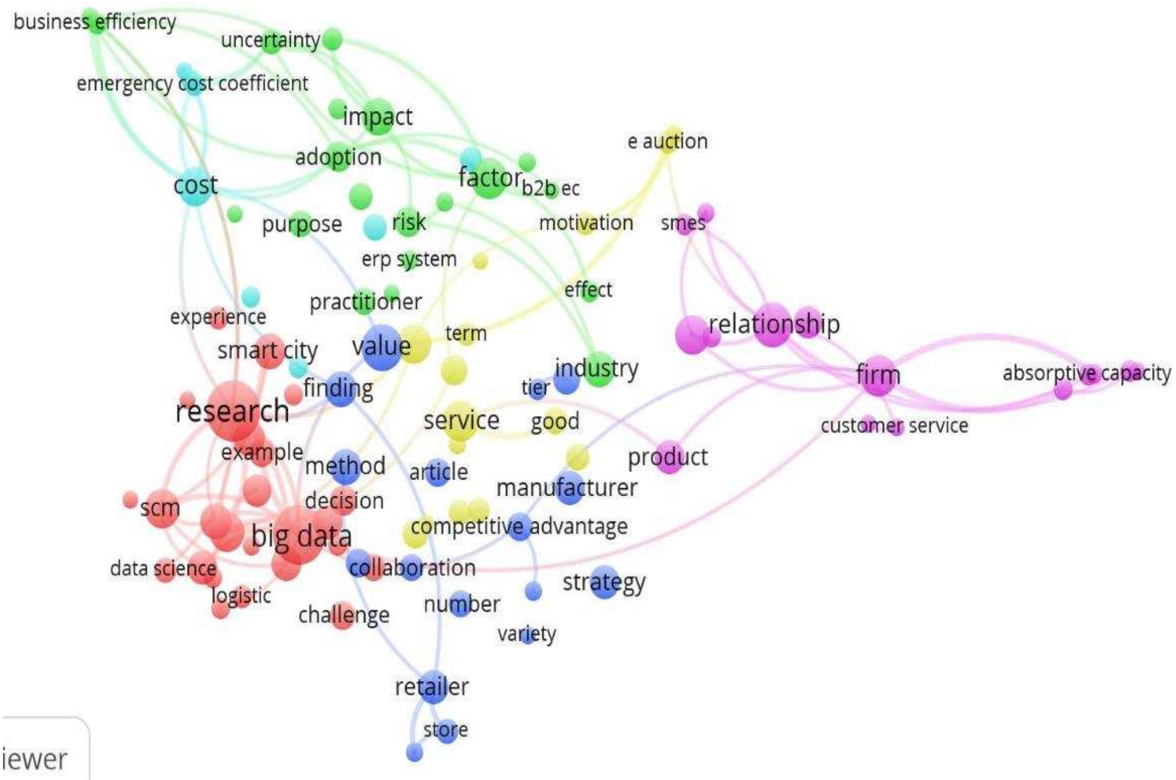

Figure 1: Visuals of dominant themes and relationships

4.1 Dominant methodological approaches

The methodological approach was a mixed bag of quantitative methods and literature reviews often leading to a conceptual and/or theoretical framework. In the main, the studies analysed were finely poised with 18 being mostly quantitative and 17 non-quantitative in nature. There seems to be an increasing level of novelty and complexity in the methods employed. Perhaps this is a reflection on the nature of the task. For instance, the quantitative studies were often preceded by mathematical modelling followed by experiments or case studies to test the model derived. For example, Bock and Isik (2015) developed a mathematical model to quantify intricacy in purchasing order sizing processes. They then followed this up by using 'structured' big data related to costs and the structure of recent deviations from desired target values, to conduct scenario analyses in order to determine the reliability of the model. Only four studies relied on primary quantitative data collected by means of survey questionnaires, these include Chen et al. (2015) who surveyed 161 U.S.-based companies, van der Spoel et al. (2015) who collected data from 230 truckers and Schoenherr and Speier-Pero (2015) study involving 531 SCM professionals. The other study by (Papadopoulos et al., 2015) made use of both survey data as well as data from social media such as tweets and Facebook. However, most of the studies actually made use of secondary "big data."

Big data employed was of both the structured and unstructured nature. In Li et al. (2015) the authors analysed structured data obtained from an omni-channel marketing platform to test a new enterprise networks integration architecture. On the other hand, Papadopoulos et al. (2015) analysed data in the form of tweets and Facebook, to help develop a framework for supply chain resilience. Other means of gathering big data included RFID-enabled intelligent from a shop floor environment (Zhong et al., 2016). The non-quantitative method (as we have termed it) was dominated by conceptual and/or theoretical framework development often involving a systemic survey of the literature. It is interesting to note that only one article in this category actually employed a qualitative approach a la conventional type. In this study Kumar et al. (2016) adopted a case study approach at the product level to gauge at the impact big data was having on the traditional supply chain.

4.2 Opportunities presented by BDA to SCM

Our findings suggest that the potential of BDA as a valuable resource to the organization has not gone unnoticed. However, there are two sides to this "opportunities coin," this is because, there is much by way of commentaries, theoretical framing, simulations and experiments and still little empirical evidence. Much of the conversation tend to be past oriented, in the sense that authors consider existing supply chain analytics and speculate on how the integration of BDA could potentially help to optimise supply chain activities. A good example is provided by

Rehman et al. (2016). In this work the authors introduce the idea of early big data reduction (BDR) and go on to make a case of how it can be leveraged in supply chains to improve efficiency. Antagonism about the lack of empirical work aside, there is widespread sentiment that BDA is a necessary evil that increasingly firms ignore at their own perils. In this context, the value potential of BDA unearthed by the findings can be themed around (1) capabilities and, (2) performance.

4.2.1 Capabilities

Drawing on the findings of the empirical studies, scholars generally recognised that BDA could be leveraged in different parts of the supply chain in order to create value. In other words, there is widespread acceptance that BDA is valuable when it is used to create distinct capabilities. In the main, the findings reveal that organizations are making use of the predictive proclivities of BDA to strengthen their decision making capabilities (Schoenherr & Speier-Pero, 2015) in a number of key supply chain activities. Some of these big-data enabled capabilities include market sensing (Chae, 2015; Lee, 2016; Li et al., 2015; Liu & Wang, 2016), planning and forecasting in different areas such as in logistics (Liu & Wang, 2016; Zhong et al., 2015) and demand and sales (Schoenherr & Speier-Pero, 2015), risk management (Papadopoulos et al., 2015; Wu et al., 2015; Zhao et al., 2015; Zou et al., 2016) and innovation (Tan et al., 2015).

4.2.2 Performance

Distinct capabilities often bestow performance advantages on the incumbent firms. The BDA-enabled performance advantages revealed by the findings include cost efficiencies (Bock & Isik, 2015; Hofmann, 2015; Li & Wang, 2015; Li et al., 2015; Liu et al., 2016; Schoenherr & Speier-Pero, 2015), enhanced customer services (Lee, 2016; Liu & Wang, 2016), agility, in terms of speed, flexibility and responsiveness (Giannakis & Louis, 2016; Kumar et al., 2016) and business growth (Chen et al., 2015).

4.3 Barriers to adopting BDA in SCM

Whilst the value of BDA to firms was recognised, this was also greeted by an air of cautious optimism by many authors. The conversations around barriers in implementing BDA in SCM mostly took place in the non-empirical articles. Cost associated with currently available solutions (Schoenherr & Speier-Pero, 2015) was one of the most cited drawbacks, for example, Rehman et al. (2016) observed that cloud service utilization costs increase because of big data analytics and value creation activities. On a related note, Hazen et al. (2014) pondered on data quality noting that the costs of poor data quality have been estimated to be as high as 8% to 12% of revenues for a typical organization and may generate up to 40% to 60% of a service organization's expenses. Issues around organizational readiness have also been highlighted. For instance, Schoenherr and Speier-Pero (2015) found employees being inexperienced, time constraints, lack of integration with current systems, change management issues, lack of appropriate predictive analytics solutions for SCM, as well as the perception of SCM predictive analytics being overwhelming and difficult to manage amongst the primary barriers for adopting BDA in SCM. It is therefore not surprising that Sonka (2014) concluded that successful application of big data and associated analytical techniques in the agriculture supply chain would also depend on organizational and managerial factors.

5. Discussion and conclusion

In this paper we sought to understand the potential of BDA in the SC to improve performance, as well as the barriers, if any, that constrain its adoption in SC. We conducted a detailed literature survey using "big data" and "supply chain" as key words. The search returned 139 publications. Applying VOS MAX to the search results we complemented the initial search with visualisations and cluster analysis that helped us understand the dominant themes and the nature of the relationships between them. Thus we provided a robust approach to analyse the existing literature.

Our analysis suggests an increasing trend in the number of publications, year on year, with authors from more countries getting involved. Particularly, the last five years have seen a shift from the US towards the east in terms of the number of articles published, with an increasing participation of authors from emerging economies such as Brazil, Russia and India, with China leading the way. Perhaps this trend could be attributed to faster uptake of BDA in SCM particularly in China, which in turn could reflect the relative costs of such technological adoption in China compared to the US. Furthermore, the findings reveal that "business and management area" is lacking behind computer science and engineering in terms of representativeness. This could be explained by the fact that most of the empirical work are of a quantitative nature often involving complex modelling using advanced algorithms which make use of secondary data. Perhaps scholars from these specialisms, that is, engineering and computer science, are better equipped with the relevant skills for these approaches. The dominant methodologies in the business and management discipline generally revolve around primary data collection through surveys or case studies, both of which requiring greater efforts to collect data while access to companies, especially in the western economies, becoming ever so harder to secure.

The findings suggest that adoption of BDA in the SC has the potential to impact on firm performance. However, we found that the value of BDA to firms is not inherent in the BDA per se but rather by being embedded within supply chain activities. In other words, BDA is causally efficacious in enabling the emergence and evolution of unique capabilities. In particular, the findings reveal that firms are leveraging the predictive proclivities of BDA to strengthen their decision making capability for various SC activities, here we wish to single out sensing and innovation capabilities. These capabilities are quintessential for competing in dynamic environments. For example, sensing capabilities feature as one of the cornerstones of Teece's (2007) microfoundations of dynamic capabilities (DC). Furthermore, innovation is often cited as the linchpin of dynamic competition (Schumpeter, 2000; Zollo & Winter, 2002). According to Eisenhardt and Martin (2000: 1107), dynamic capabilities are "the firm's processes that use resources - specifically the processes to integrate, reconfigure, gain and release resources - to match and even create market change." The emergence of anticipatory shipping, based on BDA-enabled sensing capabilities from Amazon can be thus be seen in the light of creating market changes which therefore chimes with the idea of DC. Here we wish to press home the important point that DC is a strategic enterprise. Hence, the findings serve to reiterate the strategic importance of a firm's supply chain particularly as a means of sustainable competitive advantage (Opresnik & Taisch, 2015; Rehman et al., 2016) and firm growth (Chen et al., 2015). Therefore, the findings serve to position BDA in a firm's strategic conversation particularly illuminating its role as enabler of dynamic capabilities and sustainable competitive advantage.

The findings also reveal that the adoption of BDA improves supply chain agility, particularly noticeable in performance metrics of flexibility, responsiveness and speed (Giannakis & Louis, 2016). These operations performance measures are linked with supply chain resilience. In fact a number of authors have accentuated the criticality of BDA to SC risk management with Wu et al. (2015) arguing that applying relevant analytical techniques to unstructured and structured data can help firm identify areas of risks that may impact on the sustainability of the supply chain. It is now widely accepted that SC resilience can be a dynamic capability. However, whilst DC tend to focus on a particular functional area, such as R&D for instance, resilience speaks to a whole system's system approach to adaptation and can be both proactive and reactive. Supply chain resilience is widely considered as an important component of supply chain management, however, it is surprising that only two articles make direct reference to it. Nevertheless, resilience presents firms with a different route to adaptation and the findings provide some useful insights as to the dynamics of BDA-enabled resilient supply chains.

5.1 A look into the future

The litmus test for the adoption of BDA in SCM rests on a number of permutations accessibility being one of them. The findings reveal cost of the available options as one of the key barriers.

This scenario tends to put particularly small firms at a disadvantage, in fact it seems to be the case that the uptake of this technological innovation is a lot more pronounced in large and well established firms; Amazon, Ikea, Toyota, to name but a few. The cost of adoption may also be one of the reasons behind the surge of publications from China. It may be the case that the cost of adoption in China is relatively cheaper than in other parts of the world encouraging a higher level of adoption. This could potentially result in a wider diversity of firms making use of BDA in their operations and thus providing for a larger population of organizations for researchers to draw from. Whilst we agree with Fosso Wamba and Akter (2015) in that BDA has the potential to take supply chain (SC) transformation to a level never before achieved, barriers such as costs and organizational readiness issues highlighted would need to be addressed a priory.

Furthermore, BDA is a relatively recent addition in the SCM domain and perhaps still exudes an air of cautious optimism. More needs to be done in terms of researching cost effective solutions that are currently being used. Most of the empirical studies that we have analysed tend to be detached from real life settings. There is a real world out there where these technologies are being put to use on a daily basis and companies are reaping the rewards. We are thus advocating for research of the longitudinal case study type, perhaps of an ethnographic approach, in order to further our understanding as to the process of capabilities emergence and evolution in the context of BDA. It is our conviction that by doing so scholars will help to disseminate best practices more widely which in turn could increase the rate of adoption and bring down the costs involved whilst also exposing the benefits to a wider audience. As we see it, BDA and SCM may well be in a marriage of convenience particularly for those organizations that boast a healthy balance sheet.

References

AGHAEI CHADEGANI, A., SALEHI, H., YUNUS, M. M., FARHADI, H., FOOLADI, M., FARHADI, M. & ALE EBRAHIM, N. 2013. A comparison between two main academic literature collections: Web of Science and Scopus databases. *Asian Social Science,* 9**,** 18-26.

BAILY, M. N., MANYIKA, J. & GUPTA, S. 2013. US productivity growth: An optimistic perspective. *International Productivity Monitor***,** 3.

BAR-ILAN, J. 2008. Which h-index?—A comparison of WoS, Scopus and Google Scholar. *Scientometrics,* 74**,** 257-271.

BOCK, S. & ISIK, F. 2015. A new two-dimensional performance measure in purchase order sizing. *International Journal of Production Research,* 53**,** 4951-4962.

BORNMANN, L., MARX, W., SCHIER, H., RAHM, E., THOR, A. & DANIEL, H.-D. 2009. Convergent validity of bibliometric Google Scholar data in the field of chemistry—Citation counts for papers that were accepted by Angewandte Chemie International Edition or rejected but published elsewhere, using Google Scholar, Science Citation Index, Scopus, and Chemical Abstracts. *Journal of informetrics,* 3**,** 27-35.

CHAE, B. 2015. Insights from hashtag #supplychain and Twitter analytics: Considering Twitter and Twitter data for supply chain practice and research. *International Journal of Production Economics,* 165**,** 247-259.

CHEN, D. Q., PRESTON, D. S. & SWINK, M. 2015. How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems,* 32**,** 4-39.

EISENHARDT, K. M. & MARTIN, J. A. 2000. Dynamic capabilities: what are they? *Strategic management journal,* 21**,** 1105-1121.

FAHIMNIA, B., TANG, C. S., DAVARZANI, H. & SARKIS, J. 2015. Quantitative models for managing supply chain risks: A review. *European Journal of Operational Research,* 247**,** 1-15.

FAWCETT, S. E. & WALLER, M. A. 2014. Supply chain game changers-mega, nano, and virtual trends-and forces that impede supply chain design (i.e., Building a Winning Team). *Journal of Business Logistics,* 35**,** 157-164.

FOSSO WAMBA, S. & AKTER, S. 2015. Big data analytics for supply chain management: A literature review and research agenda. *Workshop on Enterprise and Organizational Modeling and Simulation*: Springer, 61-72.

FOSSO WAMBA, S., AKTER, S., EDWARDS, A., CHOPIN, G. & GNANZOU, D. 2015. How 'big data'can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics,* 165**,** 234-246.

GIANNAKIS, M. & LOUIS, M. 2016. A multi-agent based system with big data processing for enhanced supply chain agility. *Journal of Enterprise Information Management,* 29**,** 706-727.

HAHN, G. J. & PACKOWSKI, J. 2015. A perspective on applications of in-memory analytics in supply chain management. *Decision Support Systems,* 76**,** 45-52.

HAZEN, B. T., BOONE, C. A., EZELL, J. D. & JONES-FARMER, L. A. 2014. Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics,* 154**,** 72-80.

HOFMANN, E. 2015. Big data and supply chain decisions: the impact of volume, variety and velocity properties on the bullwhip effect. *International Journal of Production Research*.

HU, H., WEN, Y., CHUA, T. S. & LI, X. 2014. Toward scalable systems for big data analytics: A technology tutorial. *IEEE Access,* 2**,** 652-687.

KANG, Y. S., PARK, I. H., RHEE, J. & LEE, Y. H. 2016. MongoDB-Based Repository Design for IoT-Generated RFID/Sensor Big Data. *IEEE Sensors Journal,* 16**,** 485-497.

KUMAR, M., GRAHAM, G., HENNELLY, P. & SRAI, J. 2016. How will smart city production systems transform supply chain design: a product-level investigation. *International Journal of Production Research*, 1-12.

LEE, C. K. H. 2016. A GA-based optimisation model for big data analytics supporting anticipatory shipping in Retail 4.0. *International Journal of Production Research*, 1-13.

LI, D. & WANG, X. 2015. Dynamic supply chain decisions based on networked sensor data: an application in the chilled food retail chain. *International Journal of Production Research*.

LI, Q., LUO, H., XIE, P. X., FENG, X. Q. & DU, R. Y. 2015. Product whole life-cycle and omni-channels data convergence oriented enterprise networks integration in a sensing environment. *Computers in Industry,* 70, 23-45.

LIU, Y. Q. & WANG, H. 2016. Order allocation for service supply chain base on the customer best delivery time under the background of big data. *International Journal of Computer Science and Applications,* 13, 84-92.

LIU, Y. Q., WANG, H., ZHANG, Y. & CAI, C. 2016. Order allocation of logistics service under background of big data. *Shenyang Gongye Daxue Xuebao/Journal of Shenyang University of Technology,* 38, 190-195.

MANYIKA, J., CHUI, M., BROWN, B., BUGHIN, J., DOBBS, R., ROXBURGH, C. & BYERS, A. H. 2011. Big data: The next frontier for innovation, competition, and productivity.

MEHMOOD, R., MERITON, R., GRAHAM, G., HENNELLY, P. & MUKESH, K. 2016. Exploring the influence of big data on city transport operations: A Markovian Approach. *International Journal of Operations and Production Management (forthcoming)*.

ÖBERG, C. & GRAHAM, G. 2016. How smart cities will change supply chain management: A technical viewpoint. *Production Planning and Control,* 27, 529-538.

OPRESNIK, D. & TAISCH, M. 2015. The value of big data in servitization. *International Journal of Production Economics,* 165, 174-184.

PAPADOPOULOS, T., GUNASEKARAN, A., DUBEY, R., ALTAY, N., CHILDE, S. J. & FOSSO-WAMBA, S. 2015. The role of Big Data in explaining disaster resilience in supply chains for sustainability. *Journal of Cleaner Production*.

RADKE, A. M. & TSENG, M. M. 2015. Design considerations for building distributed supply chain management systems based on cloud computing. *Journal of Manufacturing Science and Engineering, Transactions of the ASME,* 137.

REHMAN, M. H. U., CHANG, V., BATOOL, A. & WAH, T. Y. 2016. Big data reduction framework for value creation in sustainable enterprises. *International Journal of Information Management,* 36, 917-928.

RUSSOM, P. 2011. Big data analytics. *TDWI Best Practices Report, Fourth Quarter*, 1-35.

SCHOENHERR, T. & SPEIER-PERO, C. 2015. Data science, predictive analytics, and big data in supply chain management: Current state and future potential. *Journal of Business Logistics,* 36, 120-132.

SCHUMPETER, J. A. 2000. Entrepreneurship as innovation. *Entrepreneurship: The social science view*, 51-75.

SONKA, S. 2014. Big data and the ag sector: More than lots of numbers. *International Food and Agribusiness Management Review,* 17, 1-20.

TAN, K. H., ZHAN, Y., JI, G., YE, F. & CHANG, C. 2015. Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. *International Journal of Production Economics,* 165, 223-233.

TEECE, D. J. 2007. Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal,* 28, 1319-1350.

TIEN, J. M. 2012. The next industrial revolution: Integrated services and goods. *Journal of Systems Science and Systems Engineering,* 21, 257-296.

TIEN, J. M. 2015. Internet of connected ServGoods: Considerations, consequences and concerns. *Journal of Systems Science and Systems Engineering,* 24, 130-167.

VAN DER SPOEL, S., AMRIT, C. & VAN HILLEGERSBERG, J. 2015. Predictive analytics for truck arrival time estimation: a field study at a European distribution center. *International Journal of Production Research*.

WALLER, M. A. & FAWCETT, S. E. 2013a. Click here for a data scientist: Big data, predictive analytics, and theory development in the era of a maker movement supply chain. *Journal of Business Logistics,* 34**,** 249-252.

WALLER, M. A. & FAWCETT, S. E. 2013b. Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics,* 34**,** 77-84.

WANG, G., GUNASEKARAN, A., NGAI, E. W. T. & PAPADOPOULOS, T. 2016. Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics,* 176**,** 98-110.

WU, C. X., ZHAO, D. Z. & PAN, X. Y. 2016. Comparison on dynamic cooperation strategies of a three-echelon supply chain involving big data service provider. *Kongzhi yu Juece/Control and Decision,* 31**,** 1169-1177.

WU, K. J., LIAO, C. J., TSENG, M. L., LIM, M. K., HU, J. & TAN, K. 2015. Toward sustainability: Using big data to explore the decisive attributes of supply chain risks and uncertainties. *Journal of Cleaner Production*.

ZHAO, R., LIU, Y., ZHANG, N. & HUANG, T. 2015. An optimization model for green supply chain management by using a big data analytic approach. *Journal of Cleaner Production*.

ZHONG, R. Y., NEWMAN, S. T., HUANG, G. Q. & LAN, S. 2016. Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives. *Computers and Industrial Engineering*.

ZHONG, R. Y., XU, C., CHEN, C. & HUANG, G. Q. 2015. Big Data Analytics for Physical Internet-based intelligent manufacturing shop floors. *International Journal of Production Research*.

ZOLLO, M. & WINTER, S. G. 2002. Deliberate learning and the evolution of dynamic capabilities. *Organization Science,* 13**,** 339-351.

ZOU, X., TAO, F., JIANG, P., GU, S., QIAO, K., ZUO, Y. & XU, L. 2016. A new approach for data processing in supply chain network based on FPGA. *International Journal of Advanced Manufacturing Technology,* 84**,** 249-260.

Appendix

Table 4: Cluster analysis

Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**	**Cluster 6**
Big Data Analytics Project Data Science Decision Development Logistics Opportunity Organization Pattern Practice Predictive Analytics SCM Smart City	Adoption B2B Ecommerce Business Efficiency eCommerce Impact ERP system Impact Industry System Product Design Strategic Sourcing Transport Operation Uncertainty	Collaboration Competitive advantage Enterprise Manufacturer Physical Distribution Servitization Strategy Value Variety	Demand chain Implementation Innovation Integration Internet Same time Service Use Time	Ability Absorptive capacity Composite framework Customer Customer service Firm Firm performance Lean strategy Product Relationship Relationship building Responsive strategy SMEs Supplier Supply chain integration	Cost Emergency cost coefficient Manager Objective

Table 5: Synthesis of literature on “big data” and “supply chain”

Authors	Approach/Industry/Big Data Type	Findings/Use of Big Data/Analytics
Qualitative/Conceptual/Theoretical/Reviews/Editorials		
1. Hahn and Packowski (2015)	• Bottom-up characterization of 41 in-memory analytics applications in SCM	Out of 41 in-memory analytics applications in SCM only one was found that readily handles big data and this was: • Systema Big Data Real Time Analysis for Manufacturing The main application is mostly used for the purpose of • Monitor and navigate
2. Sonka (2014)	• Commentary on the potential big data and big data analytics offer to SCM in the agricultural sector(Bock & Isik, 2015)	• This paper interrogates the costs and benefits of leveraging big data in the agriculture value chain for its optimization. The general suggestion is that while technology has and will continue to drive performance levers successful application of big data and associated analytical techniques in the agriculture supply chain would also depend on organizational and managerial factors.
3. Rehman et al. (2016)	• Conceptual/theoretical • Introduces the idea of early big data reduction	• The paper identifies 10 organizational activities where BDR could be employed to improve efficiency. • In terms of supply chain it argues that BDR can reduce costs of operation, in other words optimise supply chain operations. • Linked to competitive advantage and value creation
4. Opresnik and Taisch (2015)	1. Conceptual 2. Simulation is used to design the Big Data Strategy framework in servitization and to simulate the related opportunities	• Proposes a Big Data strategy to create competitive advantage by data reuse and data resell in the context of servitization in the manufacturing industry
5. Hazen et al. (2014)	Focuses on data quality but does not investigate the use of big data per se	

6. Fawcett and Waller (2014)	• Editorial	• In this editorial the authors single out the profiling capabilities of big data analytics and discusses how Amazon are making use of this predictive capability to predict and ship what a customer wants before the customer places the order.
7. Waller and Fawcett (2013b)	• Conceptual/commentary	• It provides some examples as to the applications of big data in logistics • Sets a research agenda around some of these applications
8. Waller and Fawcett (2013a)	• Conceptual/theoretical	• Takes a "maker" perspective of the supply chain and try to unpack the nature of the producer. • It also argues in favour of predictive analytics as an important component in the theory development process around the supply chain domain. • Sets out a future research agenda for BDA in SCM.
9. Radke and Tseng (2015)	This paper outlines the procedures that should be considered in building distributed supply chain systems based on cloud computing	
10. Tien (2015)	Conceptual – builds on the 2012 paper	
11. Tien (2012)	• Conceptual/theoretical	• Speculative at best. • This paper investigates the modalities of transformation to mass customization in the context of IoT and big data analytics. • It proposes that the adoption of big data analytics within the supply chain offers the flexibility that supports the principles of mass customization
12. Öberg and Graham (2016)	• Commentary/conceptual	• In the context of smart cities the paper discusses the dynamics between smart cities technologies and the supply network. • In particular, it highlights the role of big data in transforming the logistics

		and transportation domain of supply chains.
13. Lee (2016)	• Cluster-based association rule mining applied to big data related to purchasing patterns of customers	1. In the context of omni-channel application of big data analytics helped to predict customer purchasing patterns. 2. This allows for optimal solution for anticipatory shipping which in turn improves the customer experience thus a source of competitive advantage.
14. Kumar et al. (2016)	• Product level analysis • Conceptualisation through literature review followed by • A case study approach by analysis the manufacturing of 4 different products	Application of big data analytics has the potential to affect the supply chain as follows: 1. quick response to demand; 2. production of products with minimum modularity; The impact of this on upstream includes: 1. reduction in supply chains nodes, 2. nodes size and 3. limited number tiers (reduce complexity)
15. Hazen et al. (2014)	• Conceptual/theoretical	• In this paper the authors discuss the potential of analytics to improve the SCM practices. • These are discussed in terms of 3 analytical methods, descriptive, predictive and prescriptive (has a few references in there that might be useful)
16. Zhong et al. (2016)	• Review/commentary/future research agenda	• This is a review paper that discusses the potential uses of big data and big data analytics in service and manufacturing SCM. • It is informative in the sense that it provides a good narrative on the big data phenomenon.
17. Wang et al. (2016)	Review paper	

Mostly Quantitative Studies		
18. Li et al. (2015)	• Structured data • A data convergence oriented enterprise networks integration architecture with relative enabling technologies is conceptualised • The architecture is thereafter tested with a case study using data from an omni-channel marketing platform of an air purifier provider in China	• Used in the context of omni-channel marketing management • It is proposed that the method can improve product tracking and marketing channel management capabilities and reduce the operation cost (Sensing)
19. Kang et al. (2016)	• Big data repository design protocol • Plus experimentation	• This work does not really address the performative aspects of supply chain but rather discusses the protocols for designing a repository able to store different types of big data generated from RFID/IoT
20. Zou et al. (2016)	• Simulation - Kalman filter algorithm and median filter algorithm for FPGA vs. traditional analytical methods	• In light of the intricacies of big data in relation to supply chain network this paper discusses a new approach to data processing based on field programmable gate array (FPGA). • It is argued that this approach helps to mitigate against the risks posed by untimely data processing to SCN (Risk management)
21. Liu and Wang (2016)	• Two tier analysis of unstructured data based on: 1. Predictive analysis using big data 2. Model optimization • Predictive analysis of customers' demands on the basis of hits and browsing time • The behaviour of clients is analysed according to the their location data	• Big data can better help companies understand customer's interest, for service improvement. (Forecasting) • It can reduce the cost of logistics services under the premise of improving the level of customer service • It can enhance the credibility of the enterprise, which is conducive to long-term development of enterprises (in other words it can help companies achieve sustainable competitive advantage)

22. Tan et al. (2015)	• Unstructured data • Deduction graph modelling and case study for testing purposes.	• Enhances supply chain innovation capabilities
23. Chae (2015)	• Unstructured data • Descriptive analytics • content analytics and • Network analytics on #supplychain tweets	The findings indicate that the knowledge extracted from big data is used in mostly two domains: 1. Professional use – learning; promoting; networking 2. Organizational use – stakeholder engagement; hiring; demand shaping and sales; market sensing and new product/service development; risk management (sensing and forecasting)
24. Bock and Isik (2015)	• Mathematical modelling of a new approach to quantify intricacy in purchasing order sizing processes • Scenario analysis involving four scenarios to test the proposed 2-dimentional matrix • Uses structured big data related to costs and the structure of recent deviations from desired target values	The findings suggest as follows: 1. The model can detect efficiency in purchasing order sizing process 2. This information can be used to train managers 3. Which in turn can lead to significant reduction of intricacy costs
25. Papadopoulos et al. (2015)	• Analysis of unstructured big data in the form of tweets, Facebook etc., and • Survey questionnaire followed by content analysis and CFA.	Used big data analysis to develop a framework for supply chain resilience in the context of recovery from a disaster. The findings suggests that following as key in shaping supply chain resilience and critical infrastructure resilience: 1. Swift trust, 2. Public private partnerships, and 3. Quality information sharing 4. Subsequently, community resilience and resources resilience and hence resilience in a supply chain network may enable the achievement of sustainability

26. Schoenherr and Speier-Pero (2015)	• Large-scale survey involving supply-chain professionals on the costs and benefits of big data predictive analytics	The findings suggests that managers believe that leveraging big data analytics in the supply chain operations can result in: 1. Informed decision making capabilities; 2. Ability to improve supply chain efficiencies; 3. Enhanced demand planning capabilities; 4. Improvement in supply chain costs; 5. Increased visibility (planning)
27. Giannakis and Louis (2016)	• Multi-agent-based modelling of supply chain management system • Incorporates big data analytics	Findings show that the proposed model can potentially provide enhanced levels in each of the dimensions of supply chain agility of: 1. Flexibility, 2. Responsiveness, and 3. Speed
28. Chen et al. (2015)	• Survey data collected from 161 U.S.-based companies followed by CFA	Draws on dynamic capabilities theory to show that: 1. Organizational level BDA use has significant impacts on two types of supply chain value creation: **asset productivity and business growth** 2. The impacts of BDA use to supply chain value creation are moderated by environmental dynamism 3. Technological factors have a direct influence on organizational BDA use 4. Organizational readiness and environmental (i.e., competitive pressure) factors have an indirect influence on organizational BDA use through top management support 5. Links to firm performance and competitive advantage
29. Lee (2016)	• Cluster-based association rule mining applied to big data related to purchasing patterns of customers	1. In the context of omni-channel application of big data analytics helped to predict customer purchasing patterns.

		2. This allows for optimal solution for anticipatory shipping which in turn improves the customer experience thus a source of competitive advantage. (Forecasting)
30. Wu et al. (2016)	• Unstructured data • Big data analytics using social media, quantitative and qualitative data	Applying relevant analytical technique to unstructured and structured data can help firm identify areas of risks that may impact on the sustainability of the supply chain. This study found that deficiencies in following aspects as being more prone to risks: 1. Capacities – capacities for long-term prevention of shortages; Capital efficiency; Margin improvement; 2. Operations - Labour relations; Compliance with supply chain partners; Resource use and availability
31. Zhong et al. (2015)	• Structured Data analysis • Big Data Analytics applied to data collected from RFID-enabled intelligent shop floor environment.	• Big Data Analytics for RFID logistics in the context of RFID-enabled intelligent shop floor (smart-manufacturing) can help to optimise logistics planning and scheduling. (forecasting and planning)
32. van der Spoel et al. (2015)	• Survey of 230 truckers, • A data analysis and a data mining experiment, using real traffic and weather data.	1. Big data analytics was used to predict arrival time of truckers at distribution centres (but predictor power not as high as expected) 2. Provides information on how to optimise logistics in terms of synchronisation of coming and going cargo at distribution centres
33. Zhao et al. (2015)	• Scenario analysis in conjunction with • Big data analytics (data acquisition and quality control)	Applied in the context of green supply chain management. The results show that big data analytics could be used to determine: 1. The best scenario in terms of reducing risk and carbon emissions,

		2. Promoting improvements in green supply chain management, and 3. Achieving long-term commercial success
34. Hofmann (2015)	• System dynamics modelling • Simulation study involving three levers of big data; velocity, volume, variety	1. Data velocity, volume and variety all contribute to reduce the bullwhip effect in supply chains with velocity having the greatest impact.
35. Liu and Wang (2016)	• Structured data • Application analytic technique to big data collected from the sensor networks of chilled foods • Scenario analysis	• Application of big data analytics to sensor network data for a supermarket chain (from 4 of its retails stores) predicted more accurate product shelf-life information in real time compared to traditional method. • This enabled optimisation of pricing in a dynamic way as a function of deterioration in quality of the products. (supply chain efficiency)