BIG DATA AND FIRM PERFORMANCE

Vania Sena[1], Sumon Bhaumik[2], Abhijit Sengupta[3] and Mehmet Demirbag[4]

[1]Essex Business School, Department of Management Science and Entrepreneurship, Gateway Building, Colchester, Essex SS2 6ER, UK. Email: vsena@essex.ac.uk

[2]University of Sheffield Management School, Conduit Road, Sheffield S10 1FL, UK. Email: s.k.bhaumik@sheffield.ac.uk.

[3]University of Kent, Kent Business School, University of Kent, Canterbury CT2 7FS, UK. Email: a.sengupta-2408@kent.ac.uk.

[4]University of Essex, Essex Business School, Elmer Approach, Southend on Sea SS1 1LW, UK. Email:

mdemirc@essex.ac.uk.

1. Introduction

Big Data – or extremely large amount of structured and/or unstructured data continuously generated from a number of sources including documents, social media and sensors among others - is a very popular buzzword. The interest in Big Data started around 2011 with statements like "data is everywhere" or "data is the next oil" becoming ubiquitous. This interest has been fuelled by technological developments such as Internet of Things and Artificial Intelligence as well as by the drastic fall in the costs of the equipment for data processing and storage. Additionally, globalization implies that performance depends on businesses being agile and managers making the best possible decision in a timely manner. As a result, the deployment of business analytics (BA) – or advanced techniques to exploit large volumes of data - across companies has started to become common (McAfee and Brynjolfsson, 2012). Retailers have been very proactive in exploiting their data holdings to improve the customer experience[1] while in manufacturing and operations management, Big Data has supported automation and reduced inefficiencies along the supply chain (Davenport et al., 2012).

Eventually, the concept of Big Data has attracted the attention of scholars too. Academics have been exploring how the availability of Big Data has transformed businesses (e.g. Akter et al., 2016; Erevelles et al., 2016; Wamba et al., 2017). As a result, academic research has provided evidence suggesting that exploitation of routinely collected data is an essential element for firms seeking to gain competitive advantage (McAfee and Brynjolfsson, 2012). The use of business analytics has allowed management to assess the effectiveness of its activities through a "data lens" (Brands, 2014) as Big Data exploitation has become an important component of the decision-making process potentially at every level of the company (Davenport et al., 2012). Within businesses, managerial processes have changed as a result of data-driven decision-making strategy, which has led to changes in operations, human resource management and other management practices (Davenport, 2014). Marketing is another area where Big Data has made significant strides in terms of value capture: Big Data have been strengthening customer relationships, lowering management risk, improving operation efficiency which can lead to more effective marketing strategies (Kiron and Bean, 2013). Big data has then not only transformed business models and enhanced productivity but more importantly has encouraged the development of new strategic decision-making models based on data rather than on intuition (McAfee and Brynjolfsson, 2012).

Despite the growing recognition of the importance of big data across industries and sectors (Gandomi and Haider, 2015) and the mounting body of knowledge (McAfee and Brynjolfsson, 2012), there are two main challenges that this field of research faces: first, there is limited synthesis of the literature across the different management disciplines suggesting that an integrated approach to Big Data is necessary to enhance our understanding of the subject (see Amankwah-Amoah, 2016); second and more importantly, as the concept of Big Data has changed over time and its complexity has increased thanks to Machine Learning and Artificial Intelligence, value-creation from Big Data is still a major challenge for businesses. Indeed, data-driven management mechanisms that could work well in the first wave of Big Data may not

[1] See for instance the experience of Tesco and other large supermarket chains in the UK.

work so in the context where new forms of Big Data have emerged over the last few years.

Against this background, the purpose of this paper is to summarize the existing literature on how to capture value from Big Data that can be relevant both to managers and researchers. By doing so, we aim to identify current themes around Big Data and performance in the hope it will be used to lead future efforts in this area. Next, the paper reviews the existing definitions of Big Data and clarifies the relationship between business analytics and Big Data. The main conclusion from this section is that while there is no real consensus on the definition of Big Data, in reality the practice around Big Data and its exploitation is well-established and from the practice it is possible to identify key themes that guide us when trying to establish the key channels through which Big Data and analytics influence business performance. For these reasons, the paper presents two case studies illustrating how the deployment of analytics in two large companies have allowed them to create value and improve their performance. The two case studies allow us to identify the main channels behind Big Data value creation and will lead into the theoretical part of the paper which will discuss the main theoretical frameworks employed by researchers to discuss value creation driven by Big Data exploitation. Finally, the last section of the paper discusses the papers included in this special issue while some concluding remarks are offered in Section 5.

2. Big Data: Is there a definition?

It is important to start by pointing out that even if there is no consensus on what Big Data are, information systems scholars point out that Big Data have some unique features that do not make them equivalent to large databases. In addition, Big Data tends to be confused with Business Analytics that refers to the variety of analytical methods drawn from mathematics that can be used for the exploitation of large volumes of data.

From an operational standpoint, practitioners tend to define Big Data as high-volume, high-velocity and/or high-variety information assets (Beyer and Laney, 2012). Importantly, the Vs that are typically associated to the concept of Big Data, namely volume, variety and velocity, were used at an early stage of the development of this notion (see Beyer and Laney, 2012; Kwon et al., 2014). However, some authors suggest that veracity (referring to the biases, noises and abnormality of the data), validity (whether the data are correct and accurate) and volatility are important characteristics of Big Data. Some authors have pointed out that Big Data is a "moving definition" which varies with time as well as with industries (e.g., Manyika et al., 2011, pp. 1). As the volume of data increases over time, then setting up a threshold above which all data will be labeled as Big Data is not very helpful. Crucially, technological developments like Internet of Things or Artificial Intelligence (AI) have increased the variety of available data: these are collected from a variety of sources like web sites, smart devices and social media. Formats of data have changed as well, for example, unstructured data are now as common as structured data and most of these are produced real time suggesting that complexity is the most important feature of Big Data nowadays.

3. Big Data in Practice

As mentioned in the Introduction, the use of data mining technologies and tools in order to harvest and process information is becoming as ubiquitous as is the availability of big data sources themselves. In this section, we examine two cases of firms implementing and using such tools to enhance their dynamic response to the changing business environment. The two cases, covering different sectors, uses and technologies available, will be discussed briefly. The interested reader may follow the links and references for further details. Note that we do not cover firms whose core proposition and business models are based on big data and analytics (for instance Google, Uber, Amazon and numerous others), as the linkage between big data analytics and firm performance is largely tautological in these cases. Rather, the focus is on firms who offer a more traditional manufacturing or service oriented business model, and who have incorporated big data analytics and mining solutions as an input into their overall production process. Also, as numerous cases of use of big data analytics for customer retention and other marketing purposes have been documented elsewhere, including here in this special issue, we focus on alternative uses which have also created a big impact on performance.

Case 1: Fraud and loss detection by Jaeger

Jaeger is a mid-sized British fashion brand retailer, who implemented data mining technologies during the 2007-08 recession to reduce losses through leakages, such as shoplifting, internal theft, poor processes and wastage by suppliers[2]. Prior to the 2009-2010 recession, such leakages could be "tolerated", but an overall downward trend in the market forced the retailer to implement a centralized system which collected data from various internal systems, including "points of sale", warehouses, store alarms, etc. Given Jaeger's mid-sized IT budget, it decided to implement a relatively cheaper solution offered by a smaller solution provider, rather than the more expensive well known loss detections systems available in the market at that time. The workflow within this solution also involved human intervention in the form of expert queries and interventions to reduce instances of "false positives" in the alerts thrown up by the system. Jaegaer started making a return on investment within nine months of this project going live, and resulted in a significant improvement in margins and net profits (see Footnote 1).

In terms of barriers to implementing the datamining solution, the primary source of difficulty was the complexity of the overall organization and the incompatibility among the multiple data feeds from various store applications into a new centralized system. Thus, the barrier was technical, and not related to human resources or management, as the leadership was open to datamining and analytics based solutions from the outset. This is especially relevant and hints at a certain degree of flexibility within the decision making hierarchy of the company, given that such systems were not as popular in 2008 as they are now, and the awareness about these was also limited at that time. The choice of cheaper solutions, rather than market leading expensive ones, is also relevant and is yet another indicator of the flexibility within the company's processes. The willingness to dedicate additional resources in the form of experts dealing with false positives is also indicative of the same. Overall, given the backing received from the management, the flexibility within the organization, and the corresponding positive returns from investment, the implementation of the data analytics solution within Jaeger can be seen as a success story.

Case 2: Route optimization by UPS

---

2 Source:
(1) Datamining strikes stock shrinkage gold for Jaeger, Computerweekly February 2008. https://www.computerweekly.com/news/2240084951/Data-mining-strikes-stock-shrinkage-gold-for-Jaeger
(2) See Abdulrahman and Abdulaziz (2015).

United Postal Service or UPS is a well-known American multinational package delivery and logistics company, whose adoption of route optimization using big data and analytics has received a lot of attention among practitioners of analytics[3]. The rollout of route optimization by UPS has largely centred in its US delivery operations, which in itself is extremely complex. The company estimates that each of its drivers make on average 120 delivery stops per day, implying that the number of alternative ordering of stops for a single driver is in the order of 10^{198}. Optimizing a single such route, based on constraints of delivery times, road regulations and restrictions, is a massively complex task in itself, and it needed to be carried out jointly for 55000 different routes covered by UPS drivers in the US. It was estimated that a reduction of just one mile per driver each day, would save the company $50 million in the course of year. In addition, a reduction in idling time of 1 minute by each driver per day would save the company $14.6 million in a year.

UPS has implemented several IT solutions over two decades, including assistance for loading vehicles and sensors on vehicles to track their movement, but it is only recently that they implemented a holistic route optimization system, which the firm describes as “the highest level of analytics maturity” using a customized system named ORION (On Road Integrated Optimization and Navigation). ORION uses detailed mapping technology, integrated with route optimization algorithms and vehicle tracking technologies, making it one of the “world’s largest operations research projects” within the commercial domain.

The benefits from previous analytics projects and ORION are also apparent, both in terms of significant cost savings (estimated around 10 million gallons of fuel and over $300 million in costs), but also in terms of being able to offer additional services and enhancements to customers. UPS has also invested a significant amount in training its workforce in optimal use of the new technologies. The success of ORION in the US has led to its introduction in global

3 Sources:
(1) At UPS, the Algorithm is the Driver, The Wall Street Journal, 16 February 2015, https://www.wsj.com/articles/at-ups-the-algorithm-is-the-driver-1424136536
(2) Algorithm will Tell All UPS Drivers Where To Go, Supply Chain 24-7, 19 February 2015, https://www.supplychain247.com/article/algorithm_will_tell_all_ups_trucks_where_to_go/
(3) UPS Pressroom, 2 March 2015, https://www.pressroom.ups.com/pressroom/ContentDetailsViewer.page?ConceptType=PressReleases&id=1426329559785-791
(4) Unhappy Truckers and Other Algorithmic Problems, Nautilus, 18 July 2013, http://nautil.us/issue/3/in-transit/unhappy-truckers-and-other-algorithmic-problems

operations, leading to further cost savings. These efficiencies have increased its competitiveness against its rivals, both in the US and abroad.

The implementation of ORION has not been without severe teething problems. The primary problem was that the system was mechanical, while the solutions provided by it required human drivers for implementing them. This meant that in many cases, mathematically optimized solutions were difficult for drivers, who were creatures of habits and intuitions. Ironically, the drivers were able to come up with *better* solutions in many cases, than those which emerged when behavioural constraints were incorporated into ORIONS algorithms directly. Feedback from drivers who were given ORION optimized routes was often very negative and pointed towards an increasingly unhappy workforce. In the end, the solution lay in incorporating human expert intervention within the algorithms, and often involved looking for approximate *sub-optimal solutions* rather than the most optimized ones. The development and incorporation of ORION within UPS's processes was in the end, a very incremental and long drawn process, involving several iterations.

Discussion

The cases described above covered two different companies coming from entirely different sectors and significantly different in scale, with entirely separate end use for the analytics solution each implemented. In both cases, the solution resulted in high returns on the investment in the form of reduction in costs, thus impacting profitability positively. In case of the larger UPS, a customer facing analytics approach additionally led to an expansion in its service offerings, and hence can be viewed as a strategic investment, while in case of Jaeger, the analytics solution acted as a tactical response to adverse conditions. Both faced barriers to implementation, although the nature of the barriers were different, and a flexible approach helped in overcoming them and improving performance. While both firms can be classified as large, the use of big data and analytics based solutions need not be restricted to this class of firms only. Well known examples exist where SMEs have also reaped significant benefits these technologies, particularly in the areas of customer analytics and digital marketing[4].

4 Space constraints prevented us from discussing more examples involving SMEs, but interested readers may want to see the following:

4. Value-creation with Big Data: the theoretical perspective

4.1 Resource based view and dynamic capabilities

In this section, we will start exploring the ways in which the management literature has conceptualized the relationship between Big Data and performance and the main mechanisms for value creation from Big Data. It has been widely accepted that a firm's ability to sustain its competitive advantage lies in its ability to develop/acquire and harness resources that are valuable, rare, imperfectly imitable, and not substitutable by other resources to the same strategic end (Barney, 1991). This resource-based view (RBV) of a firm's competitive advantage has been discussed extensively in the literature, and the inimitability of strategic resources has received particular attention.

The early RBV literature explicitly recognized that access to a valuable, rare, imperfectly imitable and imperfectly substitutable strategic resource does not guarantee that the resultant competitive advantage will last forever. Barney (1991; pp. 103), for example, argues that

> …. that a competitive advantage is sustained does not imply that it will "last forever". It only suggests that it will not be competed away through duplication efforts of other firms. Unanticipated changes in the economic structure of an industry may make what was, at one time, a source of sustained competitive advantage, no longer valuable for a firm, and this not a competitive advantage.

An extension of this line of argument is that when firms experience unanticipated changes in the industry, whereby existing resources that provided the basis for a firm's competitive advantage are no longer as useful, it would be incumbent on the firm to identify ahead of its competitors the resources that would sustain its competitiveness in the new environment. This is exactly what Jaeger and UPS did (see Section 3), that is identified the key resources (bespoke data and algorithms) ahead of its competitors, enabling them to thrive in a new environment and get ahead of competition.

(1) https://www.imd.org/research-knowledge/articles/how-a-small-company-used-big-data-to-increase-its-sales/
(2) https://www.bluecorona.com/blog/big-data-opportunities-small-businesses/

The changing nature of competitive advantage, which is contingent on the nature of the product and the associated technology was recognized by proponents of the dynamic RBV (e.g., Helfat and Petraf, 2003) who argue that “the resource-based view must incorporate the evolution over time of the resources and capabilities that form the basis of competitive advantage” (pp. 998). One implication of this dynamic view of RBV is that organizations have to evolve with their product cycles because the nature of the resources that provide competitive advantage varies over product cycles. Put differently, the organizational capability of a firm has to change dynamically as industries, markets and products change, either organically or in response to exogenously driven disruptive changes. Helfat and Petraf (2003) argue that the six Rs of a firm’s capability life cycle are as follows: “retirement (death), retrenchment, renewal, replication, redeployment, and recombination” (pp. 2003).

This requirement for sustainability of competitive advantage, namely, that the capability of firms adapt to the environment in which they operate dynamically, has been enshrined in the dynamic capabilities (DC) approach to competitive advantage (See McKelvie and Davidsson, 2009). Following Teece, Pisano and Shuen (1997), DCs can be defined as “the firms’ capacities to integrate, build, and reconfigure internal and external resources/competences to address and shape rapidly changing business environments” (Katkalo, Pitelis and Teece, 2010; pp. 1177-1178). They “include the sensing, seizing, and transforming needed to implement a business model” that can give a firm a sustainable competitive advantage (Teece, 2018; pp. 43), e.g., the Power-by-the-Hour business model adopted by Rolls Royce for its jet engine business. However, the importance of first mover advantages is not obvious. While the early literature on the RBV suggests that a firm might benefit from first mover advantage (Wernerfelt, 1984; Barney, 1991), Teece argues that in the DC paradigm sometimes late adoption of new business models – e.g., late adoption of containerization by Maersk – may be to a firm’s advantage because standards and technologies may evolve and late adopters who wait until the relevant standards and technologies mature may be at an advantage. Instead, the emphasis is on changing the organizational culture of firms (e.g., Harreld, O’Reilly and Tushman, 2007).

Given the organizational culture of a firm, the sensing of new opportunities (and indeed threats) and the refining of business models while seizing these opportunities (Teece, 2018; Fig. 1) requires access to information (as UPS realized, once its routing algorithms were implemented successfully). The central role of information in providing competitive advantage has also been recognized in the context of banking (Boot, 2000) and risk management (Nocco and Stulz, 2006). In the management literature, the role of (smooth) information flow in efficient management of supply chains (Cachon and Fisher, 2000) and inventories (Wang and Toktay, 2008), as well as the role of information in shaping distribution of product sales across products (Brynjolfsson, Hu and Smith, 2010) – and eventually product design and pricing – have also been recognized (see Ambrosini et al., 2009). Unsurprisingly, it has been argued that "information management capability plays an important role in developing other firm capabilities for customer management, process management, and performance management" (Mithas, Ramasubbu and Sambamurthy, 2011; Easterby -Smith, and Prieto, 2008). While some of this information is internal to the firm, and can be harvested using widely available resources such as enterprise resource planning systems, other information such as those about customer preferences can be "bought for less than they are worth to the buyers [i.e., firms] because this may be considerably more than they are worth to the seller [e.g., customers]" (Katkalo, Pitelis and Teece, 2010; pp. 1176-1177).

Indeed, as discussed earlier in this paper, rapid technological development enables today's firms to harvest, assimilate a huge amount of data from a number of sources, and identify patterns within the "big data" (Chen, Chiang and Storey, 2012), with attendant implications for dynamic capabilities, competitive advantage and, by extension, firm performance (Groves, Kayyali, Knott and Kuiken, 2013; Braganza et al., 2017; Wamba et al., 2017). In the next section, we discuss the state of our understanding of how big data affects firm performance and the main channels.

4.2 Big Data and Performance

A keyword search of the main management databases reveals that several communities of management researchers have an interest in Big Data, such as operations management, supply chain management and marketing. A common theme across these different domains is the assumption that Big Data is a critical asset that is key to organizational success (Russom, 2011; Dutta and Bose, 2015). Typically, Big Data has been conceptualized as an IT capability i.e. the "firm's ability to mobilize and deploy Big Data-based resources in combination or co-present with other resources and capabilities" (Bharadwaj, 2000; Bhatt & Grover, 2005; Santhanam &

Hartono, 2003) which in turn is rooted in the RBV (Ryu & Lee, 2013; Zee & Jong, 1999). Thus, the main argument of the RBV is that competitive advantage is really driven by firm-level resources that cannot be imitated or easily substituted. Big Data is an example of such a resource as its deployment is quite unique to the business and its effectiveness will be conditioned by other resources like skills and infrastructure.

However, the RBV has been criticized on two counts: first, it failed to explain how Big Data resources were used to create competitive advantage (Kraaijenbrink et al., 2010) and second, it does not explain how the dynamic of resources evolves (Kraaijenbrink et al., 2010). As a result, a number of researchers have started to conceptualize the contribution of Big Data to business performance in terms of dynamic capabilities (DCV) (Teece, 2007). In particular the DCV has helped to determine how a Big Data-based competitive advantage can be achieved in a dynamic environment (as in the case of Jaeger, for example).

In these studies, Big Data is embedded in specific capabilities that build the business' competitive advantage. According to this approach, the value of Big Data is conditioned by the ability of the management to integrate it in the processes and routines. An important implication of the shift from the RBV to dynamic capabilities view is that researchers have focused on the distinct and measurable dimensions of the Big Data capabilities (Teece, 2007; Pavlou & El Sawy, 2010; Mikalef et al., 2016; Mikalef & Pateli, 2016). For example, Bhatt and Grover (2005) suggested that Big Data capability is characterized by value, heterogeneity, and imperfect mobility with the last feature being "necessary for sustained advantage". They further conceptualized three different types of capabilities: value capability (e.g., quality of Big Data infrastructure), competitive capability (e.g., quality of Big Data business expertise), and dynamic capability (e.g., intensity of organizational learning) in order to better understand the sources of Big Data-based competitive advantage.

Mikalef et al. (2016) suggest that the Big Data capabilities are really sensing, learning, coordinating, integrating and reconfiguring capabilities. In their framework, a coordinating capability is "the ability to orchestrate and deploy tasks and resources, and synchronize activities" (Pavlou & El Sawy, 2011). For instance, through social media and CRM, businesses can synchronise the activities of several departments and align their activities to the preferences of consumers. Similarly, coordination of activities across teams can be enhanced thanks to the use of real-time data (Vera-Basquero et al., 2013). An integrating capability includes the ability to integrate and exploit internal and external resources, ability that can be enhanced by the use of analytics which allows to gather information from multiple sources and assemble them in new ways (Woldesenbet et al., 2012). Finally, a reconfiguring capability is the ability to change its strategy and operations when external environment changes (Lin & Wu, 2014). In this case big data and business analytics are particularly relevant since they provide real-time data that can increase operational agility. In this sense, both Jaeger and UPS in our case studies, have sensed, learned, coordinated, integrated and reconfigured their capabilities in response to a challenging environment, using their data resources in a timely manner.

The use of the DCV allows researchers to go beyond the concept of Big Data resources and focus on the Big Data- enabled dynamic capabilities that directly contribute to business performance. For instance, CRM analytics can be considered a Big Data resource but it needs to be transformed into a dynamic capability that allow the business to improve its performance. This theoretical perspective allows to explore the contingencies that condition the relationship between capabilities and business performance. Indeed, the impact of Big Data-enabled capabilities on competitive performance has empirically been proven to be and indirect one, mediated by other organizational capabilities and contingent upon business strategy (Barreto, 2010).

5. Emerging Issues

The collection of papers in this special issue offers interesting insights into the current academic thinking around Big Data and business performance. An important starting point is the paper by Batistic and van der Laken (2019) which provides a very important overview of the extant literature on the organizational value of big data analytics using bibliometric techniques. The approaches used the authors (a co-citation analysis, an algorithmic historiography and bibliographic coupling) indicate the growing importance of this literature alongside the growth of the underlying technologies encompassing big data analytics. Their findings show that the foundations of this research area have been based primarily within ten research clusters, and progressed along two major independent research streams. These streams are (1) empirical: focussing on the use of statistics and algorithms for predictive finance and customer analytics;

and (2) strategic: focussing on the resource based view, dynamic capabilities and knowledge based frameworks to examine how big data analytics impacts a firm. Interestingly, the other papers in the special issue fall under one of the two research streams.

The papers from Frynas et al. (2019), Mikalef et al. (2019) and Wantao et al. (2019) contribute to research under the strategic stream but at the same time highlight the role of additional resources when exploiting Big Data. Indeed, Frynas et al. (2019) focus on multidisciplinary teams and use the RBV of the firm and on unique data collected from 240 big data experts working in global agrifood networks, examine the links between the use of big data-savvy (BDS) teams' skills, big data-driven (BDD) actions and business performance. The results suggest that the presence of multidisciplinary team enables firm to exploit their Big Data holdings that in turn leads to an improvement in performance.

Mikalef et al. (2019) examine the relationship between big data analytics capability (BDAC) and a key determinant of intangible firm resource, namely, innovation capabilities. Their research, which draws on RBV and the dynamic capabilities views, examines whether BDAC enhances both incremental and radical innovation capabilities. They note that a firm's ability to leverage its BDAC capability may be contingent on factors such as the firm's decision-making structure and organizational learning, which are influenced by the firm's organizational culture. Their empirical research uses survey data from a sample of 175 Greek firms. The empirical analysis suggests that a firm's ability to realise value from BDAC is influenced by complementary factors such as technical and managerial capability, as well as the presence of a data driven organisational culture.

Wantao et al. (2019) approach big data from perspective of Teece's (1986) complementary assets framework and conceptualises data driven supply chain orientation (DDSCO) as an intangible firm resource. Their paper focuses on "innovation focused complementary assets", and examines their role in the context of the Chinese manufacturing industry. Their empirical analysis demonstrates that while DDSCO improves financial performance of companies, the impact of DDSCO is more pronounced when a company's innovation focused complementary assets are high.

The paper by Amankwah et al. (2019) falls under the empirical stream. The papers uses social media text analytics to examine the impact of managerial responses to online reviews provided by customers, on returning customers' behaviour within the hospitality sector. Examining a large dataset composed of reviews of hotels, they examine the effect of managerial responses on returning customers' future responses. They show that "responding to returning customers' online reviews can be an effective mechanism for firms to…improve competitiveness…". The paper makes important theoretical and practical contributions to the literature by directly linking the use of a social media strategy to competitiveness. Equally Ding et al. (2019) use a number of Big Data methodologies to improve the performance of the existing volatility models and therefore the resilience of existing supply chains.

Finally, Gunersekaran et al. (2019) and Wang et al. (2019) attempt to answer the key question as whether there are additional theoretical perspectives that can be used to better understand the relationship between Big Data and performance. Gunesekaran et al. (2019) attempt to integrate three theoretical perspectives, namely, institutional theory, resource based view and big data culture in developing and testing a model examining the influence of big data predictive analytics on manufacturing performance. They argue that advancements in IT and related technologies diminish technology complexity; therefore the technology complexity may not be a hindrance any more in building manufacturing firms' analytics capability. Taking institutional view, authors further argue that inter-organizational pressure is a major influence in firms' BDPA adoption decision and they also identify big data culture playing significant moderating role that enhances the effects of tangible resources and human skills on BDPA adoption.

Wang et al. (2019) use the configuration theory and have developed a conceptual model of BDA success that aims to investigate how BDA capabilities interact with complementary organizational resources and organizational capabilities in multiple configuration solutions leading to improvements in performance. They use fuzzy-set qualitative comparative analysis to test their hypothesis in the context of Medicare and Medicaid Services. The results highlight the importance of skills and their complementarity with other resources when exploiting Big Data.

6. Conclusions

This special issue has explored the relationship between Big data and performance using a number of theoretical lenses, RBV and DCV being the common denominators. A common theme across all the papers is that the technological capability of exploiting Big Data needs to be complemented by a number of additional organizational resources that can eventually lead to enhanced performance. While it can be argued that this view is consistent with the RBV of the firm, in reality it goes well beyond RBV and highlights the importance of developing new theoretical frameworks to understand the relationship between Big data and performance.

In this respect, the use of the configuration theory in the context of Big Data and performance may be a useful avenue that may offer interesting insights on how resources need to be combined so that organizations can take advantage of their data holdings. More importantly, alternative theoretical frameworks can offer interesting insights as well. For instance, papers in this special issue have explored the role of culture and institutional theory in identifying the mechanisms that help organizations in exploiting their Big Data.

What are the future directions that research in Big Data can take? A theme that is not explored too much in the context of business performance is data ethics. While there is a lot of debate on data ethics and privacy in the context of re-use of Big Data for public policy, there is very little discussion of the implication of this debate on business performance. Clearly personalization may be an avenue to improve customer's experience and eventually business performance; at the same time, it also implies that businesses are allowed to collect large volumes of data at the expense of the customers' privacy. Limiting the amount of data businesses can store is a regulatory solution to this issue but at the same time, this is the type of regulation that can limit the capability of business to maximize their profits.

A second and related theme is that in most instances best practice around Big data exploitation has partially emerged. While up to decade ago there was no clear evidence of how/when organizations could benefit from Big Data, by now we know a lot about best practice in this field although mostly relegated to the specific types of Big Data. For instance, the use of social media to support the marking function is very well established and it is clear how they can be used to improve customers' experience. Still there are still areas though whether best practice has not emerged. For instance, although businesses store large volumes of data on supply chains, it not clear how these can be used to mitigate risks associated to breakdown of the supply chain. Of course, this does not mean that it may never emerge; it simply implies that more research is needed to better assess how Big Data can be used to for risk management.

References

Abdulrahman, A.R., Abdulaziz, A.R. (2015). Overview of business intelligence through data mining, in *Business Intelligence: Concepts, Methodologies, Tools, and Applications: Information Resources Management Association (IRMA)*, IGI Global: 49-72.

Akter, S., Wamba, S.F., Gunasekaran, A., Dubey, R., and Childe, S. J. Childe. How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, 182: 113–131, 2016.

Amankwah-Amoah, J., (2016), Emerging economies, emerging challenges: mobilising and capturing value from big data, *Technol. Forecasting and Social Change*, 110, pp. 167-174.

Amankwah-Amoah, J., Sheng, J., Wang, X and Khan, Z. (2019) Competing with big data analytics: Exploring online reviews and customers' ratings, British Journal of Management, this issue.

Ambrosini, V., C. Bowman and N. Collier(2009). 'Dynamic capabilities: an exploration of how firms renew their resource base', *British Journal of Management*, 20, pp. S9–S24.

Barney, J. (1991). Firm resources and sustained competitive advantage, *Journal of Management*, 17(1): 99-120.

Batistic, S. and Van der Laken, P. (2019), The history, evolution, and future of big data & analytics: a bibliometric analysis of its relationship to performance in organizations, British Journal of Management, this issue.

Boot, A.W.A. (2000). Relationship banking: What do we know? *Journal of Financial Intermediation*, 9: 7-25.

Braganza, A. et al. (2017). Resource management in big data initiatives: Processes and dynamic capabilities, *Journal of Business Research*, 70: 328-337.

Brands, K. 2014. "Big Data and Business Intelligence for Management Accountants." *Strategic Finance* 95 (1): 64–65.

Brynjolfsson, E., Hu, Y., Smith, M.D. (2010). Long tails vs. superstars: The effect of information technology on product variety and sales concentration patterns, *Information Systems Research*, 21(4): 737-747.

Cachon, G.P., Fisher, M. (2000). Supply chain management and the value of shared information, *Management Science*, 46(8): 1032-1048.

Chen, H., Chiang, R.H.L., Storey, V.C. (2012). Business intelligence and analytics: From big data to big impact, *MIS Quarterly*, 36(4): 1165-1188.

Davenport, T.H. (2014), How strategists use “big data” to support internal business decisions, discovery and production, Strategy & Leadership, Vol. 42 Issue: 4, pp.45-50

Davenport, T. H., P. Barth, et al. (2012), How Big Data Is Different, MIT Sloan Management Review 54(1): 43-46.

Ding, S., Zhang, Y. and Duygun, M. (2019), Modeling Price Volatility based on a Genetic Programming Approach, British Journal of Management, this issue.

Erevelles, S., Fukawa, N. and Swayne, L. (2016), Big data consumer analytics and the transformation of marketing, *Journal of Business Research*, 69(2):897–904.

Easterby□Smith, M. and I. Prieto (2008). ‘Dynamic capabilities and knowledge management: an integrative role for learning?’, *British Journal of Management*, 19, pp. 235–249.

Frynas, G., Akhtar, P., Mellahi,K. and Ullah, S. (2019), Big data-savvy teams’ skills, big data-driven actions and business performance, British Journal of Management, this issue.

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

Groves, A. P., Kayyali, B., Knott, D., Van Kuiken, S. (2013). The ‘big data’ revolution in healthcare: Accelerating value and innovation, *McKinsey Quarterly*, January, 1-19.

Dubey, R., Gunasekaran, A., Childe, S., Blome, C. and Papadopoulos, T. (2019), Impact of big data and predictive analytics on manufacturing performance: the effect of institutional pressures on strategic resources of the firm, British Journal of Management, this issue.

Harreld, J.B., O’Reilly III, C.A., Tushman, M.L. (2007). Dynamic capabilities at IBM: Driving strategy into action, *California Management Review*, 49(4): 21-43.

Helfat, C.E., Petraf, M.A. (2003). The dynamic resource-based view: Capability lifecycles, *Strategic Management Journal*, Special issue: Why is there a resource-based view? Toward a theory of competitive heterogeneity, 24(10): 997-1010.

Kiron, D., Prentice, P. K., Ferguson R. B. 2014. The analytics mandate. *MIT Sloan management review*, 55, 1-25.

Kraaijenbrink, J., Spender, J.-C., Groen, A.J. (2010). The resource-based view: A review and assessment of its critiques, *Journal of Management*, 36(1): 349-372.

Kwon, O., Lee, N., Shin, B., (2014), Data quality management, data usage experience and acquisition intention of big data analytics, *International Journal of Information Management*, 34 (3), 387-394.

McKelvie, A. and P. Davidsson (2009). ‘From resource base to dynamic capabilities: an investigation of new firms’, *British Journal of Management*, 20, pp. S63–S80.

McWilliams, A., Siegel, D.S. (2011). Creating and capturing value: Strategic corporate social responsibility, resource-based theory, and sustainable competitive advantage, *Journal of Management*, 37(5): 1480-1495.

Mikalef, P., Boura, M., Lekakos, G. and Krogstie, J. (2019), Big data analytics capabilities and innovation: The mediating role of dynamic capabilities and moderating effect of the environment, British Journal of Management, this issue.

Mithas, S., Ramasubbu, N., Sambamurthy, V. (2011). How information management capability influences firm performance, *MIS Quarterly*, 35(1): 237-256.

Nocco, B.W., Stulz, R.M. (2006). Enterprise risk management: Theory and practice, *Journal of Applied Corporate Finance*, 18(4): 8-20.

Petraf, M.A. (1993). The cornerstones of competitive advantage: A resource-based view, *Strategic Management Journal*, 14(3): 179-191.

Priem, R.L., Butler, J.E. (2001). Is the resource-based “view” a useful perspective for strategic management research? *Academy of Management Review*, 26(1): 22-40.

Rau, S.B. (2014). Resource-based view of family firms, In: Melin, L., Nordqvist, M., Sharma, P. (Eds.) *The SAGE Handbook of Family Business*, pp. 321-339.

Teece, D.J. (2018). Business models and dynamic capabilities, *Long Range Planning*, 51(1): 40-49.

Teece, D.J., Pisano, G., Shuen, A. (1997). Dynamic capabilities and strategic management, *Strategic Management Journal*, 18(7): 509-533.

Verbeke, A., Yuan, W. (2013). The drivers of multinational enterprise subsidiary entrepreneurship in China: A new resource-based view perspective, *Journal of Management Studies*, 50(2): 236-258.

Wamba, S.F. et al. (2017). Big data analytics and firm performance: Effects of dynamic capabilities, *Journal of Business Research*, 70: 356-365.

Wang, Y., Kung, L.A., Gupta, S. and Ozdemir, S. (2019), Leveraging Big Data Analytics to Improve Quality of Care in Health Care: A Configurational Approach, British Journal of Management, this issue.

Wang, T., Toktay, B.L. (2008). Inventory management with advance demand information and flexible delivery, *Management Science*, 54(4): 716-732.

Wernerfelt, B. (1984). A resource-based view of the firm, *Strategic Management Journal*, 5(2): 171-180.

Yu, W., Jacobs, M., Chavez, R. and Feng, M. (2019), Data-driven supply chain orientation and financial performance: The moderating effect of innovation focused complementary assets, British Journal of Management, this issue.