Configurations of Big Data Analytics for Firm Performance: An fsQCA approach

Completed Research

Patrick Mikalef
Norwegian University of Science and Technology
patrick.mikalef@ntnu.no

Maria Boura
Athens University of Economics and Business
mboura@aueb.gr

George Lekakos
Athens University of Economics and Business
glekakos@aueb.gr

John Krogstie
Norwegian University of Science and Technology
john.krogstie@ntnu.no

Abstract

With big data analytics growing rapidly in importance, academics and practitioners have been considering the means through which they can incorporate the shifts these technologies bring into their competitive strategies. Early empirical evidence suggests that big data analytics can enhance a firm's performance; yet, there is a lack of understanding on complementary organizational factors coalesce to drive performance gains, under what conditions they are more appropriate, as well as how they can complement a firm's dynamic capabilities under turbulent and fast-paced market conditions. To address this question, this study builds on the big data analytics capability literature and examines the fit between big data analytics resources and governance practices, dynamic capabilities, and environmental conditions in driving performance gains. Survey data from 175 chief information officers and IT managers working in Greek firms is analyzed by means of fuzzy set qualitative comparative analysis (fsQCA). Results show that that different configurations of resources, practices, and external factors coalesce to drive performance gains. We show that there are multiple configurations that can lead in high and low levels of performance.

Keywords

Big Data Analytics, Firm Performance, Information Governance, Environmental Uncertainty, fsQCA

Introduction

The value of big data analytics in directing organizational decision making has attracted much attention over the past few years. A growing number of firms are accelerating the deployment of their big data analytics initiatives with the aim of developing critical insight that can ultimately provide them with a competitive advantage (Constantiou and Kallinikos 2015). Prompted by the rapidly expanding data volume, velocity, and variety, there have been significant developments in terms of techniques and technologies for data storage, analysis, and visualization. Yet, there is a surprisingly large proportion of companies that are still struggling to realize value from their investments, and even some that argue that the cost of investing in big data analytics does not outweigh the performance gains (Arnott and Pervan 2016). A commentary by Sharma et al. (2014) argues that while there is some evidence that big data analytics can create value, the claim that such investments can be a source of competitive performance gains requires a deeper analysis. Several research editorials and survey papers stress the importance of conducting empirical studies to identify if and under what conditions big data investments can produce value, and lead to performance gains (Abbasi et al. 2016; Grover et al. 2018; Günther et al. 2017). Yet, to date, most studies have emphasized on infrastructure, intelligence, and analytics tools, while other related resources such as human skills and knowledge have been largely disregarded (Gupta and George 2016). Furthermore, the role of big data analytics while argued to be a central one in enhancing organizational decision-making, is often examined separately from the context in which such decisions are applied.

The business environment in which firms are required to operate is becoming increasingly more complex and turbulent, which urges them to develop dynamic capabilities to survive and thrive (Teece 2007). The role of big data analytics in enhancing dynamic capabilities and enabling organizations to compete under such conditions is also becoming increasingly more complex (Mikalef et al. 2019b; Wamba et al. 2017). Yet, there is very limited research on what big data analytics resources firms should invest in, how they should be orchestrated through appropriate governance practices, and how these differ based on the environment in which firms operate and compete (Abbasi et al. 2016; Mikalef et al. 2019a; Ramakrishnan et al. 2018). This is quite striking when considering that that largest challenge for most companies in realizing performance gains from their big data investments is not related to technology; rather, the largest hurdles are of an organizational nature and include leveraging big data analytics to support core organizational capabilities (Vidgen et al. 2017). New external conditions can quickly erode the strategic advantage firms have as turbulence increases. It is therefore important to examine the dynamics that evolve among big data analytics resources and governance practices, dynamic capabilities, and the external environment (El Sawy et al. 2010). Through this way it is possible to capture the complex interactions amongst these elements simultaneously. By adopting such a perspective, we will be in a better position to explain if, and under what conditions big data analytics can result in performance gains for firms.

To answer these questions, we build on a sample of 175 survey responses from IT managers in Greek firms and employ a configurational theory approach to examine the patterns of elements that lead to high levels of performance. We do so through the novel methodological tool fsQCA, which allows the examination of such complex phenomena and the reduction of solutions to a core set of elements. The rest of the paper is structured as follows. In section 2 we provide an overview of literature on big data analytics, dynamic capabilities and business value under turbulent conditions. In section 3, we introduce the methodology of the study, including the data, measurements, and reliability and validity tests, while in section 5 we present the results of the fsQCA analyses. In closing, we draw on the theoretical and practical implications.

Background

Big data analytics has been in the spotlight of attention for past few years, with many regarding such technologies as the next frontier for innovation, competition, and productivity (Manyika et al. 2011). Consequently, there has been much attention from both academics and practitioners on the value that organizations can create through the use of big data analytics. Despite the prominent role that big data analytics can potentially play in building a competitive advantage, particularly in fast-paced and turbulent business environments, the existing body of knowledge typically stops at examining the dyadic relationship between big data analytics and business value (Gupta and George 2016; Wamba et al. 2017). In fact, there are very few studies that examine the confluence of big data analytics under varying environmental conditions in shaping dynamic capabilities and realizing performance gains (Mikalef et al. 2018b). Research to date has acknowledged that big data analytics requires different complementary resources to produce value, which can be categorized in to technical big data infrastructure (i.e. the hardware, software and data), human resources and skills (including technical and managerial), as well as other relational resources such a focus on data-driven culture (Gupta and George 2016). Furthermore, other studies have recognized that resources such as the previously mentioned cannot yield any value if they are not manage appropriately leveraged (Saldanha et al. 2017), and doing so requires that firms implement governance practices that establish the roles, processes and structures around big data orchestration (Tallon 2013). These studies, as the ones revolved around big data analytics, underscore the importance of being able to leverage IT to achieve performance gains, and doing so necessitates a broader view towards other complementary resources (Saldanha and Krishnan 2011).

Yet, these two streams of research remain surprisingly disconnected, and even more importantly do not incorporate the conditions of the external environment when considering how organizations should build and leverage their big data analytics investments (Günther et al. 2017). Since novel technologies such as big data analytics are becoming increasingly more embedded in strategic operations of firms, it is important to examine not only what resources are important, but also what governance practices are put into place to mobilize them and orchestrate them effectively. This problem becomes increasingly more complex when taking into account that not all firms compete under similar conditions, and there are likely several ways in which investments can be focused and practices deployed to realize performance gains depending on the external environment (Popovič et al. 2018). The value of adopting a configurational approach to explore such tensions between IT, strategy and environment has been discussed in many IS paper over the past few years and most notably that of El Sawy et al. (2010). Nevertheless, when it comes to big data analytics and its use in directing strategic actions in turbulent and continuously changing business conditions there is a lack of work following such approaches. The argument made in this paper is that if we want to understand if, and under what conditions big data analytics can result in performance gains, it is important that we adopt a holistic approach when doing so going beyond investigating dyadic relationships (Woodside 2013). Hence, this study applies a configurational approach that seeks to uncover the combinations of big data resources and governance practices, the dynamic capabilities, and their confluence with environmental uncertainty in realizing performance gains.

Method

Data Collection

To explore the configurations of business analytics eco-dynamics that lead to high performance, a survey instrument was developed and sent out to key informants within firms. The survey-based approach is regarded as a pertinent method to accurately capture the level of sophistication of firm's resources and capabilities. According to Straub et al. (2004), the survey-based method is appropriate in exploratory settings and predictive theory. All the constructs used in this study and their corresponding items were based on previously published latent variables that had been assessed towards their validity. To develop the respective constructs, we utilized a 7-point Likert scale, which is regarded as an appropriate method where no standard measures exist for quantifying notions such as resources and capabilities (Kumar et al. 1993). To ensure the validity of the measures, a small-cycle pre-test study with 17 firms was conducted prior to the main study. The pre-testing allowed the assessment of the face and content validity of items. For the main study, a mailing list of 1500 Chief Information Officers and IT managers based in Greece was used. Invitations were sent out by emails followed by three reminders with a 2-week interval between each of them. To ensure that all questions were answered appropriately, respondents were instructed to ask other employees in their firms for information that they were not so knowledgeable about. The data gathering process lasts roughly three months (April 2017–July 2017), and the average completion time of the survey was 14 min. A total of 193 firms completed the survey, with 175 providing complete responses and the rest being to some extent incomplete. The firms that comprised our final sample operated in various industries, the largest of which was from the ICT sector (20.0%), followed by bank & financials (10.8%), consumer goods (9.7%), technology (9.1%), while a significant part came from other sectors (30.8%). Most of the companies were medium-sized firms, accounting for 30.2% of the sample, while high percentages were obtained from large-sized (26.2%) and small firms (24.0%). In terms of respondent profiles, most of them held senior positions relating to business and IT management as was originally planned for, while their experience with big data analytics differed, with most having at least 2 years prior engagement in such technologies.

	Sample (N=175)	Percentage (%)
Industry		
Bank & financials	19	10.8%
Consumer goods	17	9.7%
Oil & gas	5	2.8%
Industrials (construction & industrial goods)	13	7.4%
ICT and telecommunications	35	20.0%
Technology	16	9.1%
Media	13	7.4%
Transport	3	1.7%

Other (shipping, basic materials, consumer services etc.)	54	30.8%
Firm size (number of employees)		
1–9	34	19.4%
10–49	42	24.0%
50–249	53	30.2%
250+	46	26.2%

Table 1. Descriptives of the sample respondents

To investigate the possibility of non-response bias in our sample, the profiles of the respondents from the mailing list were used to gather information about the industry, type and size-class of the firm. After performing Chi-square analyses on these attributes between responding and non-responding firms, no systematic response bias was detected. In addition, late response bias was tested for by comparing early (first two weeks) and late responses (last two weeks) on the main constructs of the study. Outcomes confirmed that there was no statistically significant difference between the two sub-groups.

Measurements

The operationalizations for the constructs used in this study were adopted from prior literature and have therefore been previously tested in empirical studies. In Appendix A we provide a summary of the scales.

Technical Big Data Infrastructure was developed as a Type II second-order construct, comprising of two dimensions, technology (e.g. software and hardware), and data (Gupta and George 2016; Wamba et al. 2017). The items used to capture he two underlying dimensions were based on past empirical studies (Gupta and George 2016; Mikalef et al. 2019b).

Human Skills and Knowledge was captured as a Type II second-order construct. The two dimensions that form the construct include technical skills and managerial skills in accordance with past literature (Gupta and George 2016). Technical skills are concerned with the ability to handle the technological components and analytical requirements of big data, while managerial skills are revolved around recognizing the value of big data and understanding where to apply insight efforts (Akter and Wamba 2016).

Relational Resources was operationalized as a Type II second-order construct with the two underlying dimensions being the level of data-driven culture, and the propensity of organizational learning (Gupta and George 2016). A data-driven culture describes the level to which organizational members rely on insight derived from data analysis to make decisions (McAfee et al. 2012). Organizational learning refers to the concentrated efforts of firm members to exploit existing knowledge and explore new knowledge in order to keep up with unpredictable market conditions (Teece 2015).

Information governance is defined and developed in accordance to the study of Tallon et al. (2013 as different capabilities or practices for the creation, capture, valuation, storage, usage, control, access, archival, and deletion of information over its life cycle. Building on the framework of Peterson (2004, and on related work on information governance (Tallon 2013), three practices are identified and quantified. The information governance practices that are include in this study are structural, procedural, and relational practices. The scales for these constructs were based on past empirical studies (Mikalef et al. 2018a).

Dynamic Capabilities (DC) were developed as a Type II second order construct, comprised of five first order constructs (Jarvis et al. 2003). Thus, first-order constructs are theoretically distinct and contribute a unique component to the second-order construct. The first-order constructs that comprise a dynamic capability include (1) sensing, (2) coordinating, (3) learning, (4) integrating, and (5) reconfiguring routines, which are adapted from past empirical studies (Mikalef and Pateli 2017; Pavlou and El Sawy 2011).

The pressures of the external environment were captured through two constructs, market uncertainty and demand volatility (Newkirk and Lederer 2006). Market uncertainty is defined as the rate and unpredictability of environmental change. Demand volatility is defined as the availability of key resources and the level of competition in the external environment. The scales used to measure these constructs are based on the work of Newkirk and Lederer (2006).

Performance was developed a latent variable measuring profitability, market share, growth, innovativeness, cost leadership, and delivery cycle time in relation to main competitors (Liu et al. 2013; Rai and Tang 2010). These measures are representative of the potential value that can be realized as a result of leveraging big data resources and capabilities (Wamba et al. 2017). Performance was developed as a first-order reflective construct consisting of 10 indicators.

Measurement Model

Due to the fact that the model contains both reflective and formative constructs, different assessment criteria were used to evaluate each. First-order reflective latent constructs were assessed in terms of reliability, convergent validity, and discriminant validity tests. Reliability was gauged at the construct and item level. At the construct level we looked at Composite Reliability (CR), and Cronbach Alpha (CA) values to make sure that their values were above the threshold of 0.70 (Nunnally 1978). Indicator reliability was determined by examining if construct-to-item loadings were above the threshold of 0.70. Convergent validity was examined by looking if AVE values were above the lower limit of 0.50, with the lowest observed value being 0.56. To test for discriminant validity we looked at three different measures. The first examined if each constructs AVE square root in order to verify that it is greater than its highest correlation with any other construct (Fornell-Larcker criterion). The second tested if for each indicator the outer loading was greater that its cross-loadings with other constructs (Farrell 2010). Finally, we examined the Heterotrait-Monotrait Ratio of Correlations (HTMT) which is calculated based on the average of the correlations of indicators across constructs measuring different aspects of the model, relative to the average of the correlations of indicators within the same construct. All obtained values were below the threshold of 0.85 indicating sufficient discriminant validity. The abovementioned results (Table 2) suggest that first-order reflective measures are valid to work with and support the appropriateness of all items as good indicators for their respective constructs (Ruiz et al. 2008).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) Data	**n/a**															
(2) Technology	0.53	**n/a**														
(3) Managerial	0.54	0.49	**0.91**													
(4) Technical	0.45	0.57	0.68	**0.88**												
(5) Data-driven Culture	0.51	0.45	0.49	0.45	**0.87**											
(6) Organizational Learning	0.50	0.47	0.71	0.53	0.55	**0.94**										
(7) Structural practices	0.51	0.34	0.54	0.39	0.55	0.42	**0.90**									
(8) Procedural practices	0.50	0.55	0.67	0.50	0.54	0.67	0.58	**0.83**								
(9) Relational practices	0.42	0.48	0.61	0.49	0.46	0.51	0.67	0.56	**0.93**							
(10) Sensing	0.33	0.37	0.29	0.29	0.22	0.38	0.34	0.40	0.37	**0.80**						
(11) Coordinating	0.37	0.31	0.25	0.44	0.31	0.27	0.42	0.48	0.37	0.29	**0.88**					
(12) Learning	0.33	0.37	0.21	0.44	0.40	0.35	0.36	0.54	0.25	0.44	0.41	**0.91**				
(13) Integrating	0.19	0.36	0.12	0.23	0.24	0.31	0.18	0.58	0.27	0.34	0.36	0.50	**0.70**			
(14) Reconfiguring	0.35	0.43	0.35	0.34	0.35	0.39	0.36	0.50	0.52	0.43	0.50	0.42	0.38	**0.8**		
(16) Heterogeneity	0.25	0.44	0.31	0.27	0.45	0.43	0.33	0.37	0.29	0.29	0.22	0.33	0.37	0.27	**0.87**	
(17) Hostility	0.21	0.44	0.40	0.35	0.35	0.48	0.20	0.31	0.25	0.44	0.31	0.37	0.31	0.25	0.35	**0.81**
Mean	4.98	4.61	5.07	4.51	5.01	5.17	4.45	5.03	4.10	4.10	4.88	4.58	4.51	5.31	4.67	4.79
Standard Deviation	1.72	2.02	1.84	1.82	1.81	1.50	1.95	1.82	1.51	1.53	1.45	1.38	1.37	1.29	1.45	1.64
AVE	n/a	n/a	0.82	0.77	0.75	0.89	0.81	0.68	0.86	0.86	0.64	0.77	0.82	0.59	0.87	0.89
Cronbach's Alpha	n/a	n/a	0.93	0.90	0.83	0.96	0.76	0.88	0.84	0.92	0.72	0.85	0.81	0.71	0.91	0.86
Composite Reliability	n/a	n/a	0.95	0.93	0.90	0.97	0.89	0.91	0.92	0.95	0.84	0.91	0.93	0.78	0.92	0.89

Table 2. Assessment of reliability and validity of reflective constructs

For formative indicators and second-order formative constructs, we examined the weights and significance of their association with their respective construct. All first-order constructs and items had positive and highly significant effects. Next, to evaluate the validity of the items of formative constructs, we followed MacKenzie, Podsakoff, and Podsakoff (2011) guidelines using Edwards (2001) adequacy coefficient (R^2a). To do so we summed the squared correlations between formative items and their respective formative construct and then divided the sum by the number of indicators. All R^2a value exceeded the threshold of 0.50, suggesting that the majority of variance in the indicators is shared with the overarching construct, and that the indicators are valid representations of the construct. Next, we examined the extent to which the indicators of formative constructs presented multicollinearity, with Variance Inflation Factor (VIF) values of 3.3 being the cut-off threshold (Petter, Straub, & Rai, 2007). All values of first-order, second-order, and third-order constructs indicated an absence of mutlicollinearity.

Findings

To examine what configurations of big data analytics resources, information governance practices, environmental factors and dynamic capabilities lead to lead to high performing outcomes we employed a fuzzy-set Qualitative Comparative Analysis (fsQCA). FsQCA is a set-theoretic method that in based on Boolean algebra (i.e. set membership) to determine how configurations of elements are linked to specific outcomes. The technique follows the principles of complexity theories and allows for the examination of interplays that develop between elements of a messy and non-linear nature (Fiss 2011). The differentiating feature of fsQCA with other methods of analyzing data is that it supports the notion of equifinality. Equifinality means that a particular outcome (e.g. high levels of firm performance) may be caused by different configurations of elements, and that these configurations may differ depending on context. Such an approach is is particularly pertinent to the case of big data analytics since depending on the areas towards which analytics are applied, the aspects that are core contributors to firm performance may vary significantly (Abbasi et al. 2016). As such, it is important to identify the configurations of big data eco-dynamics that contribute to high performance, as well as to examine which are those that low-performing. Conducting analyses through FsQCA allows this since it is geared towards reducing elements for each configuration to the fundamentally necessary and sufficient conditions. Furthermore, fsQCA further supports the occurrence of causal asymmetry, which in a short means that for an outcome to occur, the presence and absence of a causal condition depend on how this causal condition combines with one or more other causal conditions (Fiss 2011).

The first step of performing the fsQCA analyses is to calibrate dependent and independent variables into fuzzy or crisp sets. Performance is set as the dependent variable of our study, while the independent variables that are used include those that fall under the categories of big data resources, information governance practices, dynamic capabilities and environmental factors. Fuzzy sets in this analysis range anywhere on the continuous scale from 0, which denotes an absence of set membership, to 1, which indicates full set membership. To calibrate continuous variables such as the ones we have utilized in the survey into fuzzy sets we followed the method proposed by Ragin (2009). Following this procedure, the degree of set membership is based on three anchor values. These include a full set membership threshold value (fuzzy score = 0.95), a full non-membership value (fuzzy score = 0.05), and the crossover point (fuzzy score = 0.50) (Woodside 2013). Since this study uses a 7-point Likert scale to measure all constructs, we follow the suggestions of Ordanini et al. (2014) to calibrate them into fuzzy sets. Following these guidelines, and based on prior empirical research (Fiss, 2011; Ragin, 2009), we computed percentiles for each construct so that the upper 25 percentiles serve as the threshold for full membership; the lower 25 percentiles for full non-membership; and the 50 percentiles represent the cross-over point.

Fuzzy set qualitative comparative analyses

To extract the configurations that lead to high and low performance outcomes we relied on the software fsQCA 3.0 (Ragin 2009). By performing two separate analyses, the fsQCA algorithm produces truth tables of 2^k rows, where k is the number of predictor elements, and each row indicates a possible combination. FsQCA then sorts all the 175 observations into each of these rows based on their degree of membership of all the causal conditions. As a result, some truth table rows contain several observations while others just a few or even none. During this step we reduce the number of rows according to two rules: (1) a row must contain a minimum number of cases, this value was set to a frequency threshold of 5 cases (Ragin 2009); and (2) selected rows must achieve a minimum consistency level of 0.80. Thus, configurations that do not adhere to this threshold are not included in the analyses. To obtain results we use the method proposed by Ragin and Fiss (2008) that identifies core conditions that are part of both parsimonious and intermediate solutions, and peripheral conditions that are eliminated in the parsimonious solution and only appear in the intermediate solution (Fiss 2011). Outcomes of the fuzzy set analyses for high and low levels of firm performance are presented in Table 3. The black circles (●) denote the presence of a condition, while the crossed-out circles (⊗) indicate the absence of it (Mikalef et al. 2015). Core elements of a configuration are marked with large circles, peripheral elements with small ones, and blank spaces are an indication of a don't care situation in which the causal condition may be either present or absent.

Configuration	**High-performing configurations**			**Low-performing configurations**	
	H1	H2	H3	L1	L2
Resources					
Technical Big Data Infrastructure	●		●	●	⊗
Human Skills and Knowledge	●	●	•		●
Relational Resources		●		⊗	
Information Governance					
Structural practices	●	●		⊗	
Procedural practices		●	●		
Relational practices	●	●		⊗	
Dynamic Capabilities	●	●			⊗
Environmental Factors					
Market Uncertainty	●	●	⊗	●	●
Demand Volatility		●			•
Consistency	0.882	0.932	0.912	0.902	0.887
Raw Coverage	0.312	0.198	0.182	0.103	0.097
Unique Coverage	0.134	0.073	0.086	0.065	0.059
Overall Solution Consistency	0.853			0.875	
Overall Solution Coverage	0.323			0.107	

Table 3. Configurations leading to high and low performance

The outcomes of the analysis for high performance yields three different solutions. The solutions present some common features but also some differences between them. Specifically, H1 and H2 correspond the external conditions that are characterized by turbulence since in H1 market uncertainty is a present condition, while in H2 adding to market uncertainty also demand volatility characterized the cluster of firms in this group. Under such conditions we find that the presence of dynamic capabilities is a core condition for achieving performance gains. Nevertheless, H1 and H2 differ in some ways also. While in configuration H1 the presence of strong technical big data infrastructure coupled with solid human skills and knowledge are seen as core conditions, in H2 the significance shifts from technical infrastructure to developing relational resources. This can be justified by the fact that in such conditions, investing in infrastructure and data can be seen as a commodity and the differentiating element is fostering strong relational links (i.e. by promoting a data-driven culture and organizational learning). Furthermore, in both cases establishing strong information governance practices is found to be a core component of high performance, but even more importantly in H2 where all three types of practices need to be established. On the other hand, we find that H3 also corresponds to firms that achieve high performance. Nevertheless, these firms are in much more stable conditions, where market uncertainty in absent and demand volatility doesn't play any significant role. These firms can achieve high performance even without the presence of strong dynamic capabilities. The requirements in terms of big data resources and information governance practices are also less highlighted, as the presence of strong technical infrastructure coupled with procedural practices is sufficient to drive performance gains, and the presence of strong skills is only seen as a peripheral (non-core) condition. When looking at low performing firms that have adopted big data analytics we find two unique configurations of conditions that explain such outcomes. In both cases L1 and L2, these companies operate in uncertain market conditions, and where there is not a presence of dynamic capabilities. Nevertheless, for L1 despite investments in technical infrastructure and data, the absence of relational resources, as well as the non-establishment of structural and relational practices when it comes to information governance results in low levels of performance. Respectively for L2, the presence of strong human skills and knowledge is insufficient in the absence of investments in technical big data infrastructure and dynamic capabilities. These solutions demonstrate that imbalances in ways firms invest in big data resources and a lack of alignment with organizational practices and capabilities can result in performance losses rather than gains.

Discussion

This study builds on the current situation that organizations face, where the increase of environmental turbulence, the flood of data, and the required speed of organizational change are creating a complex, high-paced, and messy business environment in which they have to navigate. The aim of this research is to build on the growing literature in the IS domain that adopts a more systemic and holistic perspective in examining technology-related phenomena drawing on a digital eco-dynamics perspective and applying a configurational approach (El Sawy et al. 2010). We therefore followed a survey-design study approach and collected data from 175 IT executives from Greece, and by analyzing data through a fsQCA approach demonstrated what configurations of conditions and relevant resources are capabilities need to co-exist in order to drive performance gains. Furthermore, we examined the configurations of conditions that result in low performance with the aim of outlining where organizations often fail with their big data adoption.

By examining such configurations, we add to the emerging literature on big data analytics and business value (Mikalef et al. 2017). From a theoretical perspective, the findings of this study add to literature by demonstrating how a configurational approach can be empirically explored in the context of big data analytics. Such ways of examining phenomena pertinent to big data are seldom investigated in quantitative studies. Furthermore, there have been a lot of anecdotal claims regarding the value that big data analytics can produce, yet, there is considerably less literature examining such claims and specifically looking into the specific conditions under which this can be achieved. The findings of our analysis showcase that investments in particular types of resources in isolation will most likely not yield any value, but rather it is the combination of how these resources are invested in that will produce performance gains. In addition, we demonstrate that apart from the pre-requisite of investing in resources, there is also a need to orchestrate resources effectively, and to be able to align them with organizational capabilities (Ramakrishnan et al. 2016). Our results highlight that in conditions that are considered as turbulent and uncertain, there is a need that such analytics are put into action by appropriate practices and that dynamic capabilities are fostered by firms. Big data analytics in this sense can help enhance the routines that comprise dynamic capabilities and promote evolutionary fitness by responding to changes in the external environment.

From a practical point of view, the results of this study can be used by managers to formulate different strategies around their big data analytics initiatives, depending on what type of environment their firm operates. In particular, our results showcase something that is often mentioned by technology consultants, that is that firms need to establish practices and develop approaches that break down organizational silos, so that data can be easily accessed and utilized from different departments. Furthermore, promoting such a data-culture requires that all departments are knowledgeable about the potential business cases that big data analytics can be applied to support and that they are co-creators of them during their design and implementation. Our findings highlight such effects, denoting the importance of establishing plans for adopting and diffusing big data analytics into operations.

Nevertheless, while the results of this research shed some light on the complex relationships that characterize big data resources, information governance, dynamic capabilities and the environment in the attainment of performance, they must be considered under their limitations. First, the sample of our analysis consists of companies operating in Greece. It is probable that firms that operate in other countries may have slightly different configurations of factors that drive performance. Second, while we examine performance outcomes, we do not control for the different types of performance indicators which can be more closely associated with impacts of analytics. The different functional areas in which big data analytics are applied are likely to yield varying configurations of resources to enhance or create business value. Third, although fsQCA allows us to examine the configurations of resources and the contextual factors under which performance is enhanced, the process through which it produces this outcome is not well explained. A complementary study suing a qualitative approach would likely reveal more insight on how value is produced from such investments.

Acknowledgments

This research has been partially funded by the Research Center of the Athens University of Economics and Business.

REFERENCES

Abbasi, A., Sarker, S., and Chiang, R. H. 2016. "Big Data Research in Information Systems: Toward an Inclusive Research Agenda," *Journal of the Association for Information Systems* (17:2).

Akter, S., and Wamba, S. F. 2016. "Big Data Analytics in E-Commerce: A Systematic Review and Agenda for Future Research," *Electronic Markets* (26:2), pp. 173-194.

Arnott, D., and Pervan, G. 2016. "A Critical Analysis of Decision Support Systems Research Revisited: The Rise of Design Science," in *Enacting Research Methods in Information Systems*. Springer, pp. 43-103.

Constantiou, I. D., and Kallinikos, J. 2015. "New Games, New Rules: Big Data and the Changing Context of Strategy," *Journal of Information Technology* (30:1), pp. 44-57.

El Sawy, O. A., Malhotra, A., Park, Y., and Pavlou, P. A. 2010. "Research Commentary—Seeking the Configurations of Digital Ecodynamics: It Takes Three to Tango," *Information Systems Research* (21:4), pp. 835-848.

Farrell, A. M. 2010. "Insufficient Discriminant Validity: A Comment on Bove, Pervan, Beatty, and Shiu (2009)," *Journal of Business Research* (63:3), pp. 324-327.

Fiss, P. C. 2011. "Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organization Research," *Academy of Management Journal* (54:2), pp. 393-420.

Grover, V., Chiang, R. H., Liang, T.-P., and Zhang, D. 2018. "Creating Strategic Business Value from Big Data Analytics: A Research Framework," *Journal of Management Information Systems* (35:2), pp. 388-423.

Günther, W. A., Mehrizi, M. H. R., Huysman, M., and Feldberg, F. 2017. "Debating Big Data: A Literature Review on Realizing Value from Big Data," *The Journal of Strategic Information Systems*).

Gupta, M., and George, J. F. 2016. "Toward the Development of a Big Data Analytics Capability," *Information & Management* (53:8), pp. 1049-1064.

Jarvis, C. B., MacKenzie, S. B., and Podsakoff, P. M. 2003. "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research," *Journal of consumer research* (30:2), pp. 199-218.

Kumar, N., Stern, L. W., and Anderson, J. C. 1993. "Conducting Interorganizational Research Using Key Informants," *Academy of management journal* (36:6), pp. 1633-1651.

Liu, H., Ke, W., Wei, K. K., and Hua, Z. 2013. "The Impact of It Capabilities on Firm Performance: The Mediating Roles of Absorptive Capacity and Supply Chain Agility," *Decision Support Systems* (54:3), pp. 1452-1462.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A. H. 2011. "Big Data: The Next Frontier for Innovation, Competition, and Productivity,").

McAfee, A., Brynjolfsson, E., and Davenport, T. H. 2012. "Big Data: The Management Revolution," *Harvard business review* (90:10), pp. 60-68.

Mikalef, P., Boura, M., Lekakos, G., and Krogstie, J. 2018a. "Complementarities between Information Governance and Big Data Analytics Capabilities on Innovation," in: *European Conference on Information Systems (ECIS)*. Portsmouth, UK: AIS.

Mikalef, P., Boura, M., Lekakos, G., and Krogstie, J. 2019a. "Big Data Analytics and Firm Performance: Findings from a Mixed-Method Approach," *Journal of Business Research* (98), pp. 261-276.

Mikalef, P., Boura, M., Lekakos, G., and Krogstie, J. 2019b. "Big Data Analytics Capabilities and Innovation: The Mediating Role of Dynamic Capabilities and Moderating Effect of the Environment," *British Journal of Management* (In press).

Mikalef, P., Framnes, V. A., Danielsen, F., Krogstie, J., and Olsen, D. H. 2017. "Big Data Analytics Capability: Antecedents and Business Value," *Pacific Asia Conference on Information Systems*, Langkawi, Malaysia.

Mikalef, P., Pappas, I. O., Krogstie, J., and Giannakos, M. 2018b. "Big Data Analytics Capabilities: A Systematic Literature Review and Research Agenda," *Information Systems and e-Business Management* (16), pp. 1-32.

Mikalef, P., and Pateli, A. 2017. "Information Technology-Enabled Dynamic Capabilities and Their Indirect Effect on Competitive Performance: Findings from Pls-Sem and Fsqca," *Journal of Business Research* (70), pp. 1-16.

Mikalef, P., Pateli, A., Batenburg, R. S., and Wetering, R. v. d. 2015. "Purchasing Alignment under Multiple Contingencies: A Configuration Theory Approach," *Industrial Management & Data Systems* (115:4), pp. 625-645.

Newkirk, H. E., and Lederer, A. L. 2006. "The Effectiveness of Strategic Information Systems Planning under Environmental Uncertainty," *Information & Management* (43:4), pp. 481-501.

Nunnally, J. 1978. "Psychometric Methods." New York: McGraw-Hill.

Ordanini, A., Parasuraman, A., and Rubera, G. 2014. "When the Recipe Is More Important Than the Ingredients: A Qualitative Comparative Analysis (Qca) of Service Innovation Configurations," *Journal of Service Research* (17:2), pp. 134-149.

Pavlou, P. A., and El Sawy, O. A. 2011. "Understanding the Elusive Black Box of Dynamic Capabilities," *Decision Sciences* (42:1), pp. 239-273.

Peterson, R. 2004. "Crafting Information Technology Governance," *Information systems management* (21:4), pp. 7-22.

Popovič, A., Hackney, R., Tassabehji, R., and Castelli, M. 2018. "The Impact of Big Data Analytics on Firms' High Value Business Performance," *Information Systems Frontiers* (20:2), pp. 209-222.

Ragin, C. C. 2009. "Qualitative Comparative Analysis Using Fuzzy Sets (Fsqca)," *Configurational comparative methods* (51).

Ragin, C. C., and Fiss, P. C. 2008. "Net Effects Analysis Versus Configurational Analysis: An Empirical Demonstration," *Redesigning social inquiry: Fuzzy sets and beyond*), pp. 190-212.

Rai, A., and Tang, X. 2010. "Leveraging It Capabilities and Competitive Process Capabilities for the Management of Interorganizational Relationship Portfolios," *Information Systems Research* (21:3), pp. 516-542.

Ramakrishnan, T., Khuntia, J., Kathuria, A., and Saldanha, T. 2016. "Business Intelligence Capabilities and Effectiveness: An Integrative Model," *System Sciences (HICSS), 2016 49th Hawaii International Conference on*: IEEE, pp. 5022-5031.

Ramakrishnan, T., Khuntia, J., Kathuria, A., and Saldanha, T. J. 2018. "Business Intelligence Capabilities," in *Analytics and Data Science*. Springer, pp. 15-27.

Ruiz, D. M., Gremler, D. D., Washburn, J. H., and Carrión, G. C. 2008. "Service Value Revisited: Specifying a Higher-Order, Formative Measure," *Journal of Business Research* (61:12), pp. 1278-1291.

Saldanha, T., and Krishnan, M. 2011. "Leveraging It for Business Innovation: Does the Role of the Cio Matter?," *International Conference on Information Systems (ICIS)*, Shanghai, China: AIS.

Saldanha, T. J., Mithas, S., and Krishnan, M. S. 2017. "Leveraging Customer Involvement for Fueling Innovation: The Role of Relational and Analytical Information Processing Capabilities," *MIS Quarterly* (41:1).

Sharma, R., Mithas, S., and Kankanhalli, A. 2014. "Transforming Decision-Making Processes: A Research Agenda for Understanding the Impact of Business Analytics on Organisations," *European Journal of Information Systems* (23:4), pp. 433-441.

Straub, D., Boudreau, M.-C., and Gefen, D. 2004. "Validation Guidelines for Is Positivist Research," *The Communications of the Association for Information Systems* (13:1), p. 63.

Tallon, P. P. 2013. "Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost," *Computer* (46:6), pp. 32-38.

Tallon, P. P., Ramirez, R. V., and Short, J. E. 2013. "The Information Artifact in It Governance: Toward a Theory of Information Governance," *Journal of Management Information Systems* (30:3), pp. 141-178.

Teece, D. J. 2007. "Explicating Dynamic Capabilities: The Nature and Microfoundations of (Sustainable) Enterprise Performance," *Strategic management journal* (28:13), pp. 1319-1350.

Teece, D. J. 2015. "Intangible Assets and a Theory of Heterogeneous Firms," in *Intangibles, Market Failure and Innovation Performance*. Springer, pp. 217-239.

Vidgen, R., Shaw, S., and Grant, D. B. 2017. "Management Challenges in Creating Value from Business Analytics," *European Journal of Operational Research* (261:2), pp. 626-639.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J.-f., Dubey, R., and Childe, S. J. 2017. "Big Data Analytics and Firm Performance: Effects of Dynamic Capabilities," *Journal of Business Research* (70), pp. 356-365.

Woodside, A. G. 2013. "Moving Beyond Multiple Regression Analysis to Algorithms: Calling for Adoption of a Paradigm Shift from Symmetric to Asymmetric Thinking in Data Analysis and Crafting Theory." Elsevier.

Appendix A. Survey Instrument

The survey items can be found in the following link: https://goo.gl/Y9aB5Q