



HAL
open science

Profiting from big data analytics: The moderating roles of industry concentration and firm size

Elisabetta Raguseo, Claudio Vitari, Federico Pigni

► To cite this version:

Elisabetta Raguseo, Claudio Vitari, Federico Pigni. Profiting from big data analytics: The moderating roles of industry concentration and firm size. *International Journal of Production Economics*, 2020, 229, pp.107758. 10.1016/j.ijpe.2020.107758 . hal-03032504

HAL Id: hal-03032504

<https://hal.science/hal-03032504v1>

Submitted on 30 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Profiting from Big Data Analytics: The Moderating Roles of Industry Concentration and Firm Size

Elisabetta Raguseo, Department of Management and Production Engineering, Politecnico di Torino, Torino, Italy;

Claudio Vitari, Department of Economics and Management, University of Aix-Marseille, Marseille, France;

Federico Pigni, Department of Management, Technology and Strategy, Grenoble Ecole de Management, Grenoble, France

Abstract

Big data has gained momentum as an Information Technology that is capable of supporting organizational efforts to generate new and better business value. We here contribute to the emerging literature on big data analytic (BDA) solutions by investigating the moderating roles of firm size and industry concentration in the relationship between BDA solutions and firm profitability. Using a unique panel data set that covers 13 years, from 2004 to 2016, which contains information about 176 firms, we provide robust econometric empirical evidence of the negative moderating effects of industry concentration and the positive moderating effects of firm size on the relationship between the use of BDA solutions and firm profitability. Our findings provide strong empirical evidence on the business value of BDA as well as the essential role played by contextual conditions that managers should consider.

Keywords

IT business value, big data analytics, firm profitability, econometric analysis, industry concentration, firm size.

Profiting from Big Data Analytics: The Moderating Roles of Industry Concentration and Firm Size

1 Introduction

Today, private and public organizations, in different sectors, are experiencing rapid changes in the needs, expectations and behaviors of their customers and, more generally, transformations in their industries and shifts in competitive forces. Either proactively or reactively, organizations turn to information technology (IT) to find solutions for their competitive issues (Oh et al., 2012; Shah & Shin, 2007). In this context, big data analytics (BDA), as a result of the unprecedented growth in data volume, variety and velocity, has recently been considered a new form of IT that is capable of supporting organizations in their efforts to advance new and better value propositions and business value (Huang et al., 2018; McAfee & Brynjolfsson, 2012; Müller et al., 2018). The current economic and market data reflect this trend. The market for BDA solutions is rapidly expanding and is expected to “reach \$260 billion in 2022, with a compound annual growth rate (CAGR) of 11.9 per cent over the 2017-2022 period,” according to International Data Corporation (IDC, 2018).

BDA solutions are “a holistic approach to managing, processing and analyzing the 5V data-related dimensions (i.e. volume, variety, velocity, veracity and value) to create actionable insights to deliver sustained value, measure performance and establish competitive advantages” (Fosso Wamba et al., 2015). Therefore, BDA solutions alter the way through which firms operate, work and create customer value. This potential raises new questions regarding whether and how big data contribute to value creation and competitive advantages, and which factors influence or determine the effects of big data (Sena et al., 2019).

The importance of understanding the impact of BDA on firm profitability emerges from the extent of its impact, compared to other well researched IT systems. BDA solutions change both how knowledge is exchanged between and within companies and how companies can extract value from data (e.g., Fosso Wamba et al., 2015; Baesens et al., 2016; Mikalef et al., 2020).

Although evidence supports the relationship between BDA investment and superior organizational performance (see the Appendix), the accumulated knowledge still returns a heterogeneous and contingent representation of the problem (Huang et al., 2018; Mikalef, Boura, et al., 2019b; Sena et al., 2019). Multiple theoretical perspectives have been used to explore this relationship. Among them, many studies have investigated the relationship through the lenses of resource-based, dynamic capabilities, resource orchestration or absorptive capacity (Sharma et al., 2014; Wamba et al., 2017; Grover et al., 2018; Shamim et al., 2019; Zhang et al., 2019; S. Wang et al., 2019; Mikalef, Krogstie, et al., 2019; Mikalef et al., 2020), and have shown the suitability of these perspectives. However, only a limited number of studies have proposed an assessment of business value in terms of the organizational performance (Božič & Dimovski, 2019; George et al., 2014; Müller et al., 2018; Raguseo & Vitari, 2018) or the competitive advantage (Mikalef et al., 2020; Mikalef, Krogstie, et al., 2019) and measured them with objective financial archival data.

Indeed, not all organizations benefit from big data investments in the same way and to the same extent (Hindle et al., 2020). Although the substantial amount of accumulated knowledge identifies numerous complementary organizational resources and conditions that affect BDA business value, they are largely based on the underlying assumption that organizations overall are homogenous in the context in which they operate (Mikalef, Boura, et al., 2019b). Only a few studies have explored these contextual conditions, and the available literature has mainly concentrated on environmental dynamism (Chen et al., 2015; Côte-Real et al., 2020), measured in the form of the environmental turbulence or uncertainty that is typical of DCV (Mikalef, Boura, et al., 2019b; Mikalef et al., 2020). Other contextual conditions that have been advanced, but which remain underexplored, include industry, firm size, organizational structure, and top management support (Mikalef, Boura, et al., 2019b).

However, the current assessments of industry influences have been directed toward the effects of big data investments on firm productivity, considering IT-intensive industry and competitiveness

(e.g., Müller et al., 2018) and labor markets, where similar investments by other firms have helped the development of workers with complementary technical skills (Tambe, 2014). However, productivity is not profitability, and the former does not always translate into the second (Hitt & Brynjolfsson, 1996). Productivity is the production of more output for a given quantity of inputs. Profitability considers the ability to gain competitive advantage and earn higher profits than the company would have earned otherwise (Hitt & Brynjolfsson, 1996).

Hence, a richer and broader theorizing of these contextual determinants could have important implications on guiding managers' assessments of whether investments in BDA generate business value and whether the returns are different according to the contextual conditions in which companies operate (Mikalef, Boura, et al., 2019b). Among the many contextual determinants that have still only been explored to a limited extent in BDA studies, two are particularly susceptible to playing significant roles in mediating organizational performances: company size (Oliveira & Martins, 2011) and industry concentration (F.M. Scherer & Ross, 1990).

With this study, we aim to contribute to the development of a richer and broader theorizing of the contextual determinants of the profitability of BDA solutions by analyzing the influence of industry concentration and company size. These two variables represent internal and external conditions that could influence the economic returns of company investments. Accordingly, we address the following research question: *“Do contextual conditions, namely, company size and industry concentration, impact the relationship between the use of BDA and firm profitability?”*.

Although our results will, first of all, identify the conditions under which companies benefit and profit from BDA (Gupta & George, 2016a; Müller et al., 2018; Sivarajah et al., 2017), we will also contribute to the rich stream of research that studies the impact of specific IT system investments on firm performance, which includes, for example, enterprise resource planning (ERP) (e.g., Mangin et al., 2015), supply chain management (SCM) (e.g., Dehning et al., 2007), knowledge management

systems (KMS) (Feng et al., 2005) and customer relationship management (CRM) (e.g., Müller et al., 2018)

We have grounded our work on the resource-based view (RBV) of the firm (Barney, 1996) and on the theory of industrial organization (F.M. Scherer & Ross, 1990), thereby supporting the moderating roles of firm size and industry concentration, respectively, on firm profitability. Our study addresses the research gap mentioned above by using robust econometric models. We have based our analysis on an original panel dataset that contains information on the adoption, or not, of BDA solutions in 176 companies over 13 years (from 2004 to 2016). We combined this information with financial data obtained from the Bureau Van Dijk DIANE database to estimate the signs of the moderating effects of company size and industry concentration on the relationship between BDA and firm profitability. We found that industry- and firm-specific effects matter. Larger firms as well as firms in less concentrated environments are capable of profiting more from BDA solution investments.

The remainder of the paper is structured as follows. First, we present the theoretical background in which the RBV of a firm is integrated with the theory of industrial organization, and we formulate our hypotheses. We then describe the methodology that we followed in detail and present our results. We continue with a discussion of the findings, the limits of the research, and our conclusions and guidelines for future studies.

2 Theoretical foundations and research hypotheses

To understand the impact of BDA on firm profitability and to investigate the moderating effects of contextual variables, the research framework and research hypotheses are rooted in two main theories: the Resource Based View (RBV) of a firm (Barney, 1996) and the theory of industrial organization (Frederic M. Scherer & Ross, 1990; Tirole, 1988). The reasons behind the choice of these theories are discussed hereafter.

2.1 The effect of BDA solutions on firm profitability

RBV (Barney, 1996) explains the extent to which firm resources and capabilities in general, and BDA solutions in particular, contribute to firm performance (Wamba et al., 2017). We consider that BDA, as a holistic approach, gathers resources and capabilities that are valuable, rare, inimitable and sustainable to achieve a potential competitive advantage (Barney, 1996). Superior firm performance on competitive markets can help firms achieve greater success than their direct competitors (Davenport, Thomas H., 2006; Peteraf & Barney, 2003). This approach is aligned with the theorization of BDA as developing capabilities that involve mobilizing big data resources, such as big data assets, the analytics portfolio and the available human talent. These assets are levered to build capabilities, which are applied to generate organizational value and impact (Grover et al., 2018).—In order to address the identified research gap, we extend the current framework to incorporate the contributions of BDA solutions to firm performance and explore industry- and company-specific moderating factors.

Several studies have already investigated (see the Appendix) the effects of big data resources on firm performance (e.g., Krishnamoorthi & Mathew, 2018; Mikalef et al., 2020), and have demonstrated that the effect of these resources on firm performance depends on the solution scale and time horizon (Matthias et al., 2017; G. Wang et al., 2016; Y. Wang et al., 2018). Scale concerns the extent of the big data application and may be narrow when covering a single activity of a broader process (e.g., a product recommendation within the purchasing process) or wider when entailing entire business domains (e.g., the whole supply chain). The time horizon is the temporal orientation of the solution which can alternatively focus on historical data, with a descriptive orientation (e.g., auditing solutions), or on the present, with a predictive orientation (e.g., real-time trading tools). Finally, a BDA solution can have a future, and hence prescriptive, orientation, as in the case of strategic decision-support systems (Gunasekaran et al., 2017). The possible combinations of scales and time horizons explain why BDA solutions can have the potential to

provide business value for the most diverse activities of any organization (Fosso Wamba et al., 2015; Hindle et al., 2020, 2020; Lycett, 2013; Y. Wang et al., 2018).

Financial performance is used extensively as a dependent variable when measuring a firm's competitive advantage (Kaufman, 2015). Other evidence has already shown the ability of BDA solutions to positively impact financial performance (Akter & Wamba, 2016; Božič & Dimovski, 2019; Chen et al., 2015; Fosso Wamba et al., 2015; Mikalef, Krogstie, et al., 2019; Wamba et al., 2017). The results of these studies show that, for example, BDA solutions can improve Returns on Investments (ROI) for retailers or e-commerce by helping to complete the purchase process. Accordingly, we adopt profitability-based variables as synthetic indicators of the effect of BDA solutions on firm performance.

2.2 The moderation of firm size

Firm size is a significant moderator of firm performance in RVB studies. Firm size can change the degree to which certain postures, structures, and tactics boost firm performance in consideration of different strategic goals (Covin et al., 1994). More specifically, firm size is a variable that can affect how firms invest in and profit from IT (Oliveira & Martins, 2011). A variety of justifications, in line with the RBV of a firm, such as resource bases, scale and scope economies, pre-emptive move capabilities, formalization levels, decentralization patterns, specialization trends and innovativeness levels (Dong & Yang, 2020; Eisenhardt & Martin, 2000; Forés & Camisón, 2016; Kirca et al., 2011), have been advanced to explain the moderating effect of firm size on the relationship between IT investments and firm performance.

In contrast to this variety of justifications, other studies have instead shown opposite moderation effects of firm size on firm performance. In general, larger organizations exhibit higher degrees of differentiation and formalization, more decentralized managerial decision-making authority systems, higher levels of task specialization, more complex forms of communication and greater organizational inertia. Furthermore, a high level of bureaucracy in larger firms makes them slower

in making strategic moves, struggle longer to respond to changing conditions (Eisenhardt & Martin, 2000) and less innovative (Forés & Camisón, 2016; Wagner et al., 2012). Overall, large firms benefit less than small firms from the same type of IT investment.

A similar debate is taking place as far as BDA solutions are concerned. Some aspects of BDA solutions favor smaller companies, while other features are favorable for larger organizations (Dong & Yang, 2020; Mikalef, Boura, et al., 2019b). Moreover, some previous results did not show any difference on the basis of firm size (Bughin, 2016). For example, BDA as a Service is advanced as particularly beneficial for smaller companies, as it grants radical innovation possibilities, flexibility, efficiency and filling, and eventually overcomes the complementary resource gap with larger companies (Ebner et al., 2014). On the other hand, stronger IT capabilities and advanced information systems, better training and more secured data management seem to explain why larger companies could be better off taking advantage of big data (Kamioka & Tapanainen, 2014).

Overall, although we believe that some smaller companies are better placed than larger organizations, smaller companies are in general less prepared and equipped to benefit from BDA. Hence, on average, larger organizations could ultimately be in a more favorable position to combine the external resources available on the market with internal big data assets to create valuable, rare, inimitable and sustainable resources and capabilities (Ebner et al., 2014). Therefore, we expect big data analytic solutions to have a stronger effect on the profitability of larger firms.

H1: The positive effect of BDA solutions on firm profitability is lower in small firms than in large ones.

2.3 The moderation of industry concentration

While RBV can explain the extent to which firm resources, in general, and BDA solutions, in particular, contribute to firm performance, the theory of industrial organization can be used to

analyze the differences in performance that stem from the structural characteristics of industries, such as industry concentration.

The theory of industrial organization is grounded in the structure-conduct-performance paradigm. This paradigm assumes that an industry's structural elements impact the conduct of firms operating in that industry, and ultimately explains the differences in the performance of the firms in that industry and of the industry as a whole (F.M. Scherer & Ross, 1990). The theory of industrial organization helps us understand the importance, for a firm, of the contextual variables of the market structure. This theory indicates that firms must find the right fit between their internal characteristics and the characteristics of the external environment to improve their organizational performance (Melville et al., 2007). Hence, the good combination of endogenous mechanisms with external variables could help firms achieve a competitive advantage (Burns & Stalker, 1994; Thompson et al., 1992).

One prominent characteristic of the industrial organization is the industry concentration (Porter & Advantage, 1985). The more an industry is concentrated around a small number of players, the weaker the competitive pressure. The literature shows that competitive pressure is generally associated with higher organizational efficiencies (F.M. Scherer & Ross, 1990). Empirical evidence has demonstrated, for example, that the effects of human resource management systems on labor productivity are moderated by the degree of the industry's capital intensity, industry growth and industry differentiation (Datta et al., 2005).

Empirical studies have also shown that the industry of a firm has an increasingly important influence on IT resource investment and the subsequent performance of IT (Chiasson & Davidson, 2005; Schryen, 2013). Like the broad management and economics literature, the information systems literature recognizes that industry concentration is a significant contextual moderator of the impact of IT on firm productivity (Melville et al., 2007) and on competitive advantage (Li & Richard Ye, 1999). Indeed, among the various contextual variables, industry concentration emerges

as one of the major moderators of firm performance (Bain, 1951; Belman & Heywood, 1990; Donsimoni et al., 1984; Levy, 1985).

The theory of industrial organization can be used to explain performance differences when the same type of IT is used across competitive regimes. Under competitive pressure, firms become more innovative by leveraging on their IT assets (Basole et al., 2013; Melville et al., 2007). The absence of competitive pressure allows firms to build up slack and other inefficiencies yet still be able to stay in business (Melville et al., 2007). Therefore, the same kind of IT assets can be more efficiently used in highly competitive industries than in industries with less competition.

In the BDA domain, a specific research framework, based on the strategic business value of the BDA solution (Grover et al., 2018), advances a structure to help us identify which of an industry's competitive dynamics moderate the effects of BDA solutions on performance. Although environmental dynamism has largely been explored (e.g., Chen et al., 2015; Mikalef, Boura, et al., 2019b; Wamba et al., 2019), other aspects have been analyzed less. Empirically, industry competitive intensity has been observed to be positively linked to the use of BDA (Malladi & Krishnan, 2013). Similarly, industry hostility, such as the availability of key resources and the level of competition, is recognized as having an influence on the relationship between BDA and firm performance (Mikalef, Boura, et al., 2019b). According to the theory of industrial organization, structural characteristics, such as industry concentration, could play a moderating role in the relationship between big data investments and firm profitability. Since few studies have investigated the moderating role of contextual factors and because several moderators can influence the relationship between big data and performance, we expect that BDA solutions have a greater effect on firm profitability when the competitiveness of the industry is high, because the industry concentration is low.

H2: The higher the level of industry concentration is, the lower the contribution of BDA solutions to firm profitability.

Figure 1 summarizes the research framework and the two hypotheses investigated in this study.

--- Figure 1 around here ---

3 Methodology

3.1 Data collection

We administered our questionnaire to medium- and large-sized French firms to evaluate the impact of the big data business value on firm performance. As our study considers the effects at the firm level, we followed previous studies that had targeted the Chief Information Officer (CIO) as the main informant. We implemented a random sampling method to select the medium and large French companies we wanted to interview from a population of 19,875 medium and large companies from the sectors shown in Table 1 and registered in the Bureau Van Dijk DIANE database, which is one of the main sources of financial information on firms in France. We aimed to gather 200 questionnaires, thus ensuring a 95% confidence level and a 6.9% confidence interval that our sample would reflect the entire population. Firms were categorized by size on the basis of their revenues and according to the European definition where small companies generate less than €10 million, medium-sized firms between €10 and €50 million, and large companies more than €50 million.

The questionnaire consisted of two sections. The first section, which all the companies answered, assessed the presence or lack of BDA solutions. This discrimination consisted of triangulating various questions to assess the velocity, variety and volume of a firm's big data. We included a question for each dimension of the big data Vs in the questionnaire.

The second section assessed the model variables and was made available to those companies that were found to have adopted BDA solutions. We included scales based on the previously published multi-item scale variables (Gregor et al., 2006; Ji-fan Ren et al., 2016; Mithas et al., 2011; Vorhies & Morgan, 2005) in this section. In the second section, we double checked whether the

companies had adopted a BDA solution: visual analytics software, scripting languages, in-memory analytics software, MapReduce and Hadoop software, machine learning software, natural language processing, social media analytics software and/or predictive analytics software.

Specifically, in order to obtain 200 compiled questionnaires, we conducted a pilot study on the first 30 companies that answered the questionnaire, from a subsample of the population of 142 companies (response rate of 21.13%) to test the comprehensibility of the questions, to identify possible response issues and to establish the expected response rate. All the questions led to appropriate answers and thus did not require any further changes. Therefore, the final questionnaire remained unchanged.

After the pilot test, another 170 companies, from 1,962 randomly selected companies from our population of 19,875 medium and large companies (response rate of 10%), took part in the survey. The data gathering process involved three steps. In the first step, we contacted the company to inform them about the aim of the research study and to ask permission to contact the CIO. In the second step, the CIO was contacted and asked about his/her willingness to participate in the survey. When the CIO was not available at the time agreed upon in the first call, we made a second appointment. Therefore, the questionnaire was completed, in either the second or third step, according to the availability of the CIO. When the CIO was unable to answer the questionnaire, we identified another qualified respondent who was knowledgeable about the firm's investments and the adoption of BDA solutions. Overall, 200 questionnaires, 30 from the pilot test and 170 from the main survey delivery phase, were completed (Table 1).

In short, since financial data were not available for 24 companies, we built a unique panel dataset that contained information about the adoption, or not, of the BDA solutions of 176 companies over a period of 13 years (from 2004 to 2016). We combined data gathered through the survey with financial data obtained from the Bureau Van Dijk DIANE database. In this way, we were able to

estimate the signs of the moderating effects of company size and industry concentration on the relationship between BDA and firm profitability.

--- Table 1 around here ---

3.2 Measures

3.2.1 Dependent variable

Profitability. We assessed profitability using multiple economic performance measures. We measured the impact of BDA on profitability by considering the annual change in return on sales (ROS), the return on assets (ROA) and the return on equity (ROE) as the dependent variables (e.g., Geringer et al., 2000).

3.2.2 Independent variables

Big Data Analytics. The BDA variable refers to the use of at least one of the eight BDA solutions identified in the survey conducted to double check whether the companies had adopted a BDA solution and to establish their experience over the years, which was measured with the following question: “*How many years ago did you first start using these technologies?*”. This variable was made equal to 1 if the company had used at least one of the eight BDA solutions in the considered year and 0 otherwise.

Industry concentration. Industry concentration measures the extent to which industry output is produced by a few firms. This variable is widely used in the literature as the inverse proxy of competitiveness (Bharadwaj et al., 1999; Porter & Sakakibara, 2004). The industry concentration data were computed using economic data obtained from the DIANE database. The variables were lagged to show the effect of the industry concentration of the previous year on the economics of the year after. We computed the Herfindahl-Hirschman index (HHI) for each sector, since it is a

commonly accepted measure of market concentration (Porter & Sakakibara, 2004). This index is calculated as the sum of the squared market share of each company that competes in a sector. Its value can range from zero to 10,000. The measure can then be expressed as follows: $HHI = s_1^2 + s_2^2 + s_3^2 + \dots + s_n^2$ (where s is the integer market share of each firm in the sector). Specifically, we computed the HHI value for each industry covered by our sample (for a total of 48 industries defined at the two-digit level NAF code) following a three-step approach. First, we downloaded data from the DIANE database for each industry: the list of all the companies that belong to the considered industry as well as the value of their revenues from 2004 to 2016. Second, we computed the squared market share of all the companies in each industry for every year by dividing the revenues of the considered company with the total revenues of the sector the company belongs to. This step was performed for each company and every year. Third, we computed the HHI for each sector by summing the market share of the first 100 companies with the highest market shares every year.

Firm size. In line with the assumption that revenues are a proxy of company size (Qian & Lee, 2003), the sales revenue value, in millions of euros, was included in the logarithmic form in each model to measure the size of the company.

3.2.3 Control variables

Age. This variable represents the age of a firm and is measured as the logarithmic form of the actual age of the firm, beginning from the year it started its operations (Autio *et al.*, 2000).

Years. This measure controls for the year we referred to in each observation. It was operationalized for each company through twelve dummy variables, one for each year from 2005 to 2016 (Melville *et al.*, 2007). To simplify the readability of the models, we omitted the coefficients of these variables in the table that shows the regression results.

Firm risk. This variable is measured as the standard deviation of the accounting based returns (*Kim et al., 1993*). We used a 3-year average starting from before the year to which the observation referred

Past performance. This variable is measured by considering the average of all the accounting-based returns for the three years before the year to which the observation refers. Past performance may affect the availability of slack resources and therefore the risk-taking ability (*Autio et al., 2000*). When ROA was used as a dependent variable in the model, we included the average of ROS and ROE; in the case of ROE, we included the average of ROA and ROS; and in the case of ROS, we included the average of ROE and ROA.

Dynamism. We followed the approach used by previous authors (*Stoel M.D., & Muhanna, W.A., 2009*) to operationalize the environmental factors, and we used industry data from the DIANE database (Bureau Van Dijk), which includes financial measures for all French companies. To do so, we considered the industries using the NACE classification at the two-digit level. We measured dynamism by considering turbulence in the distribution of revenues within each industry using firm-level data from DIANE. Specifically, we operationalized the dynamism of an industry as the variability in annual industry sales: the total industry-level sales for 5 years were regressed on the year variable for each sector, and dynamism was measured as the standard error of the regression slope coefficient of the annual industry sales divided by the industry mean of the 5-years period.

Table 2 synthesizes the operationalization of the variables used in this study.

--- Table 2 around here ---

3.3 The empirical model

The hypotheses were tested using a unique panel dataset of 176 French companies over a thirteen-year period from 2004 to 2016. We tested the moderating effects of the two moderation variables, that is, industry concentration and firm size, on the relationship between the use of BDA solutions and firm profitability, and for each measure of profitability. The independent variables were lagged

to account for potential simultaneity bias (Melville et al., 2007). The independent variables used to compute the interaction effects were standardized, since we included the interaction variable to evaluate the moderation effects in the models.

The nine equations used to verify the moderating effects of industry concentration and firm size on the ROS, ROA and ROE have the following form:

$$ROS_t = a_1 + b_1 \text{Big_data_analytics}_{t-1} + b_2 \text{Industry_concentration}_{t-1} + b_3 \text{Firm_size}_{t-1} + b_4 \text{Big_data_analytics}_{t-1} * \text{Firm_size}_{t-1} + b_5 X_t + \varepsilon_t \text{ (Equation 1)}$$

$$ROS_t = a_2 + b_6 \text{Big_data_analytics}_{t-1} + b_7 \text{Industry_concentration}_{t-1} + b_8 \text{Firm_size}_{t-1} + b_9 \text{Big_data_analytics}_{t-1} * \text{Industry_concentration}_{t-1} + b_{10} X_t + \varepsilon_t \text{ (Equation 2)}$$

$$ROS_t = a_3 + b_{11} \text{Big_data_analytics}_{t-1} + b_{12} \text{Industry_concentration}_{t-1} + b_{13} \text{Firm_size}_{t-1} + b_{14} \text{Big_data_analytics}_{t-1} * \text{Industry_concentration}_{t-1} + b_{15} \text{Big_data_analytics}_{t-1} * \text{Firm_size}_{t-1} + b_{16} X_t + \varepsilon_t \text{ (Equation 3)}$$

$$ROA_t = a_4 + b_{17} \text{Big_data_analytics}_{t-1} + b_{18} \text{Industry_concentration}_{t-1} + b_{19} \text{Firm_size}_{t-1} + b_{20} \text{Big_data_analytics}_{t-1} * \text{Firm_size}_{t-1} + b_{21} X_t + \varepsilon_t \text{ (Equation 4)}$$

$$ROA_t = a_5 + b_{21} \text{Big_data_analytics}_{t-1} + b_{22} \text{Industry_concentration}_{t-1} + b_{23} \text{Firm_size}_{t-1} + b_{24} \text{Big_data_analytics}_{t-1} * \text{Industry_concentration}_{t-1} + b_{25} X_t + \varepsilon_t \text{ (Equation 5)}$$

$$ROA_t = a_6 + b_{26} \text{Big_data_analytics}_{t-1} + b_{27} \text{Industry_concentration}_{t-1} + b_{28} \text{Firm_size}_{t-1} + b_{29} \text{Big_data_analytics}_{t-1} * \text{Industry_concentration}_{t-1} + b_{30} \text{Big_data_analytics}_{t-1} * \text{Firm_size}_{t-1} + b_{31} X_t + \varepsilon_t \text{ (Equation 6)}$$

$$ROE_t = a_7 + b_{32} \text{Big_data_analytics}_{t-1} + b_{33} \text{Industry_concentration}_{t-1} + b_{34} \text{Firm_size}_{t-1} + b_{35} \text{Big_data_analytics}_{t-1} * \text{Firm_size}_{t-1} + b_{36} X_t + \varepsilon_t \text{ (Equation 7)}$$

$$ROE_t = a_8 + b_{36} \text{Big_data_analytics}_{t-1} + b_{37} \text{Industry_concentration}_{t-1} + b_{38} \text{Firm_size}_{t-1} + b_{39} \text{Big_data_analytics}_{t-1} * \text{Industry_concentration}_{t-1} + b_{40} X_t + \varepsilon_t \text{ (Equation 8)}$$

$$ROE_t = a_9 + b_{41} Big_data_analytics_{t-1} + b_{42} Industry_concentration_{t-1} + b_{43} Firm_size_{t-1} + b_{44} \\ Big_data_analytics_{t-1} * Industry_concentration_{t-1} + b_{45} Big_data_analytics_{t-1} * Firm_size_{t-1} + \\ b_{46} X_t + \varepsilon_t (Equation 9)$$

where X_t is a set of control variables that could influence the profitability of a company.

4. Analysis and results

4.1 Descriptive statistics

Table 3 shows the descriptive statistics of the variables, including the minimum and maximum values, the means and standard deviations, and the Spearman correlation matrix. We can draw three main conclusions from an analysis of the correlation coefficients. First, the size of the company is positively and significantly correlated with the usage of BDA solutions. This positive link implies that larger companies invest more in big data solutions since they possess significantly more resources to commit to these types of investments. Second, the age of the company is not correlated with the adoption of the investigated BDA solutions, and the correlation coefficient is not statistically significant (Table 3). This result supports the idea that new and established firms do not invest in different ways in these solutions. Finally, another interesting finding has emerged concerning the positive correlation between the variable that refers to the year and BDA solutions. As expected, this correlation is positive and statistically significant, thus demonstrating an increasing adoption of these solutions in recent years.

--- Table 3 around here ---

Table 4 shows the development of the BDA diffusion rate in our dataset over time and split into HHI intervals and firm size (revenues as a proxy). In alignment with the recent trend which shows that the market for BDA solutions is rapidly expanding, the BDA diffusion rate increased from 2004 to 2016, from 8.4% in 2004 to 35.8% in 2016, that is over a period of thirteen years of observation.

When comparing the diffusion rate of high and low market concentrations, as well as larger or smaller firms, some differences emerge as statistically significant while others do not, although the majority of them are not different, especially for the last years of observation. We conducted an analysis of variance (ANOVA) to verify the significant differences, using F-tests to statistically test the equality of means.

--- Table 4 around here ---

4.2 Panel data analysis

It is possible to control for variations in random or fixed effects in the error terms when using panel data by considering systematic time-series effects. Random or fixed effects can be used, since potential deviations from the assumptions of an ordinary least squared regression can occur as unaccounted effects across firms and across time. For example, not accounting for managerial practices may lead to bias in the models, as their effects could be attributed to BDA or other regressors. This effect may lead to different values and therefore to an incorrect assessment of the impacts of BDA (Melville et al., 2007).

A Hausman specification test was performed to determine the suitability of random versus fixed-effects models and to establish the appropriateness of the random effects model compared to that of the fixed-effects model by computing an X^2 test to determine whether the random component of the variance model was appropriate. The Hausman test indicated that the fixed effect was the most appropriate estimation method for our case. Accordingly, nine fixed-effects regression models were run with robust standard errors to verify the hypotheses. We used STATA 14.0 software to perform the calculations.

4.3 Findings

4.3.1 Regression findings

Table 5 shows the results obtained from our models when both hypotheses were tested (H1 and H2). We followed a five-step approach to test the hypotheses. First, we verified the risk of multicollinearity. Second, we ran three fixed-effects regression models (Model 1, Model 4 and Model 7) with robust standard errors to test the first hypothesis. We included the interaction term between the use of BDA solutions and the firm size level in Model 1 to test the effect of this interaction on the ROS of the companies in our sample. We tested the same interaction on the ROA level in Model 4 and, in Model 7, considering ROE. Third, we ran three additional fixed models (Model 2, Model 5 and Model 8) to test Hypothesis 2. We included the interaction term between the use of BDA solutions and industry concentration in Model 2 to test the effect of this interaction on the ROS of the companies in our sample. We tested the same interaction on the ROA level in Model 5 and, in Model 8, considering ROE. Fourth, we tested the full models. We included the effects of all the interaction terms on profitability, as proxied by ROS, ROA and ROE, in Model 3, Model 6 and Model 9. Finally, we here discuss how we addressed the endogeneity issue, which can be a major concern when evaluating the impact of IT on profitability.

--- Table 5 around here ---

Although the low correlation coefficients among the variables hinted at a limited risk of multicollinearity, we computed the variance inflation factors (VIFs) to further exclude any potential multicollinearity problems that may have arisen due to the correlation between the variables included in the models. Since the variables had adequate VIFs (between 1.026 and 2.839); that is, well below the suggested threshold of 10 (Hair et al., 1998), multicollinearity was not a problem for these analyses.

Hypothesis 1 proposed that the positive effect of BDA solutions on firm profitability is higher in larger firms than in smaller ones. As shown in Models 1, 4 and 7, the interaction term between BDA solutions and firm size was positive and significant for all the considered types of profitability. This result means that larger companies may expect higher returns from investing in

BDA solutions than smaller ones. For this reason, it is possible to state that Hypothesis 1 was confirmed.

Hypothesis 2 proposed that the higher the level of industry concentration is, the lower the contribution of BDA solutions to firm profitability. The results of the interaction effects of Models 2, 5 and 8 show that the interaction effect between BDA and firm concentration is negative and statistically significant. Therefore, the higher the concentration of the market is, the lower the contribution of BDA to firm profitability. Since industry concentration can be used to measure the extent to which industry output is produced by a few firms, and is widely used in the literature as an inverse proxy for competitiveness (e.g., Melville et al., 2007), the investment in BDA solutions in competitive environments enables firms to achieve higher levels of profitability. For this reason, we can say that Hypothesis 2 was confirmed.

As previously mentioned, we tested the full model in Models 3, 6 and 9, thereby obtaining further confirmation of the two hypotheses. These models confirmed the previously found moderating effects.

Endogeneity could have been a concern in the previous models, which included the lagged independent variable as a control for endogeneity. However, like many estimations from empirical studies in the IT value literature, the estimate presented in the previous table can be subject to endogeneity concerns, including omitted variable bias and reverse causality. Although many papers use econometric approaches to eliminate such problems (Tambe & Hitt, 2012), biased profitability estimates, due to the difficulty of finding adequate instruments, remain a concern. We used fixed-effects regression models to address the endogeneity concerns, but time-varying factors and IT investments could still have imposed an upward bias on the estimates (Tambe, 2014). One of the concerns that can arise is that unobserved firm-level factors, such as the managerial capabilities of firms, can determine an upward bias on the coefficient estimate of BDA. Although these biases are

difficult to remove completely, the following paragraphs discuss the causal interpretation of the estimates presented above on the basis of the results of robustness tests (Tambe, 2014).

An explanation that favors reverse causality or simultaneity would be consistent with a pattern of estimates in which higher profitability firms systematically make BDA investments in the subsequent year. In order to test this aspect, we ran a fixed-effects regression model in which the dependent variable was the use of BDA solutions and the independent variable was the profitability for the year before the observation of the level of adoption of such technologies:

$$Big_data_analytics_t = a_{10} + b_{47} ROS_{t-1} + b_{48} Industry_concentration_{t-1} + b_{49} Firm_size_{t-1} + b_{50} X_t + \varepsilon_t$$

(Equation 10)

$$Big_data_analytics_t = a_{11} + b_{51} ROA_{t-1} + b_{52} Industry_concentration_{t-1} + b_{53} Firm_size_{t-1} + b_{54} X_t + \varepsilon_t$$

(Equation 11)

$$Big_data_analytics_t = a_{12} + b_{55} ROE_{t-1} + b_{56} Industry_concentration_{t-1} + b_{57} Firm_size_{t-1} + b_{58} X_t + \varepsilon_t$$

(Equation 12)

where X_t is a set of control variables that could influence the profitability of a company.

The results in Table 6 support the absence of reverse causality or simultaneity on BDA technology investments and company profitability, since the effect of the value of the profitability at time $t-1$ is not significant for the big data technology investment at time t .

Therefore, we have demonstrated that endogeneity was not a concern in our models.

--- Table 6 around here ---

We conducted a post-hoc analysis to support the linear relationship between BDA and the performance measures. Specifically, we investigated whether there was an inverted U-shaped relationship or S shaped relationship, instead of a linear relationship, by running the following regressions (referring to one year only, since we only had information on the number of BDA

solutions adopted for just one period of time, namely 2016; X_t is a set of control variables that could influence the profitability of a company).

$$ROS_t = a_{13} + b_{59}Big_data_analytics_{t-1} + b_{60}Big_data_analytics_{t-1}^2 + b_{61}Industry_concentration_{t-1} + b_{62}Firm_size_{t-1} + b_{63}X_t + \varepsilon_t \text{ (Equation 13)}$$

$$ROA_t = a_{14} + b_{64}Big_data_analytics_{t-1} + b_{65}Big_data_analytics_{t-1}^2 + b_{66}Industry_concentration_{t-1} + b_{67}Firm_size_{t-1} + b_{68}X_t + \varepsilon_t \text{ (Equation 14)}$$

$$ROE_t = a_{15} + b_{69}Big_data_analytics_{t-1} + b_{70}Big_data_analytics_{t-1}^2 + b_{71}Industry_concentration_{t-1} + b_{72}Firm_size_{t-1} + b_{73}X_t + \varepsilon_t \text{ (Equation 15)}$$

As shown in Table 7, there is no inverted U-shaped relationship or S shaped relationship in the way BDA impacts economic performance, since the squared coefficient of the BDA on economic performance is not statistically significant in any of the three models. Accordingly, these results support the linear relationship hypothesized and investigated in this study.

--- Table 7 around here ---

5. Discussion

This study has made an important empirical contribution to the literature on the business value of BDA. We are the first to have analyzed the influence of industry and company factors on BDA solution performance, and to have chosen objective measures of firm profitability as the dependent variable. Indeed, the majority of previous works investigated this topic by applying a qualitative research methodology (e.g., Mikalef, Boura, et al., 2019b) or by choosing a self-assessed perceived impact (e.g., Côte-Real et al., 2020) of BDA solutions. We have been able to conduct our study by triangulating data gathered through a survey and time-series financial data from the Bureau Van Dijk DIANE database for a twelve-year period (2004-2016).

Our findings highlight that firm size is a significant moderator of firm profitability when BDA is leveraged on, thereby confirming Hypothesis 1 and extending the current body of knowledge

pertaining to the linking of big data initiatives with performance. In addition to the somewhat contrasting results and theories available in the literature about the moderating effect of firm size on firm performance (Covin et al., 1994), we offer additional evidence to the still open debate about when firms invest in IT in general (Oliveira & Martins, 2011) and BDA specifically (Dong & Yang, 2020; Mikalef, Boura, et al., 2019b). We conclude that, in terms of BDA solutions, larger firms are better equipped to seize the potential of BDA than smaller firms. The results are far from being trivial, as our data are not based on short-term data. The point of the discussion in fact lies in better understanding the reasons for this significance. The mediation role of firm size clearly proxies a more complex mechanism. This finding, within the context of RBV and the existing literature on BDA, suggests that larger firms possess the complementary (Yu et al., 2019) resources necessary to leverage on BDA and maintain these long-lasting, rare, inimitable and non-substitutable organizational value resources. The limited resources of smaller firms, ~~on the other hand~~ conversely, result in the observed lower capacity to transform the BDA potential into financial profits. Interestingly, this difference was not found to be related to the measured dynamism of the economic environment, partially contradicting previous results (Mikalef, Boura, et al., 2019b; Mikalef et al., 2020), thus suggesting that the organizational characteristics associated with firm size play critical roles. In this sense, and coherently with the theoretical framework, the success of BDA solutions on firm profitability may depend on the availability of resources. Complementarity may then emerge, because of the wider pool of resources large firms can leverage on. In addition, dynamic capabilities for deploying the resources can be key complements of the firm's resource-picking ability in the value creation process by BDA (Grover et al., 2018; Makadok, 2001). We then provide a different argument on dynamism where more than the speed of change, the value of BDA lays in its accumulation or growth of the available resources (Mikalef et al., 2020).

A similar explanation could also be advanced for industry concentration.

Our results demonstrate that the relationship between BDA solutions and company profitability is moderated to a great extent by the level of industry competitiveness, thereby confirming Hypothesis 2. A higher level of competitiveness seems to make big data-related investments generate better results than what they produce in highly concentrated industries. The findings we obtained, by adopting an alternative empirical model, confirm previous observations (Melville et al., 2007; Mikalef, Boura, et al., 2019b). Our results suggest that policies that ease competition may contribute to profitability when BDA solutions are levered on. A possible explanation may be related to the differentiating strength that BDA, and more generally IT-based innovation, has in the absence of large market leaders. Interestingly, competitiveness does not erode the appropriability of the benefits resulting from BDA solutions, even in the long term. Although we cannot determine whether this finding is idiosyncratic for BDA solutions, it suggests a more articulated role of IT investments than what has traditionally been accepted (Brynjolfsson, 1996). The findings of this study point out that the availability of resources, both internal and external to an organization, is a critical factor for the success of BDA initiatives. At an organizational level, larger firms that have access to a wider pool of resources show a higher performance, thanks to BDA. At the industry level, a concentrated market, [that could proxy a matured or declining market](#) (Vitari & Raguseo, 2019), leaves fewer resources available, thereby limiting the success and effects of BDA initiatives. Therefore, the understanding of the mechanism adopted to pool, accumulate and orchestrate these resources is of great importance, as here demonstrated by their contextual significance.

Overall, our study highlights the relevance of industry- and company-specific effects in explaining the business value of big data. Our effort joins the existing literature on industry- (**Chen et al., 2015; Mikalef, Boura, et al., 2019a, 2019b; Mikalef et al., 2020; Mikalef, Krogstie, et al., 2019; Wamba et al., 2019**) and company- (Akter et al., 2016; Côte-Real et al., 2020; Dong & Yang, 2020; Dubey et al., 2019; Krishnamoorthi & Mathew, 2018; Seddon et al., 2017; Yu et al., 2019) specific effects that have been considered to explain the business value of big data. In conclusion,

although the overall idea that BDA impacts the competitiveness of companies is not new, our original specification, operationalization and findings together provide a richer and more detailed picture of how big data impact organizational performance.

From a managerial point of view, this study has many implications. Managers, before making BDA investments, should include both industry-specific features and organizational characteristics in their evaluations. BDA returns on profitability in fact differ according to the level of concentration of the industry as well as of the firm size. The observed limited effect of dynamism may imply a stronger focus for managers in making the necessary pool of strategic resources for building a competitive advantage based on BDA. Managers should then focus on the [capabilities](#) necessary [for](#) accumulating and [deploying these resources](#) and in selecting those that most contribute to the [value creation process](#). The ability to accumulate rare and inimitable data sources and deploy BDA by leveraging the complementarity at organizational and industry level has therefore emerged as an additional requirement for decision makers. This discussion suggest then to managers to attentively consider the context and contingencies in which they operate. This study supports managerial practice by shedding light on the determinants of the profitability of BDA solutions and policy making and by providing the first rationale for the incentivization of BDA investments. In particular, we advise small and medium enterprise managers that the publicized benefits of BDA investments could be lower, and may even be unsatisfactory, for smaller companies. However, enterprises that experience strong competitive pressure could find some relief, at least temporarily, from this pressure, by investing in BDA.

6. Limitations and future research

This study enriches the literature on the business value of BDA and the contextual conditions that matter in such relationships. We confirm the results of previous studies on BDA and have observed the positive impact of BDA on performance and the significant moderating role of industry specificities ([Huang et al., 2018](#); [Mikalef, Boura, et al., 2019a, 2019b](#); [Mikalef et al., 2020](#);

Raguseo & Vitari, 2018). Although we provide evidence on the relationship between BDA and firm profitability, we cannot exclude the influence of other indirect and unobserved effects. Even though our findings can be supported by existing studies, further studies should consider extending the nomological network to identify and study a larger “context-contingent set of synergistic combinations of IT and other organizational resources” (Melville et al., 2007).

This study is robust from a methodological point of view, since it is based on a large unique dataset that covers a period of 13 years. However, as in the study of econometrics, it suffers from some limitations, since it has not been possible to observe all the interrelated factors, such as managerial practices or environmental features, that can influence firm profitability. Future studies should also consider other output variables, such as: market share, growth, innovativeness, cost leadership and delivery cycle time (Mikalef, Krogstie, et al., 2019 Big data analytics and firm performance: Findings from a mixed-method approach).

These findings should motivate future research to assess additional firm- and industry-level effects and to increase our understanding of their roles in remove firm profitability. The effects that could be considered in future studies include industry effects, such as R&D intensity, and firm effects, such as the managerial capabilities of the CIO/CEO.

7. Conclusions

The market for BDA solutions is rapidly expanding and is expected to “reach \$260 billion in 2022 with a compound annual growth rate (CAGR) of 11.9 per cent over the 2017-2022 period”, according to the International Data Corporation. In our research, we have investigated the impact of BDA solutions on firm profitability. We found that industry- and firm-specific effects matter. Larger firms as well as firms facing more competitive environments can profit more from investments in BDA solutions. This calls for greater efforts to understand the role played by BDA solutions, which, in this study, differs from the role described in prior studies; competitiveness does

not erode the appropriability of the benefits resulting from BDA investments, even in the long term. Overall, these findings contribute to enriching the emerging body of literature on the business value of BDA.

References

- Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: A systematic review and agenda for future research. *Electronic Markets*, 26(2), 173–194. <https://doi.org/10.1007/s12525-016-0219-0>
- Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, 182, 113–131. <https://doi.org/10.1016/j.ijpe.2016.08.018>
- Autio, E., Sapienza, H. J., & Almeida, J. G. (2000). Effects of Age at Entry, Knowledge Intensity, and Imitability on International Growth. *Academy of Management Journal*, 43(5), 909–924. <https://doi.org/10.2307/1556419>
- Aydiner, A. S., Tatoglu, E., Bayraktar, E., Zaim, S., & Delen, D. (2019). Business analytics and firm performance: The mediating role of business process performance. *Journal of Business Research*, 96, 228–237. <https://doi.org/10.1016/j.jbusres.2018.11.028>
- Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., & Zhao, J. L. (2016). Transformational Issues of Big Data and Analytics in Networked Business. *MIS Quarterly*, 40(4), 807–818. <https://doi.org/10.2307/4311111>
- Bain, J. S. (1951). Relation of profit rate to industry concentration: American manufacturing, 1936–1940. *The Quarterly Journal of Economics*, 65(3), 293–324.
- Barney, J. B. (1996). The Resource-Based Theory of the Firm. *Organization Science*, 7(5), 469.
- Basole, R. C., Seuss, C. D., & Rouse, W. B. (2013). IT innovation adoption by enterprises: Knowledge discovery through text analytics. *Decision Support Systems*, 54(2), 1044–1054.
- Belman, D., & Heywood, J. S. (1990). THE CONCENTRATION-EARNINGS HYPOTHESIS: RECONCILING INDIVIDUAL AND INDUSTRY DATA IN US STUDIES. *Oxford Bulletin of Economics and Statistics*, 52(3), 293–302.
- Bharadwaj, A. S., Bharadwaj, S. G., & Konsynski, B. R. (1999). Information technology effects on firm performance as measured by Tobin's q. *Management Science*, 45(7), 1008–1024.
- Božič, K., & Dimovski, V. (2019). Business intelligence and analytics use, innovation ambidexterity, and firm performance: A dynamic capabilities perspective. *The Journal of Strategic Information Systems*, 28(4), 101578. <https://doi.org/10.1016/j.jsis.2019.101578>
- Brynjolfsson, E. (1996). The contribution of information technology to consumer welfare. *Information Systems Research*, 7(3), 281–300.
- Bughin, J. (2016). Big data, Big bang? *Journal of Big Data*, 3(1), 2.
- Burns, T., & Stalker, G. M. (1994). *The Management of Innovation* (Revised edition). Oxford University Press.
- 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), 4–39. <https://doi.org/10.1080/07421222.2015.1138364>
- Chiasson, M. W., & Davidson, E. (2005). Taking industry seriously in information systems research. *Mis Quarterly*, 591–605.
- Côrte-Real, N., Ruivo, P., & Oliveira, T. (2020). Leveraging internet of things and big data analytics initiatives in European and American firms: Is data quality a way to extract business value? *Information & Management*, 57(1), 103141. <https://doi.org/10.1016/j.im.2019.01.003>

- Covin, J. G., Slevin, D. P., & Schultz, R. L. (1994). Implementing Strategic Missions: Effective Strategic, Structural and Tactical Choices. *Journal of Management Studies*, 31(4), 481–506. <https://doi.org/10.1111/j.1467-6486.1994.tb00627.x>
- Datta, D. K., Guthrie, J. P., & Wright, P. M. (2005). Human resource management and labor productivity: Does industry matter? *Academy of Management Journal*, 48(1), 135–145.
- Davenport, Thomas H. (2006). Competing on Analytics. *Harvard Business Review*, 84(1), 98–107.
- Dehning, B., Richardson, V. J., & Zmud, R. W. (2007). The financial performance effects of IT-based supply chain management systems in manufacturing firms. *Journal of Operations Management*, 25(4), 806–824.
- Dong, J. Q., & Yang, C.-H. (2020). Business value of big data analytics: A systems-theoretic approach and empirical test. *Information & Management*, 57(1), 103124. <https://doi.org/10.1016/j.im.2018.11.001>
- Donsimoni, M.-P., Geroski, P., & Jacquemin, A. (1984). Concentration indices and market power: Two views. *The Journal of Industrial Economics*, 419–434.
- Dubey, R., Gunasekaran, A., Childe, S. J., Blome, C., & Papadopoulos, T. (2019). Big Data and Predictive Analytics and Manufacturing Performance: Integrating Institutional Theory, Resource-Based View and Big Data Culture. *British Journal of Management*, 30(2), 341–361. <https://doi.org/10.1111/1467-8551.12355>
- Ebner, K., Bühnen, T., & Urbach, N. (2014). Think big with big data: Identifying suitable big data strategies in corporate environments. *2014 47th Hawaii International Conference on System Sciences*, 3748–3757.
- Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: What are they? *Strategic Management Journal*, 21(10–11), 1105–1121.
- Feng, K., Chen, E. T., & Liou, W. (2005). Implementation of knowledge management systems and firm performance: An empirical investigation. *Journal of Computer Information Systems*, 45(2), 92–104.
- Forés, B., & Camisón, C. (2016). Does incremental and radical innovation performance depend on different types of knowledge accumulation capabilities and organizational size? *Journal of Business Research*, 69(2), 831–848. <https://doi.org/10.1016/j.jbusres.2015.07.006>
- 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. <https://doi.org/10.1016/j.ijpe.2014.12.031>
- George, G., Haas, M. R., & Pentland, A. (2014). Big Data and Management. *Academy of Management Journal*, 57(2), 321–326. <https://doi.org/10.5465/amj.2014.4002>
- Geringer, J. M., Tallman, S., & Olsen, D. M. (2000). Product and international diversification among Japanese multinational firms. *Strategic Management Journal*, 21(1), 51–80.
- Gregor, S., Martin, M., Fernandez, W., Stern, S., & Vitale, M. (2006). The transformational dimension in the realization of business value from information technology. *The Journal of Strategic Information Systems*, 15(3), 249–270.
- Grover, V., Chiang, R. H., Liang, T.-P., & Zhang, D. (2018). Creating Strategic Business Value from Big Data Analytics: A Research Framework. *Journal of Management Information Systems*, 35(2), 388–423. <https://doi.org/10.1080/07421222.2018.1451951>
- Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*, 70, 308–317. <https://doi.org/10.1016/j.jbusres.2016.08.004>
- Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems*, 26(3), 191–209. <https://doi.org/10.1016/j.jsis.2017.07.003>

- Gupta, M., & George, J. F. (2016a). Toward the development of a big data analytics capability. *Information & Management*. <https://doi.org/10.1016/j.im.2016.07.004>
- Gupta, M., & George, J. F. (2016b). Toward the development of a big data analytics capability. *Information & Management*, 53(8), 1049–1064.
- H. Davenport, T. (2014). How strategists use “big data” to support internal business decisions, discovery and production. *Strategy & Leadership*, 42(4), 45–50. <https://doi.org/10.1108/SL-05-2014-0034>
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate Data Analysis* (5th ed.). Prentice Hall.
- Hindle, G., Kunc, M., Mortensen, M., Oztekin, A., & Vidgen, R. (2020). Business analytics: Defining the field and identifying a research agenda. *European Journal of Operational Research*, 281(3), 483–490. <https://doi.org/10.1016/j.ejor.2019.10.001>
- Hitt, L. M., & Brynjolfsson, E. (1996). Productivity, business profitability, and consumer surplus: Three different measures of information technology value. *MIS Quarterly*, 121–142.
- Huang, C.-K., Wang, T., & Huang, T.-Y. (2018).** Initial Evidence on the Impact of Big Data Implementation on Firm Performance. *Information Systems Frontiers*, 1–13.
- IDC. (2018, August 15). *Revenues for Big Data and Business Analytics Solutions Forecast to Reach \$260 Billion in 2022, Led by the Banking and Manufacturing Industries, According to IDC*. IDC: The Premier Global Market Intelligence Company. <https://www.idc.com/getdoc.jsp?containerId=prUS44215218>
- Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2016). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*, 1–16.
- Kamioka, T., & Tapanainen, T. (2014). Organizational Use of Big Data and Competitive Advantage-Exploration of Antecedents. *PACIS, 2014*, 18th.
- Kaufman, B. E. (2015). The RBV theory foundation of strategic HRM: Critical flaws, problems for research and practice, and an alternative economics paradigm. *Human Resource Management Journal*, 25(4), 516–540. <https://doi.org/10.1111/1748-8583.12085>
- Kim, W. C., Hwang, P., & Burgers, W. P. (1993). Multinationals’ diversification and the risk-return trade-off. *Strategic Management Journal*, 14(4), 275–286.
- Kirca, A. H., Hult, G. T. M., Roth, K., Cavusgil, S. T., Perry, M. Z., Akdeniz, M. B., Deligonul, S. Z., Mena, J. A., Pollitte, W. A., Hoppner, J. J., Miller, J. C., & White, R. C. (2011). Firm-Specific Assets, Multinationality, and Financial Performance: A Meta-analytic Review and Theoretical Integration. *Academy of Management Journal*, 54(1), 47–72. <https://doi.org/10.5465/AMJ.2011.59215090>
- Kitchens, B., Dobolyi, D., Li, J., & Abbasi, A. (2018). Advanced Customer Analytics: Strategic Value Through Integration of Relationship-Oriented Big Data. *Journal of Management Information Systems*, 35(2), 540–574.
- Krishnamoorthi, S., & Mathew, S. K. (2018). Business analytics and business value: A comparative case study. *Information & Management*, 55(5), 643–666. <https://doi.org/10.1016/j.im.2018.01.005>
- Levy, D. (1985). Specifying the dynamics of industry concentration. *The Journal of Industrial Economics*, 55–68.
- Li, M., & Richard Ye, L. (1999). Information technology and firm performance: Linking with environmental, strategic and managerial contexts. *Information & Management*, 35(1), 43–51. [https://doi.org/10.1016/S0378-7206\(98\)00075-5](https://doi.org/10.1016/S0378-7206(98)00075-5)
- Lycett, M. (2013). ‘Datafication’: Making sense of (big) data in a complex world. *European Journal of Information Systems*, 22(4), 381–386. <https://doi.org/10.1057/ejis.2013.10>
- Makadok, R. (2001). Toward a synthesis of the resource-based and dynamic-capability views of rent creation. *Strategic Management Journal*, 22(5), 387–401. <https://doi.org/10.1002/smj.158>

- Malladi, S., & Krishnan, M. (2013). *Determinants of usage variations of business intelligence & analytics in organizations—an empirical analysis*.
- Mangin, P., Hovelaque, V., & Bironneau, L. (2015). Enterprise Resource Planning contribution to firm performance: A literature review over the last 15 years. *11th Congrès International de Genie Industrial-CIGI2015, Québec, Canada*, 26–28.
- Matthias, O., Fouweather, I., Gregory, I., & Vernon, A. (2017). Making sense of Big Data – can it transform operations management? *International Journal of Operations & Production Management*, 37(1), 37–55. <https://doi.org/10.1108/IJOPM-02-2015-0084>
- McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*, 90(10), 60–68. bth.
- Melville, N., Gurbaxani, V., & Kraemer, K. (2007). The productivity impact of information technology across competitive regimes: The role of industry concentration and dynamism. *Decision Support Systems*, 43(1), 229–242.
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019a). Big Data Analytics Capabilities and Innovation: The Mediating Role of Dynamic Capabilities and Moderating Effect of the Environment. *British Journal of Management*, 30(2), 272–298. <https://doi.org/10.1111/1467-8551.12343>
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019b). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, 98, 261–276. <https://doi.org/10.1016/j.jbusres.2019.01.044>
- Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2019). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information & Management*, 103169. <https://doi.org/10.1016/j.im.2019.05.004>
- Mikalef, P., Pappas, I. O., Krogstie, J., & Pavlou, P. A. (2020). Big data and business analytics: A research agenda for realizing business value. *Information & Management*, 57(1), 103237. <https://doi.org/10.1016/j.im.2019.103237>
- Mithas, S., Ramasubbu, N., & Sambamurthy, V. (2011). How information management capability influences firm performance. *MIS Quarterly*, 237–256.
- Müller, O., Fay, M., & vom Brocke, J. (2018). The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics. *Journal of Management Information Systems*, 35(2), 488–509.
- Oh, L.-B., Teo, H.-H., & Sambamurthy, V. (2012). The effects of retail channel integration through the use of information technologies on firm performance. *Journal of Operations Management*, 30, 368–381.
- Oliveira, T., & Martins, M. F. (2011). Literature review of information technology adoption models at firm level. *Electronic Journal of Information Systems Evaluation*, 14(1), 110.
- Peteraf, M. A., & Barney, J. B. (2003). Unraveling the resource-based tangle. *Managerial and Decision Economics*, 24(4), 309–323. <https://doi.org/10.1002/mde.1126>
- Porter, M. E., & Advantage, C. (1985). Creating and sustaining superior performance. *Competitive Advantage*, 167.
- Porter, M. E., & Sakakibara, M. (2004). Competition in Japan. *Journal of Economic Perspectives*, 18(1), 27–50.
- Qian, G., & Lee, L. (2003). Profitability of Small- and Medium-Sized Enterprises in High-Tech Industries: The Case of the Biotechnology Industry. *Strategic Management Journal*, 24(9), 881–887.
- Raguseo, E., & Vitari, C. (2018). Investments in big data analytics and firm performance: An empirical investigation of direct and mediating effects. *International Journal of Production Research*, 1–16.

- Scherer, F.M., & Ross, D. (1990). *Industrial Market Structure and Economic Performance*. Houghton Mifflin Company.
- Scherer, Frederic M., & Ross, D. (1990). Industrial market structure and market performance. *Houghton Mifflin Company, Boston, USA*.
- Schryen, G. (2013). Revisiting IS business value research: What we already know, what we still need to know, and how we can get there. *European Journal of Information Systems*, 22(2), 139–169.
- Seddon, P. B., Constantinidis, D., Tamm, T., & Dod, H. (2017). How does business analytics contribute to business value? *Information Systems Journal*, 27(3), 237–269. <https://doi.org/10.1111/isj.12101>
- Sena, V., Bhaumik, S., Sengupta, A., & Demirbag, M. (2019). Big Data and Performance: What Can Management Research Tell us? *British Journal of Management*, 30(2), 219–228. <https://doi.org/10.1111/1467-8551.12362>
- Shah, R., & Shin, H. (2007). Relationships among information technology, inventory, and profitability: An investigation of level invariance using sector level data. *Journal of Operations Management*, 25(4), 768–784.
- Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information & Management*, 56(6), 103135. <https://doi.org/10.1016/j.im.2018.12.003>
- Sharma, R., Mithas, S., & Kankanhalli, A. (2014). *Transforming decision-making processes: A research agenda for understanding the impact of business analytics on organisations*. Taylor & Francis.
- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263–286. <https://doi.org/10.1016/j.jbusres.2016.08.001>
- Surbakti, F. P. S., Wang, W., Indulska, M., & Sadiq, S. (2020). Factors influencing effective use of big data: A research framework. *Information & Management*, 57(1), 103146. <https://doi.org/10.1016/j.im.2019.02.001>
- Tambe, P. (2014). Big Data Investment, Skills, and Firm Value. *Management Science*, 60(6), 1452–1469. <https://doi.org/10.1287/mnsc.2014.1899>
- Tambe, P., & Hitt, L. M. (2012). The productivity of information technology investments: New evidence from IT labor data. *Information Systems Research*, 23(3-part-1), 599–617.
- Thompson, C. A., Kopelman, R. E., & Schriesheim, C. A. (1992). Putting all one's eggs in the same basket: A comparison of commitment and satisfaction among self- and organizationally employed men. *Journal of Applied Psychology*, 77(5), 738–743. <https://doi.org/10.1037/0021-9010.77.5.738>
- Tirole, J. (1988). *The theory of industrial organization*. MIT press.
- Vitari, C., & Raguseo, E. (2016). Digital data, dynamic capability and financial performance: An empirical investigation in the era of Big Data. *Systèmes d'information & Management*, 21(3), 63–92.
- Vitari, C., & Raguseo, E. (2019). Big data analytics business value and firm performance: Linking with environmental context. *International Journal of Production Research*, 1–21. <https://doi.org/10.1080/00207543.2019.1660822>
- Vorhies, D. W., & Morgan, N. A. (2005). Benchmarking marketing capabilities for sustainable competitive advantage. *Journal of Marketing*, 69(1), 80–94.
- Wagner, S. M., Grosse-Ruyken, P. T., & Erhun, F. (2012). The link between supply chain fit and financial performance of the firm. *Journal of Operations Management*, 30(4), 340–353. <https://doi.org/10.1016/j.jom.2012.01.001>
- Wamba, S. F., Dubey, R., Gunasekaran, A., & Akter, S. (2019). The performance effects of big data analytics and supply chain ambidexterity: The moderating effect of environmental dynamism.

International Journal of Production Economics, 107498.
<https://doi.org/10.1016/j.ijpe.2019.09.019>

- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70, 356–365.
- 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.
<https://doi.org/10.1016/j.ijpe.2016.03.014>
- Wang, S., Yeoh, W., Richards, G., Wong, S. F., & Chang, Y. (2019). Harnessing business analytics value through organizational absorptive capacity. *Information & Management*, 56(7), 103152.
<https://doi.org/10.1016/j.im.2019.02.007>
- Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, 126, 3–13. <https://doi.org/10.1016/j.techfore.2015.12.019>
- Yu, W., Jacobs, M. A., Chavez, R., & Feng, M. (2019). Data-Driven Supply Chain Orientation and Financial Performance: The Moderating Effect of Innovation-Focused Complementary Assets. *British Journal of Management*, 30(2), 299–314. <https://doi.org/10.1111/1467-8551.12328>
- Zhang, D., Pan, S. L., Yu, J., & Liu, W. (2019). Orchestrating big data analytics capability for sustainability: A study of air pollution management in China. *Information & Management*, 103231. <https://doi.org/10.1016/j.im.2019.103231>

Figure

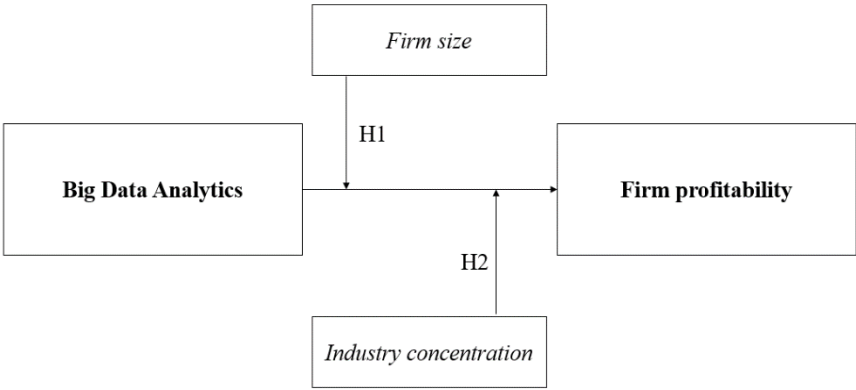


Figure 1. Research framework

Tables

Table 1. Sample characteristics

Sample characteristics	Percentage (%)
Role of the respondent	
<i>General Director</i>	26.5%
<i>Chief Executive Officer</i>	4.0%
<i>Chief Information Officer</i>	63.5%
<i>Other person qualified to make BDA investments</i>	6.0%
Industrial sectors	
<i>Manufacturing</i>	28.5%
<i>Services</i>	42.5%
<i>Retail</i>	20.5%
<i>Construction</i>	8.5%
Total	100%
Revenue interval	
<i>Medium-sized firm revenues between €10 and €50 million</i>	71.59%
<i>Large companies with more than €50 million</i>	28.41%

Table 2. Operationalization of the variables

Variable	Operationalization	Reference	Data source
<i>Firm profitability</i>	Return on Sales Return on Assets Return on Equity	(Geringer et al., 2000)	Bureau Van Dijk DIANE database
<i>Big data analytics</i>	Usage in companies with at least one of the identified BDA solutions	(H. Davenport, 2014; Müller et al., 2018)	Survey
<i>Industry concentration</i>	$HHI = s_1^2 + s_2^2 + s_3^2 + \dots + s_n^2$ (where s is the integer market share of each firm in a sector)	(Bharadwaj et al., 1999; Porter & Sakakibara, 2004)	Bureau Van Dijk DIANE database
<i>Firm size</i>	Sales revenues in millions of euros (logarithmic form)	(Qian & Lee, 2003)	Bureau Van Dijk DIANE database
<i>Age</i>	The logarithmic form of the actual existence of the firm since the year it started operation	(Autio et al., 2000)	Bureau Van Dijk DIANE database
<i>Years</i>	Dummy variables, one for each year	(Melville et al., 2007)	Survey
<i>Firm risk</i>	Standard deviation of accounting-based returns using a 3-year average before the year to which the observation refers	(Kim et al., 1993)	Bureau Van Dijk DIANE database
<i>Past performance</i>	The average of all the accounting-based returns for the previous three years before the year to which the observation refers	(Autio et al., 2000)	Bureau Van Dijk DIANE database
<i>Dynamism</i>	Standard error of the regression slope coefficient of the annual industry sales divided by the industry mean for the 5 year period	(Stoel M.D., & Muhanna, W.A., 2009)	Bureau Van Dijk DIANE database

Table 3. Descriptive statistics and Spearman's correlation matrix

N	Variable	Mean	Standard deviation	Min.	Max.	1	2	3	4	5	6	7	8
1	ROS	3.395	6.861	-23.542	54.323	1.000							
2	ROA	6.994	13.963	0.023	45.628	0.310*	1.000						
3	ROE	9.345	14.948	-39.957	34.891	0.799*	0.131*	1.000					
4	BDA	0.000	1.000	0.069*	0.004	0.041*	1.000				
5	HHI	201.534	361.526	2.960	2359.08	0.085*	-0.125*	0.003	0.018	1.000			
6	Firm size	9.612	1.147	4.390	13.907	0.019	0.007	0.067*	0.059*	0.061*	1.000		
7	Age	3.043	0.758	0.693	4.574	0.020	-0.146*	-0.026	-0.027	-0.210*	0.218*	1.000	
8	Year	2004	2016	-0.038	-0.055*	-0.069*	0.192*	-0.142*	0.118*	0.179*	1.000

Note: * $p < 5\%$.

Table 4. Overall BDA diffusion rate according to the HHI intervals and firm size (revenues) intervals (percentage values of the subgroups)

BDA diffusion rate	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Overall	8.2 %	8.4 %	13.7 %	13.7 %	15.3 %	16.3 %	17.9 %	23.7 %	25.3 %	30.0 %	33.2 %	35.8 %	35.8 %
HHI intervals													
HHI<1,500 as competitive marketplace	8.6 %	8.6 %	13.9 %	13.9 %	15.3 %	15.6 %	17.2 %	23.5 %	25.1 %	29.7 %	33.0 %	35.7 %	35.7 %
HHI>=1,500 as moderately/highly concentrated marketplace	0.0 %	0.0 %	0.0%	0.0%	14.3 %	14.3 %	28.5 %	28.5 %	28.5 %	37.5 %	37.5 %	37.5 %	37.5 %
F statistic	0.37	0.37	0.64	0.64	0.30	3.42 †	2.88 †	1.56	1.32	0.24	0.11	0.04	0.04
Revenue interval													
Medium-sized firm revenues between €10 and €50 million	6.0 %	7.2 %	8.9%	11.4 %	10.8 %	15.0 %	18.9 %	24.3 %	25.1 %	27.2 %	30.8 %	34.3 %	34.8 %
Large companies more than €50 million	3.4 %	5.0 %	17.2 %	15.5 %	17.2 %	17.8 %	20.4 %	27.4 %	28.8 %	40.0 %	41.2 %	41.2 %	41.2 %
F statistic	1.64	1.76	8.18* *	3.10 †	2.30	2.91 †	3.59 †	1.68	2.07	3.31 †	1.93	0.41	0.03

Note: ***p-value < 0.1%; ** p < 1%; * p < 5%; † < 10%.

Table 5. Results of the fixed-effects regressions

Dependent variable	ROS	ROS	ROS	ROA	ROA	ROA	ROE	ROE	ROE
Model	M1	M2	M3	M4	M5	M6	M7	M8	M9
<i>Direct effects</i>									
Big data analytics (BDA) (t-1)	0.518	1.088	1.128	0.264	0.271	0.278	0.351	0.263	0.512
	(0.573)	(0.603)	(0.653)	(0.274)	(0.270)	(0.274)	(0.827)	(0.936)	(0.947)
Industry concentration (IC) (t-1)	-2.164*	-3.510***	-3.020**	1.156***	0.927***	1.052***	-2.985***	-3.177*	-3.338*
	(1.184)	(1.241)	(1.261)	(0.344)	(0.339)	(0.348)	(0.936)	(1.813)	(1.816)
Firm size (log) (FS) (t-1)	1.312	1.448	1.300	1.405***	1.709***	1.389***	4.195***	4.683	4.182
	(0.954)	(0.923)	(0.950)	(0.444)	(0.430)	(0.443)	(1.263)	(3.797)	(3.468)
<i>Moderating effects</i>									
BDA x FS (t-1)	1.495***		1.258**	0.436*		0.456*	1.459**		1.590*
	(0.562)		(0.574)	(0.236)		(0.236)	(0.656)		(0.911)
BDA x IC (t-1)		-2.340**	-2.180*		-0.569**	-0.487*		-1.210**	-1.389**
		(1.067)	(1.136)		(0.256)	(0.260)		(0.534)	(0.572)
<i>Control variables</i>									
Firm age (log)	2.603	5.235	2.842	1.695	1.604	1.606	-3.243	-3.688	-3.358
	(5.759)	(5.643)	(5.738)	(1.163)	(1.157)	(1.163)	(3.816)	(4.270)	(4.128)
Past performance	0.0737***	0.0766***	0.0739***	0.0112*	0.0115*	0.0127**	2.901***	2.877***	2.903***
	(0.009)	(0.009)	(0.009)	(0.006)	(0.006)	(0.006)	(0.133)	(0.360)	(0.333)
Firm risk	-0.736***	-0.574***	-0.723***	0.267***	0.282***	0.261***	-0.0417**	-0.0325	-0.0408
	(0.102)	(0.077)	(0.102)	(0.037)	(0.037)	(0.037)	(0.017)	(0.024)	(0.026)
Dynamism	8.782	9.144	10.78	8.334	8.277	9.996*	-4.307	1.223	-0.576
	(9.672)	(9.098)	(9.691)	(5.580)	(5.501)	(5.644)	(15.930)	(15.470)	(16.460)
Years	Included	Included	Included	Included	Included	Included	Included	Included	Included
Constant	2.926***	2.236**	2.829***	5.264***	5.225***	5.191***	-2.062	-2.267	-2.316
	(1.068)	(0.993)	(1.065)	(0.419)	(0.416)	(0.420)	(1.442)	(2.153)	(2.157)
R-squared	37.30%	37.60%	38.00%	8.80%	9.50%	9.00%	39.10%	39.10%	39.30%

Note: The variables in the models are standardized; standard errors appear in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6. Results of the fixed-effects regressions used to test for reverse causality or simultaneity

Dependent variable	Big data analytics	Big data analytics	Big data analytics
Model	M10	M11	M12
<i>Independent variables</i>			
ROS (t-1)	-0.001 (0.001)		
ROA (t-1)		-0.001 (0.001)	
ROE(t-1)			0.001 (0.001)
<i>Control variables</i>			
HHI	-0.012 (0.016)	-0.013 (0.015)	-0.016 (0.015)
Firm size (log)	0.055*** (0.019)	0.049*** (0.019)	0.046** (0.019)
Firm age (log)	0.127** (0.050)	0.134*** (0.050)	0.141*** (0.050)
Dynamism	0.577** (0.249)	0.493** (0.248)	0.609** (0.249)
Years	Included	Included	Included
Constant	0.148*** (0.019)	0.151*** (0.019)	0.146*** (0.018)
R-squared	15.10%	14.70%	15.10%

Note: The variables in the models are standardized; standard errors appear in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The "firm risk" and "past performance" control variables have been omitted from these models since they increase the VIFs of the independent variables.

Table 7. Post hoc analysis: regression models with BDA and BDA squared as independent variables

Dependent variable	ROS	ROA	ROE
Model	M13	M14	M15
<i>Independent variables</i>			
Big data analytics (BDA) (t-1)	0.520	0.558	10.73
	(1.346)	(1.302)	(6.750)
BDA squared (t-1)	0.007	-0.144	-1.410
	(0.318)	(0.306)	(1.460)
<i>Control variables</i>			
Industry concentration (IC) (t-1)	0.646	0.197	-9.686***
	(0.632)	(0.622)	(3.162)
Firm size (log) (FS) (t-1)	0.752	1.657**	14.800***
	(0.677)	(0.670)	(3.457)
Firm age (log)	0.495	-1.117*	-2.018
	(0.563)	(0.570)	(3.359)
Past performance	0.115***	-0.0391**	1.049**
	(0.020)	(0.019)	(0.454)
Firm risk	-0.268	1.384***	0.716***
	(0.206)	(0.038)	(0.028)
Dynamism	76.730**	-24.640	-33.500
	(29.620)	(29.460)	(98.130)
Constant	0.086	4.777***	-13.130**
	(1.300)	(1.120)	(6.225)
R-squared	24.50%	31.60%	32.30%

Note: Standard errors appear in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appendix

Table A1 Summary of the most relevant articles on big data analytics and performance published in the British Journal of Management, European Journal of Information Systems, Information & Management, Information Systems Journal, International Journal of Production Economics, Information Systems Research, Journal of the Association for Information Systems, Journal of Information Technology, Journal of Management of Information Systems, Journal of Strategic Information Systems and Management of Information Systems Quarterly until the end of 2019

Big data (independent)	Mediators	Moderators	Performance (dependent)	Source
Analytic leadership, enterprise-wide analytic orientation, well-chosen targets, extent to which evidence-based decision making is embedded in the DNA of the organization, functional fit of business analytic tools, readily available high quality data, analytical people, overcoming organizational inertia	Use analytic resources → Insights → Decisions → Intended value-creating actions that use the organisation's existing resources	Organizational Resources	Organisational benefits	(Seddon et al., 2017)
IT-supported data infrastructure	IT-enabled agility → Advanced customer analytics capability portfolio → Improved customer lifetime value & customer equity		Strategic value	(Kitchens et al., 2018)
BDA infrastructure (BD assets, analytics portfolio, human talent)	BDA capabilities → Value creation mechanisms → Value targets		Impact (functional value, symbolic value)	(Grover et al., 2018)
Big data analytics use		Environmental dynamism	Asset productivity, business growth	(Chen et al., 2015)
Business intelligence & analytics	Absorptive capacity → Innovation ambidexterity		Firm performance (sales growth, ROA)	(Božič & Dimovski, 2019)
Portability and interconnectivity	Work practice-working with big data in practice, Organizational-Developing organizational models, Supra-organizational-dealing with stakeholder interests		Social and economic value	(Günther et al., 2017)
Motivational, operational, supporting mechanisms			Value realization	(Surbakti et al., 2020)
Structuring big data and the business analytics resource portfolio	Orchestrating big data and business analytics resources to build capabilities → Business capability to exploit market opportunities → Value creation streams	Environmental uncertainty	Competitive advantage	(Mikalef et al., 2020)
Social media analytics, social media diversity		Firm size	Market performance (total sales)	(Dong & Yang, 2020)
Data quality (completeness, accuracy, format, currency)		Process sophistication	Competitive advantage (strategic performance, financial performance)	(Côte-Real et al., 2020)
Business analytics competency	Organizational absorptive capacity → Business analytics assimilation		Competitive advantage	(S. Wang et al., 2019)
Big data analytics capability (tangible, human skills, intangible)	Dynamic capabilities, market capabilities, technological capabilities		Competitive performance (profitability, market share, growth, innovativeness, cost leadership, delivery cycle time)	(Mikalef, Krogstie, et al., 2019)

Big data (independent)	Mediators	Moderators	Performance (dependent)	Source
Big data analytics capability	size-class, organizational structure, industry, top manager support, technical skills, managerial skills	Dynamism, heterogeneity, hostility	Business value / Performance	(Mikalef, Boura, et al., 2019b)
Big data analytics capability (tangible, human skills, intangible)	Dynamic capabilities	Environment (dynamism, heterogeneity, hostility)	Incremental innovation, radical innovation	(Mikalef, Boura, et al., 2019a)
Analytics resources (analytics technology assets, business analytics capability)		Analytics value enhancer, organizational level variables	Business performance (organizational benefits from analytics use, ROI)	(Krishnamoorthi & Mathew, 2018)
Big data analytics capability (tangible, human skills, intangible)			Market performance, operational performance	(Gupta & George, 2016b)
Big data analytics	Supply chain ambidexterity	Environmental dynamism	Organizational performance	(Wamba et al., 2019)
Big data analytics capability (Management, Technology, Talent)		Analytics capability–business strategy alignment	Firm performance	(Akteer et al., 2016)
Adoption of business analytics (data acquisition and processing, prescriptive analytics, predictive analytics, descriptive analytics)	Business process performance		Firm performance	(Aydiner et al., 2019)
Institutional pressure → tangible resources, human skills → Big data predictive analytics		Big data driven culture	Cost performance, operational performance	(Dubey et al., 2019)
Data drive supply chain orientation		Innovation focused complementary assets	Firm performance	(Yu et al., 2019)
Big data analytics		Industry concentration, firm size	Firm profitability (ROE, ROA, ROS)	The present study