The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1755-425X.htm

Big data analytics capability as a mediator in the impact of open innovation on firm performance

José Arias-Pérez and Alejandro Coronado-Medina
Department of Administrative Sciences, Universidad de Antioquia, Medellin, Colombia, and
Geovanny Perdomo-Charry
University of Technology Sydney, Haymarket, Australia and CEIPA Business School, Sabaneta, Colombia

Received 24 September 2020
Revised 1 February 2021
8 March 2021
19 April 2021
20 May 2021
Accepted 30 May 2021

Abstract

Purpose – Big data analytics capability (BDAC) is the ability of a firm to capture and analyze big data toward the generation of insights. The literature has mainly focused on analyzing the direct effects of BDAC on different aspects related to firm performance such as finances and innovation. However, the lack of works analyzing the intermediation role BDAC could play is noticeable, particularly in organizational situations that pose great challenges in terms of data processing. Thus, the aim of this paper is to analyze BDAC mediation in the relationship between open innovation (OI), particularly customer involvement, and firm performance (financial and non-financial).

Design/methodology/approach – Structural equation modeling was used to test the proposed model with survey data from a sample of 112 firms.

Findings – The results show that BDAC has a partial mediating effect on the relationship between OI and financial performance, and between OI and non-financial performance. Nevertheless, this mediation is greater in the first relationship.

Originality/value – The main contribution of the study is to offer a broader research perspective regarding the role of BDAC in the relationship between OI and firm performance. This study ultimately questions that research tradition in which this role has been reduced to that of a simple application of data analytics techniques. Instead, the results show BDAC is primarily an organizational skill that should be articulated with key processes, such as customer involvement, to maximize the financial and non-financial use of the large flow of data coming from the main OI activity of low and medium-technology companies.

Keywords Big data analytics, Open innovation, Data-driven innovation, Collaborative innovation platforms, Digital transformation, Financial and non-financial performance

Paper type Research paper

Introduction

Digital technology is increasingly important in achieving business goals, and its pervasive effects have resulted in the radical restructuring of entire industries (Peter *et al.*, 2020). Digitalization refers to the way in which new digital technologies such as analytics, cloud, mobile and social media can be used to modify existing business processes (Arias-Pérez *et al.*, 2021b). Among all these technologies, the Deloitte firm recently identified big data analytics to be in the top three of its investment priorities, all with a view to improving revenue and customer value: it may increase by 15 and 23%, respectively, in an early adoption stage, and reach 45 and 41%, respectively, when a maturity stage is reached (Gurumurthy *et al.*, 2020). It is foreseen, however, that the greatest impact of big data analytics in firms will be a greater exploitation of the knowledge captured from external sources or derived from open innovation (OI) processes, i.e. collaborative work with external actors (Dahiya *et al.*, 2021; Urbinati *et al.*, 2019).

This work was supported by Universidad de Antioquia and CEIPA Business School under project number 2019-25750.

1755-425X
DOI 10.1108/JSMA-09-2020-0262

These recent findings from consulting firms mirror the results that academic literature has repeatedly shown regarding the positive relationship between big data analytics capability (BDAC) and firm performance (Caputo *et al.*, 2016; Gupta *et al.*, 2020; Mikalef *et al.*, 2019a). BDAC is defined as the ability of a firm to capture and analyze big data toward the generation of insights by effectively orchestrating and deploying its data, technology and talent (Henao-García *et al.*, 2021). Conversely, firm performance has been defined as the achievement of financial and non-financial organizational objectives (Chadwick and Dawson, 2018; Meisinger and Moldaschl, 2020). The former relates to return on investment and sales growth, among others; the latter is concerned with customer satisfaction and employee capacity improvement, to mention a few (Ibarra-Cisneros *et al.*, 2021; Lee *et al.*, 2011).

Nevertheless, this type of studies that provide a generic analysis of the relationship between BDAC and organizational performance has been the target of certain criticism: by not placing BDAC in specific organizational settings where there is a large flow of data and information, they could be revealing positive effects without opening the black box (Ferraris *et al.*, 2019; Mikalef *et al.*, 2018). As a result, there is a growing interest in the literature to analyze BDAC in relation to specific organizational processes where data flow in large quantities, specifically with supply chain management (Bag *et al.*, 2021; Rialti *et al.*, 2019); and in an emerging fashion with decision-making (Shamim *et al.*, 2020) and OI on virtual platforms (Lozada *et al.*, 2019).

On the other hand, in the context of digitalization, OI is seen as one of the most important research topics in the coming years (Enkel *et al.*, 2020). OI is defined as the innovation process based on the purposive management of knowledge flows beyond organizational boundaries (Chesbrough and Bogers, 2014; Dogbe *et al.*, 2020). OI is subdivided into two major organizational processes: acquisition or inbound and outbound (Arias-Pérez *et al.*, 2021a; van de Vrande *et al.*, 2009).

Unfortunately, previous studies linking OI with big data analytics are marked by their focus on the application of specific techniques and solutions for taking advantage of the knowledge flow coming from external sources (Urbinati *et al.*, 2019) seeking to generate innovation ideas (Ciampi *et al.*, 2021; Wen *et al.*, 2020). In other words, these studies have mainly focused on the acquisition process and the operational dimension of data analytics, mostly paying attention to the type of external data source (Dahiya *et al.*, 2021), for example if they are customers (Zhang and Xiao, 2020) or patents (Papa *et al.*, 2021).

Therefore, on the one hand, there is a shortage of studies adopting a less operational research perspective and analyzing the relationship between big data analytics and the OI process of acquisition beyond the simple application of a specific data analytics technique. In this respect, we believe that for there to be a real use of the data and the knowledge flow created by OI, rather than learning to apply a technique, companies need to develop an organizational capacity to generate insights from big data, i.e. BDAC, which involves having adequate resources in terms of human resource and technology.

On the other hand, the OI process of acquisition implies collaborative work with external allies through specific activities such as customer involvement, external networking, crowdsourcing, outsourcing technology services, inward intellectual property (IP) licensing, among others (Burchardt and Maisch, 2019; Cappa *et al.*, 2019; van de Vrande *et al.*, 2009). Hence, more studies are needed to examine BDAC in relation to some specific acquisition activities involving collaborative work with external actors, *e.g.* customers or suppliers, from which large amounts of data emerge permanently (Lozada *et al.*, 2019).

In the context of emerging countries, where small and mid-size enterprises (SMEs) from medium- and low-technology technological sectors prevail, the core activity in the OI process of acquisition is customer involvement or co-innovation with customers (Arias-Pérez *et al.*, 2021a; van de Vrande *et al.*, 2009). We believe that BDAC plays an intermediary role in the relationship between OI with customers and firm performance (financial and non-financial).

The said relationship has been analyzed extensively in the literature (Hung and Chou, 2013; Moretti and Biancardi, 2020; Sisodiya *et al.*, 2013). That is, a great flow of data and information emerges from collaborative work with customers and using it to improve firm performance would fully rely on the mediation of BDAC. Therefore, the aim of the work is to analyze that mediating effect based on survey data collected from a sample of companies that work collaboratively in a program sponsored by an institution from the regional innovation system.

In consequence, understanding the role of big data analytics in the OI process of acquisition from a BDAC perspective is the first contribution of the article, as it breaks with that reductionist tradition that emphasizes technique and the type of external data source. Our research, thus, also contributes otherwise to the discussion on the activities associated with the OI process of acquisition and BDAC by analyzing the role of BDAC in improving firm performance from the flow of knowledge derived from collaborative work with an external actor in particular: customers.

Theoretical framework and hypotheses development

The knowledge-based view (KBV) of the firm proposes that the inventory of individual and social knowledge is the most valuable resource of the organization (Grant, 1996) and the main determinant of competitive advantage (Kogut and Zander, 1992). OI and BDAC are then considered paramount organizational skills, given their potential to generate knowledge about the environment –particularly about technology– and the market, among other aspects, which can be used to overcome competition (Enkel *et al.*, 2020; Gupta *et al.*, 2020).

Open innovation (OI) and firm performance

Acquisition or inbound involves collaborative work with external allies through activities such as customer involvement, external networking, crowdsourcing, external participation investments in new or established enterprises to gain access to their knowledge, outsourcing technology services and inward IP licensing involving the purchase or use of intellectual property, such as patents, copyrights or trademarks, of other organizations to benefit from external knowledge (Burchardt and Maisch, 2019; van de Vrande *et al.*, 2009).

In the context of emerging countries, where SMEs from medium- and low-technology technological sectors prevail, the core activity of the OI process of acquisition is customer involvement or co-innovation with customers (Arias-Pérez *et al.*, 2021a; van de Vrande *et al.*, 2009). Customer involvement is an organizational skill comprised of four dimensions (Taghizadeh *et al.*, 2016): *Dialogue* represents communication between companies and customers. *Access* refers to tools and processes that facilitate co-innovation with customers. *Risk* is related to measurements that enable customers to assess the risk involved in accepting a value proposition. *Transparency* is the degree to which a company reduces information asymmetry relating to customers (Zaborek and Mazur, 2017).

The relationship between OI, particularly its acquisition process, and firm performance has been studied extensively in the literature, and this process has been proven to energize the knowledge flow from external sources into the company, facilitating the improvement of financial and non-financial performance aspects (Hung and Chou, 2013; Liao *et al.*, 2020; Sisodiya *et al.*, 2013; Wang *et al.*, 2015). In particular, customer involvement allows knowledge capture in terms of ideas, experiences and customer needs, and the development of innovative and creative solutions (Dogbe *et al.*, 2020; Wen *et al.*, 2020), which, when implemented, are translated into improvements of different firm performance aspects, including revenue, customer satisfaction and profitability, among others (Caputo *et al.*, 2016; Moretti and Biancardi, 2020).

Mediating role of big data analytics capabilities (BDAC) in the relationship between open innovation and organizational performance

The term big data has gained influence as a paradigm for taking advantage of the growing flow of information to add value to the data for making better-supported decisions (Vidgen *et al.*, 2017). Despite the growing influence of big data, there are still divergences of concepts (Dumbill, 2012), but it could be defined as information assets with characteristics of high volume, high velocity and high variety, which demand innovative and cost-effective ways of processing to improve their understanding and making sound decisions (Chen *et al.*, 2012; Dumbill, 2012).

Volume refers to the amount of information generated from data internally or externally in organizations from sensors and internet of things (IoT) devices, transactions, social networks, among others. It is a characteristic related to technological development since its magnitude depends on the generation and storage capacity; what today could be big data might not be in the future (Demchenko *et al.*, 2013). Velocity is related to the speed of information generation, the variety of the collected data already structured; from spreadsheets to unstructured data such as text, image, audio and others (Gandomi and Haider, 2015).

To the mentioned characteristics, also known as the 3Vs, other authors add veracity and value to form the so-called 5Vs of big data (Fosso Wamba *et al.*, 2015; Gantz and Reinsel, 2011). Veracity is understood as reliability of the origin of the information and reliability of the processing method. Value is understood as the benefit generated by data processing, which can be given in the form of prediction, forecasting and making better business decisions. Additionally, dimensions could be added such as variability, associated with the variation of information flows, or complexity associated with the diversity of data sources (Gandomi and Haider, 2015).

Big data analytics capabilities refer to incorporating big data into the organizational decision-making process (Rialti *et al.*, 2019). This approach includes not only big data technical infrastructure, but human skills and the knowledge required to embrace, adopt and deploy data-driven insights extraction (Van De Wetering *et al.*, 2019). BDAC can be divided into different factors: (1) tangible, such as technology and financial resources; and (2) intangible, which relates to human skills and knowledge (Gupta and George, 2016). Human skills are relevant to derive value from information, and data culture created within companies is key to implement technologies and conduct successful processes with higher information flows and knowledge (Mikalef *et al.*, 2019b).

BDAC helps companies to identify market trends and preferences to develop products and services to meet demand (Beretta, 2019). BDAC can enable organizations to focus their strategic decision-making process (Ferraris *et al.*, 2019) related to markets more accurately, reducing uncertainty and risks at the same time (Gupta and George, 2016).

At a strategic level, BDAC became a crucial element (Hagel, 2015) since it facilitates the understanding of business dynamics for managers, identifying business trends and changes in the environment (Schläfke *et al.*, 2013), through the transformation of large volumes of data into useful information (Alnoukari and Hanano, 2017), which also has a positive impact in terms of market value, productivity and financial benefits (Brynjolfsson *et al.*, 2011). The above is reflected in the growing trend of companies toward investing in data analytics tools (Davenport *et al.*, 2010). BDAC could determine a high or low performance since it allows companies to become proactive and prospective; it can also improve their yields and customer acquisition by about 8%. (Liu, 2014).

Companies that develop better BDAC are more likely to innovate and create new products and services (Mikalef *et al.*, 2019b); its relevance for the OI debate results from its importance to identify social changes in consumption, production and its connection with the exchange of knowledge, information and skills, which are proven to be highly effective to conduct successful OI processes (Del Vecchio *et al.*, 2018). By developing BDAC, companies can take a great amount of data from external interactions between firms and external actors such as customers to identify business opportunities, expand market share, reduce costs or create new products and services.

For example, customer involvement yields data on customer experiences with the company's existing products, whether structured such as databases with customer contact information and recent orders or returns, or unstructured such as reports of customer opinions on the company's social networks or on virtual innovation platforms. Nonetheless, BDAC is essential for taking advantage of this large flow of data; it means the innovation team will have the knowledge and criteria to draw on the application of data analytics to create insights that will ultimately be translated into solutions improving firm performance. In situations such as this, sentiment analysis classifies the opinions of customers about a product, and even allows knowing in real time whether those opinions have a positive, negative or neutral connotation, facilitating the generation and implementation of innovative solutions (Zhou *et al.*, 2020).

In short, large amounts of data and information emerge from collaborative work with customers, e.g. from their experiences with new and existing products, from conversations with them, from changes in their preferences, among others. However, BDAC is the organizational ability that allows to permanently and routinely generate insights from these data so that they may be translated into actions that concretely enhance key aspects of firm performance such as revenue, customer satisfaction, among others. Therefore, this study proposes the following hypotheses:

H1. The relationship between the process of inbound OI, particularly customer involvement and financial performance, is mediated by BDAC.

H2. The relationship between the process of inbound OI, particularly customer involvement and non-financial performance, is mediated by BDAC.

Methodology

Sample and data collection

The proposed model was tested in a sample of manufacturing and service companies located in Colombia, from medium- and low-technology sectors, namely, food and plastic products manufacturing. The use of big data analytics in these industries has grown exponentially in recent years, the reason why they are changing in an accelerated manner (Nara *et al.*, 2021; O'Connor and Stephen, 2017; Shukla *et al.*, 2019). These firms work collaboratively in a program sponsored by an institution from the regional innovation system, which liaises firms with universities and digital technology providers.

In this scenario, firms have several tools at their disposal to devise new products and adopt new technologies, including digital ones, e.g. training in cloud services, and machine learning and its different applications for business. In principle, this strategy has encouraged local companies to adopt cloud services, *e.g.* data storage, to use big data analytics tools, such as sentiment analysis, and to develop analytical models to segment and anticipate customer churn or buying intention. On the other hand, companies have access to a virtual co-innovation platform on which they create innovation ideas with customers, universities and other external actors.

The questionnaire was administered via e-mail and physically to senior management from a total of 600 firms. Field work was conducted between September 2018 and October 2018. Finally, 112 responses were obtained; a sample size guaranteeing satisfactory statistical power, greater than 80% (Hair *et al.*, 2019).

Measurement scale

To measure the BDAC, the scale of Gupta and George (2016) was used, which is a construct composed of three second-order constructs: tangible, human and intangible skills. Tangible is composed of three formative constructs: data, technology and basic resources; intangible of two reflective constructs: data-driven culture and intensity of organizational learning. The scale from Taghizadeh *et al.*, 2016) comprising four dimensions – dialogue, access, risk and transparency – was used to measure OI with customers. For financial and non-financial performance, the scales proposed by Lee *et al.* (2011) were used. We decided to use these scales because they look beyond financial aspects, such as customer satisfaction – an essential indicator when analyzing innovation with customers as an independent variable. Besides, a Likert scale from *totally disagree* (1) to *totally agree* (5) was conducted.

Reliability and validity

The reliability and validity of the measurement model were examined with equations by the partial least squares (PLS) method, given the presence of formative and reflective constructs (Hair *et al.*, 2019). Since the model BDAC is an endogenous variable with more than two layers of constructs, it was decided to perform the two-stage hierarchical component model (HCM) analysis independently to simplify the PLS path model. This analysis allows to obtain the scores of the three latent constructs of this capability: tangible, human and intangible abilities, and use these scores as manifest variables in the PLS path model, particularly in the evaluation process of the measurement model as a whole and the structural model (Hair *et al.*, 2019).

Before obtaining the latent variable scores, the reliability and validity of the scale were examined (Table 1). In the case of the formative constructs, it was verified that the variance inflation factor (VIF) values were below 5, and the weights of the constructs and formative items were significant. In the case of items with no significant weights, it was verified that the loading was significant (Hair *et al.*, 2019).

Regarding the reflective constructs, it was verified that all the items had a loading equal to or greater than 0.7. It was found that all constructs presented a Cronbach's alpha (CA), composite reliability (CR) and Dijkstra–Henseler (pA) indexes greater than 0.7, and an average variance extracted (AVE) greater than 0.5. Table 1 shows the validation results of the BDAC. To establish discriminant validity, it was confirmed that all heterotrait-monotrait (HTMT) values were below the threshold of 0.85.

Mediating effect test

This study followed the procedure by Zhao *et al.* (2010), who proposed the confirmation of the statistical significance of indirect effects by the bootstrap-percentile test, as the only criterion to account for the existence of a mediating effect. This procedure constitutes an improvement compared to the one previously used (Baron and Kenny, 1986), which demanded to prove the significant direct effects between X and Y as prerequisite to test the mediation. This does not make sense since the mediation can be total, i.e. all the effect of X on Y takes the indirect way through the mediator (Hair *et al.*, 2017). To confirm the above, structural equations were used by the method of PLS (Hair *et al.*, 2019).

Results

Table 2 shows that the paths of the indirect effects between OI and BDAC (β = 0.46; t-value = 5.64), between BDAC and financial performance (FP) (β = 0.40; t-value = 2.52), and between BDAC and non-financial performance (NFP) (β = 0.41; t-value = 2.97) are significant and positive. This being so, the mentioned criterion defined in the procedure of Zhao *et al.* (2010) is met to test the mediations. Therefore, H1 and H2 are accepted. However, the direct effects paths between OI and FP (β = 0.24; t-value = 1.76), and between OI and NFP (β = 0.37; t-value = 2.99) are also significant and positive.

Constructs	Weight	Loading	VIF
BDAC (third-order)			
Tangibles (second-order)	0.40***		
Data (first-order)	0.16		
BDAC1 – We have access to very large, unstructured or fast-moving data for analysis	0.05	0.66***	1.84
BDAC2 – We integrate data from multiple internal sources into a data warehouse or mart for easy access	0.38**		2.91
BDAC3 – We integrate external data with internal to facilitate high-value analysis of our business environment	0.65***		2.21
Technology (first-order)	0.48***		
BDAC4 – We have explored or adopted parallel computing approaches (e.g. Hadoop) to big data processing	0.38***		3.85
BDAC5 – We have explored or adopted different data visualization tools	0.15	0.88***	3.93
BDAC6 – We have explored or adopted cloud-based services for processing data and performing analytics	0.04	0.84***	3.68
BDAC7 – We have explored or adopted cloud-based services for processing data and performing analytics	0.21*		3.01
BDAC8 – We have explored or adopted new forms of databases such as Not Only SQL (NoSQL) for storing data	0.33**		4.22
Basic resources (first-order)	0.43***		
BDAC9 – Our big data analytics projects are adequately funded	0.76*		4.17
BDAC10 – Our big data analytics projects are given enough time to achieve their objectives	0.26***		4.17
Human skills (second-order) (CA = 0.98; ρA = 0.98; CR = 0.98; AVE = 0.81)	0.43***		
BDAC11 – We provide big data analytics training to our own employees		0.74***	
BDAC12 – We hire new employees that already have the big data analytics skills		0.82***	
BDAC13 – Our big data analytics staff has the right skills to accomplish their jobs successfully		0.93***	
BDAC14 – Our big data analytics staff has suitable education to fulfill their jobs		0.93***	
BDAC15 – Our big data analytics staff holds suitable work experience to accomplish their jobs successfully		0.94***	
BDAC16 – Our big data analytics staff is well trained		0.94***	
BDAC17 – Our big data analytics managers understand and appreciate the business needs of other functional managers, suppliers and customers		0.94***	
BDAC18 – Our big data analytics managers are able to work with functional managers, suppliers and customers to determine opportunities that big data might bring to our business		0.92***	
BDAC19 – Our big data analytics managers are able to coordinate big data-related activities in ways that support other functional managers, suppliers and customers		0.94***	
BDAC20 – Our big data analytics managers are able to anticipate the future business needs of functional managers, suppliers and customers		0.88***	
BDAC21 – Our big data analytics managers have a good sense of where to apply big data		0.93***	

(*continued*)

Table 1.
Reliability and validity of BDAC

Table 1.

Constructs	Weight	Loading	VIF
Intangibles (second-order)	*0.27****		
Data-driven culture (first-order) (CA = 0.82; ρA = 0.85; CR = 0.90; AVE = 0.74)	0.43**		
BDAC22 – We base our decisions on data rather than on instinct		0.76***	
BDAC23 – We continuously assess and improve the business rules in response to insights extracted from data		0.92***	
BDAC24 – We continuously coach our employees to make decisions based on data		0.90***	
Intensity of organizational learning (first-order) (CA = 0.92; ρA = 0.93; CR = 0.95; AVE = 0.81)	0.64***		
BDAC25 – We are able to search for new and relevant knowledge		0.91***	
BDAC26 – We are able to acquire new and relevant knowledge		0.93***	
BDAC27 – We are able to apply relevant knowledge		0.91***	
BDAC28 – We have made concerted efforts for the exploitation of existing competencies and exploration of new knowledge		0.84***	
OI (second-order)			
Dialogue (first-order)	0.44*		2.81
OI1 – Use diversified communication channels to have dialogue sessions with consumers	0.15	0.87***	3.70
OI2 – Conduct dialogue session with consumer frequently	0.48*		4.11
OI3 – Involve internal parties during the dialogue session with consumers	0.01	0.81***	4.94
OI4 – Involve external parties during the dialogue session with consumers	0.12	0.65***	2.39
OI5 – Recognize the consumer's experience regarding to the service product	0.08	0.80***	2.99
OI6 – Emphasize the employees' effort to individual consumers	0.34	0.81***	2.34
Access (first-order)	0.14		3.55
OI7 – Offer opportunity to the consumers to share in the design process of service product	0.23	0.63***	1.50
OI8 – Emphasize more on providing experiences to the consumers than the ownership of service product	0.31*		1.83
OI9 – Provide all the necessary service product related information to the consumers	0.67***		1.37
Risk (first-order)	0.59**		2.68
OI10 – Inform potential risks of the service product offered to the consumers	0.22	0.82***	2.55
OI11 – Inform consumers about the limitation of the firm's knowledge and capability	0.53**		2.55
OI12 – Recognize the changing dynamics of consumers' need	0.39*		2.14
OI13 – Accept the consumers' complaints on service product offerings	0.11	0.71***	2.54
OI14 – Shoulder all the risk-related responsibilities upon themselves	−0.161	0.48**	1.89
Transparency (first-order)	−0.118		2.20
OI14 – Make clear to the consumers about the service product-related information	0.38*		1.77
OI15 – Disclose pricing-related information to the consumers	0.02	0.55***	1.63

(*continued*)

Constructs	Weight	Loading	VIF
OI16 – Get benefit from the information symmetry between the consumers and the firm	0.27	0.81***	2.32
OI17 – Build trust among the consumers through transparent information	0.45	0.89***	3.53
OI18 – Provide up-to-date information to consumers	0.07	0.85***	3.85
Financial performance			
FP1 – Return on investment	0.41	0.81***	2.35
FP2 – Earnings growth	0.11	0.70***	2.59
FP3 – Sales growth	0.14	0.63***	1.67
FP4 – Market share change	0.57**		1.51
Non-financial performance			
NFP5 – Customer satisfaction improvement	0.30	0.77***	1.52
NFP6 – Corporate image improvement	0.21	0.68***	1.52
NFP7 – Brand value improvement	0.32	0.83***	1.91
NFP8 – Employee capacity improvement	0.43**		1.56

Note(s): *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

Table 1.

Table 2.
Structural equations results

Trajectories	Coefficient	*t*-value
Direct effect		
OI → BDAC (R^2: 0.42)	0.46***	5.646
OI → FP	0.24*	1.766
OI → NFP	0.37**	2.992
BDAC → FP (R^2: 0.31)	0.40**	2.527
BDAC → NFP (R^2: 0.45)	0.41**	2.978
Indirect effects		
H1. OI → BDAC → FP (VAF = 43%)	0.18*	2.04
H2. OI → BDAC → NFP (VAF = 34%)	0.19*	2.32
Control variables		
Age → FP	0.20	1.596
Size → FP	−0.070	0.566
Age → NFP	0.08	0.740
Size → NFP	−0.040	0.353

Note(s): *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

Hence, this result means a part of the OI positive effect directly reaches the two firm performance aspects; however, another part of this effect requires the intermediation of BDAC to be able to become a real impact on firm performance. In other words, part of this big data flow and information coming from OI with customers is at a high risk of being wasted and will not be translated into an improvement in financial and non-financial results, unless BDAC intervenes and takes advantage of these data and generates insights that ultimately become better decisions and solutions.

With the aim of dimensioning how important this BDAC mediating role is, the variance accounted for (VAF) test was carried out (Table 2), allowing to establish the magnitude of the indirect effect with respect to the total one. In the case of FP, 43% of its variance is explained by the indirect relationship via the mediator variable, whereas in the case of NFP, such percentage is 34%. This indicates that BDAC mediation is partial in both cases since the percentages do not exceed 80% (Hair *et al.*, 2017).

Discussion and conclusions

Contrary to what was expected, BDAC mediation is partial, and the OI process of acquisition, particularly customer involvement, has a positive and significant influence on firm performance. This was an unforeseen result as the existence of such direct effect had been ruled out entirely. However, the fact that BDAC mediation represents 43 and 34% of the total effect of OI on FP and NFP, respectively, indicates this variable plays a complementary role that cannot be overlooked in the context of medium- and low-technology companies. BDAC is, in effect, maximizing the financial and non-financial use of the large flow of data produced by customer involvement.

Anyway, BDAC mediation is higher in the relationship between OI and FP. This fact could explain how BDAC improves the decision-making process inside organizations regarding markets, competitors and others, as well as how its effect impacts financial variables related to revenue, sales and cost reduction in a faster manner (Caputo *et al.*, 2019). Nevertheless, that effect could appear more delayed for non-financial variables such as customer satisfaction, reputation and innovation processes, because these variables involve other human components, such as knowledge management and absorptive capacity, which require more time to improve (Ahn *et al.*, 2016).

This article contributes to the literature showing evidence that BDAC must be understood in its relationship to OI and firm performance, not just from a technological perspective, as it has been predominantly considered, but as an organizational skill and a resource to develop the OI process successfully to finally impact firm performance. In this context, BDAC is not simply related to technological conditions that support the decision-making process, but to a capability that could drive other internal or external processes. Above all BDAC is a strategic skill that impacts OI by driving interactions and information gathering that helps companies to create networks and encourage collaborative working, which impacts OI according to this paradigm. This ultimately boosts innovation in products or services, impacting sales, customer satisfaction or market share, and having a positive effect on firm performance.

Thus, the main contribution of the study lies in the fact that it offers a new, broader study perspective on the role of BDAC in the relationship between OI and firm performance. Ultimately, our study questions the research tradition adopted by recent studies that have resorted to reducing this role to a simple application of data analytics techniques or identifying specific solutions to benefit from the knowledge flow coming from external sources (Urbinati *et al.*, 2019) or those highlighting the importance of certain external data sources (Papa *et al.*, 2021); Zhang and Xiao, 2020). Instead, our results show BDAC is far more than the application of a set of techniques. It rather is an organizational skill that should be aligned and highly articulated with the outputs of key processes – namely collaborative work with customers – yielding the data flow with the greatest potential to impact the firm performance of medium- and low-technology companies.

In consequence, a first practical recommendation is to rethink altogether the storage of structured or unstructured data coming from collaborative work with customers. Until now, storage has been carried out to facilitate consultation by other members of the organization (Wen *et al.*, 2020), *e.g.* by creating a bank of lessons learned in the corporate intranet. Nonetheless, our results compel us to understand storage as an enabler of big data analytics. Based on our results, we therefore consider the prioritization of data storage in the cloud – which greatly facilitates the implementation of big data analytics solutions – to be fitting. For example, Amazon offers Hadoop as a service (HaaS); it is a big data analytics framework that stores and analyzes data in the cloud using the Hadoop software library. This service can be very useful for companies that are just entering the world of analytics.

On the other hand, a second recommendation concerns the ongoing training of employees, mainly the product and process innovation team that interacts with customers, so that they are familiar with the main big data analytics solutions available in the market, such as HaaS. In that sense, our recommendation is not to turn the innovation team into data scientists, but to broaden their understanding of the potential of all current and emerging analytics tools, particularly their purposes, uses and benefits, so that the team can even utilize them proactively to take advantage of the large flow of data produced by customer involvement.

The main limitation of our work is that it is confined to one of the several activities associated with the OI process of acquisition or inbound, namely, collaborative work with customers. However, there are other activities within this process such as outsourcing technology services, inward IP licensing, supplier involvement, among others, from which large amounts of data and information emerge. Moreover, the flow of data produced by the OI process of exploitation or outbound is not within the scope of the present work.

On the other hand, the work was conducted in an emerging country where medium- and low-technology sectors predominate, as reflected by our sample of firms. Consequently, the extrapolation of our results to the context of developed countries is limited, as high-tech firms are more common in the latter and OI influences firm performance, albeit under different conditions, given the presence of other types of relationships between both constructs at different points in time (Fu *et al.*, 2019).

Future studies should, thus, strive to examine the mediating role of BDAC in the relationship between the other OI acquisition activities and firm performance. Secondly, it would also be worth analyzing the mediating role of BDAC in the relationship between the OI process of exploitation or outbound and firm performance, a work that could potentially make a great contribution to the field of OI in the digital age. Similarly, analyzing whether BDAC plays a moderating role in the relationship between OI and firm performance in high-tech firms would also be worth exploring.

References

Ahn, J.M., Ju, Y., Moon, T.H., Minshall, T., Probert, D., Sohn, S.Y. and Mortara, L. (2016), "Beyond absorptive capacity in open innovation process: the relationships between openness, capacities and firm performance", *Technology Analysis and Strategic Management*, Routledge, Vol. 28 No. 9, pp. 1009-1028.

Alnoukari, M. and Hanano, A. (2017), "Integration of business intelligence with corporate strategic management", *Journal of Intelligence Studies in Business*, Vol. 7 No. 2, doi: 10.37380/jisib.v7i2.235.

Arias-Pérez, J., Alegre, J. and Villar, C. (2021a), "Triggering open innovation processes through organizational emotional capability and rival's absorptive capacity orientation", *IEEE Transactions on Engineering Management*, pp. 1-11, doi: 10.1109/TEM.2019.2955439.

Arias-Pérez, J., Velez-Ocampo, J. and Cepeda-Cardona, J. (2021b), "Strategic orientation toward digitalization to improve innovation capability: why knowledge acquisition and exploitation through external embeddedness matter?", *Journal of Knowledge Management*. doi: 10.1108/JKM-03-2020-0231.

Bag, S., Pretorius, J.H.C., Gupta, S. and Dwivedi, Y.K. (2021), "Role of institutional pressures and resources in the adoption of big data analytics powered artificial intelligence, sustainable manufacturing practices and circular economy capabilities", *Technological Forecasting and Social Change*, Vol. 163, p. 120420.

Baron, R.M. and Kenny, D.A. (1986), "The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations", *Journal of Personality and Social Psychology*, Vol. 51 No. 6, pp. 1173-1182.

Beretta, M. (2019), "Idea selection in web-enabled ideation systems", *Journal of Product Innovation Management*, Vol. 36 No. 1, pp. 5-23.

Brynjolfsson, E., Hitt, L. and Kim, H. (2011), "Strength in numbers: how does data-driven decision-making affect firm performance?", *International Conference on Information Systems 2011, ICIS 2011*, Vol. 1, pp. 541-558.

Burchardt, C. and Maisch, B. (2019), "Digitalization needs a cultural change – examples of applying Agility and Open Innovation to drive the digital transformation", *Procedia CIRP*, Vol. 84, pp. 112-117.

Cappa, F., Oriani, R., Pinelli, M. and De Massis, A. (2019), "When does crowdsourcing benefit firm stock market performance?", *Research Policy*, Vol. 48 No. 9, p. 103825.

Caputo, M., Lamberti, E., Cammarano, A. and Michelino, F. (2016), "Exploring the impact of open innovation on firm performances", *Management Decision*, Emerald Group Publishing, Vol. 54 No. 7, pp. 1788-1812.

Caputo, F., Cillo, V., Candelo, E. and Liu, Y. (2019), "Innovating through digital revolution: the role of soft skills and Big Data in increasing firm performance", *Management Decision*, Emerald Group Publishing, Vol. 57 No. 8, pp. 2032-2051.

Chadwick, I.C. and Dawson, A. (2018), "Women leaders and firm performance in family businesses: an examination of financial and nonfinancial outcomes", *Journal of Family Business Strategy*, Vol. 9 No. 4, pp. 238-249.

Chen, H., Chiang, R.H.L. and Storey, V.C. (2012), "Business intelligence and analytics: from big data to big impact", *MIS Quarterly: Management Information Systems*, University of Minnesota, Vol. 36 No. 4, pp. 1165-1188.

Chesbrough, H. and Bogers, M. (2014), "Explicating open innovation: clarifying an emerging paradigm for understanding innovation", in Chesbrough, H., Vanhaverbeke, W. and West, J. (Eds), *New Frontiers in Open Innovation*, Oxford University Press, Oxford, pp. 3-28.

Ciampi, F., Demi, S., Magrini, A., Marzi, G. and Papa, A. (2021), "Exploring the impact of big data analytics capabilities on business model innovation: the mediating role of entrepreneurial orientation", *Journal of Business Research*, Vol. 123, pp. 1-13.

Dahiya, R., Le, S., Ring, J.K. and Watson, K. (2021), "Big data analytics and competitive advantage: the strategic role of firm-specific knowledge", *Journal of Strategy and Management*, Emerald Publishing. doi: 10.1108/JSMA-08-2020-0203.

Davenport, T.H., Harris, J.G. and Morison, R. (2010), *Analytics at Work : Smarter Decisions, Better Results*, Harvard Business Press, Boston, p. 240.

Del Vecchio, P., Di Minin, A., Petruzzelli, A.M., Panniello, U. and Pirri, S. (2018), "Big data for open innovation in SMEs and large corporations: trends, opportunities, and challenges", *Creativity and Innovation Management*, Life Science Publishing, Vol. 27 No. 1, pp. 6-22.

Demchenko, Y., Grosso, P., de Laat, C. and Membrey, P. (2013), "Addressing big data issues in scientific data infrastructure", *2013 International Conference on Collaboration Technologies and Systems (CTS)*, pp. 48-55.

Dogbe, C.S.K., Tian, H., Pomegbe, W.W.K., Sarsah, S.A. and Otoo, C.O.A. (2020), "Effect of network embeddedness on innovation performance of small and medium-sized enterprises", *Journal of Strategy and Management*, Emerald Publishing, Vol. 13 No. 2, pp. 181-197.

Dumbill, E. (2012), "Making sense of big data", *Big Data*, Mary Ann Liebert, Vol. 1 No. 1, pp. 1-2.

Enkel, E., Bogers, M. and Chesbrough, H. (2020), "Exploring open innovation in the digital age: a maturity model and future research directions", *R&D Management*, Vol. 50 No. 1, pp. 161-168.

Ferraris, A., Mazzoleni, A., Devalle, A. and Couturier, J. (2019), "Big data analytics capabilities and knowledge management: impact on firm performance", *Management Decision*, Emerald Group, Vol. 57 No. 8, pp. 1923-1936.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G. and 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*, Vol. 165, pp. 234-246.

Fu, L., Liu, Z. and Zhou, Z. (2019), "Can open innovation improve firm performance? An investigation of financial information in the biopharmaceutical industry", *Technology Analysis and Strategic Management*, Routledge, Vol. 31 No. 7, pp. 776-790.

Gandomi, A. and Haider, M. (2015), "Beyond the hype: big data concepts, methods, and analytics", *International Journal of Information Management*, Vol. 35 No. 2, pp. 137-144.

Gantz, J. and Reinsel, D. (2011), "Extracting value from chaos", *IDC Iview*, Vol. 1142 No. 2011, pp. 1-12.

Grant, R.M. (1996), "Toward a knowledge-based theory of the firm", *Strategic Management Journal*, John Wiley & Sons, Vol. 17 No. S2, pp. 109-122.

Gupta, M. and George, J.F. (2016), "Toward the development of a big data analytics capability", *Information and Management*, Vol. 53 No. 8, pp. 1049-1064.

Gupta, S., Drave, V.A., Dwivedi, Y.K., Baabdullah, A.M. and Ismagilova, E. (2020), "Achieving superior organizational performance via big data predictive analytics: a dynamic capability view", *Industrial Marketing Management*, Vol. 90, pp. 581-592.

Gurumurthy, R., Schatsky, D. and Camhi, J. (2020), *Uncovering the Connection between Digital Maturity and Financial Performance*, Deloitte Insights, New York, p. 23.

Hagel, J. (2015), "Bringing analytics to life", *Journal of Accountancy*, Vol. 219 No. 2, pp. 1-24.

Hair, J.F., Hult, G.T.M., Ringle, C. and Sarstedt, M. (2017), *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, Sage Publications, Los Angeles.

Hair, J., Risher, J., Sarstedt, M. and Ringle, C. (2019), "When to use and how to report the results of PLS-SEM", *European Business Review*, Emerald, Vol. 31 No. 1, pp. 2-24.

Henao-García, E., Arias-Pérez, J. and Lozada, N. (2021), "Fostering big data analytics capability through process innovation: is management innovation the missing link?", *Business Information Review*, SAGE Publications, Vol. 38 No. 1, pp. 28-39.

Hung, K.-P. and Chou, C. (2013), "The impact of open innovation on firm performance: the moderating effects of internal R&D and environmental turbulence", *Technovation*, Vol. 33 No. 10, pp. 368-380.

Ibarra-Cisneros, M.-A., Demuner-Flores, M., del, R. and Hernández-Perlines, F. (2021), "Strategic orientations, firm performance and the moderating effect of absorptive capacity", *Journal of Strategy and Management*, Emerald Publishing. doi: 10.1108/JSMA-05-2020-0121.

Kogut, B. and Zander, U. (1992), "Knowledge of the firm, combinative capabilities, and the replication of technology", *Organization Science*, INFORMS, Vol. 3 No. 3, pp. 383-397.

Lee, Y., Kim, S. and Lee, H. (2011), "The impact of service R&D on the performance of Korean information communication technology small and medium enterprises", *Journal of Engineering and Technology Management*, Vol. 28 No. 1, pp. 77-92.

Liao, S., Fu, L. and Liu, Z. (2020), "Investigating open innovation strategies and firm performance: the moderating role of technological capability and market information management capability", *Journal of Business and Industrial Marketing*, Emerald Publishing, Vol. 35 No. 1, pp. 23-39.

Liu, Y. (2014), "Big Data and predictive business analytics", *Journal of Business*, Vol. 33 No. 4, pp. 40-42.

Lozada, N., Arias-Pérez, J. and Perdomo-Charry, G. (2019), "Big data analytics capability and co-innovation: an empirical study", *Heliyon*, Vol. 5 No. 10, e02541.

Meisinger, N. and Moldaschl, M. (2020), "Reduced to the max: firm performance and organizational ambidexterity research", *Journal of Strategy and Management*, Emerald Publishing, Vol. 14 No. 1, pp. 96-106.

Mikalef, P., Pappas, I.O., Krogstie, J. and Giannakos, M. (2018), "Big data analytics capabilities: a systematic literature review and research agenda", *Information Systems and e-Business Management*, Vol. 16 No. 3, pp. 547-578.

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*, Vol. 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*, Blackwell Publishing, Vol. 30 No. 2, pp. 272-298.

Moretti, F. and Biancardi, D. (2020), "Inbound open innovation and firm performance", *Journal of Innovation and Knowledge*, Vol. 5 No. 1, pp. 1-19.

Nara, E.O.B., da Costa, M.B., Baierle, I.C., Schaefer, J.L., Benitez, G.B., do Santos, L.M.A.L. and Benitez, L.B. (2021), "Expected impact of industry 4.0 technologies on sustainable development: a study in the context of Brazil's plastic industry", *Sustainable Production and Consumption*, Vol. 25, pp. 102-122.

O'Connor, C. and Stephen, K. (2017), "Facilitating knowledge management through filtered big data: SME competitiveness in an agri-food sector", *Journal of Knowledge Management*, Emerald Publishing, Vol. 21 No. 1, pp. 156-179.

Papa, A., Chierici, R., Ballestra, L.V., Meissner, D. and Orhan, M.A. (2021), "Harvesting reflective knowledge exchange for inbound open innovation in complex collaborative networks: an empirical verification in Europe", *Journal of Knowledge Management*, Vol. 25 No. 4, pp. 669-692, doi: 10.1108/JKM-04-2020-0300.

Peter, M.K., Kraft, C. and Lindeque, J. (2020), "Strategic action fields of digital transformation", *Journal of Strategy and Management*, Emerald Publishing, Vol. 13 No. 1, pp. 160-180.

Rialti, R., Marzi, G., Ciappei, C. and Busso, D. (2019), "Big data and dynamic capabilities: a bibliometric analysis and systematic literature review", *Management Decision*, Emerald Group Publishing, Vol. 57 No. 8, pp. 2052-2068.

Schläfke, M., Silvi, R. and Möller, K. (2013), "A framework for business analytics in performance management", *International Journal of Productivity and Performance Management*, Vol. 62 No. 1, pp. 110-122.

Shamim, S., Zeng, J., Khan, Z. and Zia, N.U. (2020), "Big data analytics capability and decision making performance in emerging market firms: the role of contractual and relational governance mechanisms", *Technological Forecasting and Social Change*, Vol. 161, p. 120315.

Shukla, N., Tiwari, M.K. and Beydoun, G. (2019), "Next generation smart manufacturing and service systems using big data analytics", *Computers and Industrial Engineering*, Vol. 128, pp. 905-910.

Sisodiya, S.R., Johnson, J.L. and Grégoire, Y. (2013), "Inbound open innovation for enhanced performance: enablers and opportunities", *Industrial Marketing Management*, Vol. 42 No. 5, pp. 836-849.

Taghizadeh, S.K., Jayaraman, K., Ismail, I. and Rahman, S.A. (2016), "Scale development and validation for DART model of value co-creation process on innovation strategy", *Journal of Business and Industrial Marketing*, Vol. 31 No. 1, pp. 24-35.

Urbinati, A., Bogers, M., Chiesa, V. and Frattini, F. (2019), "Creating and capturing value from Big Data: a multiple-case study analysis of provider companies", *Technovation*, Vol. 84 No. 85, pp. 21-36.

van de Vrande, V., de Jong, J.P.J., Vanhaverbeke, W. and de Rochemont, M. (2009), "Open innovation in SMEs: trends, motives and management challenges", *Technovation*, Vol. 29 No. 6, pp. 423-437.

Van De Wetering, R., Mikalef, P. and Krogstie, J. (2019), "Strategic value creation through big data analytics capabilities: a configurational approach", *Proceedings - 21st IEEE Conference on Business Informatics, CBI 2019*, Vol. 1, Institute of Electrical and Electronics Engineers Inc., pp. 268-275.

Vidgen, R., Shaw, S. and Grant, D.B. (2017), "Management challenges in creating value from business analytics", *European Journal of Operational Research*, Elsevier B.V., Vol. 261 No. 2, pp. 626-639.

Wang, C.-H., Chang, C.-H. and Shen, G.C. (2015), "The effect of inbound open innovation on firm performance: evidence from high-tech industry", *Technological Forecasting and Social Change*, Vol. 99, pp. 222-230.

Wen, X., Wu, G., Kang, Q., Wang, L. and Zeng, J. (2020), "A study on customer knowledge management, inbound open innovation and firm performance", *Human Systems Management*, IOS Press, Vol. 39, pp. 183-195.

Zaborek, P. and Mazur, J. (2017), "Exploring links between engaging customers in value Co-creation and product innovativeness", *International Journal of Management and Economics*, Warsaw School of Economics, Collegium of World Economy, Vol. 53 No. 3, pp. 82-106.

Zhang, H. and Xiao, Y. (2020), "Customer involvement in big data analytics and its impact on B2B innovation", *Industrial Marketing Management*, Vol. 86, pp. 99-108.

Zhao, X., Lynch John, G.J. and Chen, Q. (2010), "Reconsidering Baron and Kenny: myths and truths about mediation analysis", *Journal of Consumer Research*, Vol. 37 No. 2, pp. 197-206.

Zhou, X., Guo, Y., Li, F., Wang, J., Wei, H., Yu, M. and Chen, S. (2020), "Identifying and assessing innovation pathways for emerging technologies: a hybrid approach based on text mining and altmetrics", *IEEE Transactions on Engineering Management*, pp. 1-12, doi: 10.1109/TEM.2020.2994049.

Corresponding author
José Arias-Pérez can be contacted at: jenrique.Arias@udea.edu.co