


Article

Big Data, Big Data Analytics Capability, and Sustainable Innovation Performance

Shengbin Hao ¹, Haili Zhang ^{2,*} and Michael Song ² ¹ School of Management, Harbin Institute of Technology, Harbin 150001, China; haoshengbin@hit.edu.cn² School of Economics and Management, Xi'an Technological University, Xi'an 720021, China; michael.song@xatu.edu.cn

* Correspondence: zhanghaili@xatu.edu.cn

Received: 17 November 2019; Accepted: 10 December 2019; Published: 13 December 2019



Abstract: Literature suggests that big data is a new competitive advantage and that it enhance organizational performance. Yet, previous empirical research has provided conflicting results. Building on the resource-based view and the organizational inertia theory, we develop a model to investigate how big data and big data analytics capability affect innovation success. We show that there is a trade-off between big data and big data analytics capability and that optimal balance of big data depends upon levels of big data analytics capability. We conduct a four-year empirical research project to secure empirical data on 1109 data-driven innovation projects from the United States and China. This research is the first time reporting the empirical results. The study findings reveal several surprising results that challenge traditional views of the importance of big data in innovation. For U.S. innovation projects, big data has an inverted U-shaped relationship with sales growth. Big data analytics capability exerts a positive moderating effect, that is, the stronger this capability is, the greater the impact of big data on sales growth and gross margin. For Chinese innovation projects, when big data resource is low, promoting big data analytics capability increases sales growth and gross margin up to a certain point; developing big data analytics capability beyond that point may actually inhibit innovation performance. Our findings provide guidance to firms on making strategic decisions regarding resource allocations for big data and big data analytics capability.

Keywords: big data; big data analytics capability; sustainable innovation performance

1. Introduction

In recent years, the accumulation and application of big data have attracted much attention. Both academics and executives have emphasized the importance of big data. Many firms have invested heavily in big data with the aim of deriving important insights to gain sustainable competitive advantages. Recent literature has suggested that the innovation process need to create sustainable new capabilities such as big data analytics capability (BDAC) [1–4]. BDAC can help companies process the enormous amounts of available data to identify market opportunities, predict customer needs, and analyze customer purchase decisions [5,6]. Indeed, the role of BDAC in the innovation process is becoming increasingly prominent in practice and acknowledged in the literature. Past studies suggest that BDAC capabilities are new sustainable capabilities [5,7–10]. Despite the hype surrounding big data, limited attention has been paid to the mechanisms and processes through which big data and BDAC add value to firms.

Sustainable organizational growth and performance is becoming a hot topic in the literature. The literature has suggested that sustainable organizations need to create sustainable new capabilities such as BDAC [1–4]. Prior studies have suggested that big data is an important sustainable resource for enterprises because they need new technologies to manage and analyze massive data and information [3].

The emergence of BDAC has changed the original production and operation mode of enterprises. Many large enterprises wish to develop sustainable big data analysis capability to improve their market competitiveness in order to survive [11]. Since BDAC can help enterprises to use advanced analytical skills to extract valuable information from big data, it can help enterprises obtain high operational efficiency and sustainable performance. However, past studies suggest that if organizations do not pay attention to the sustainability of BDAC, the social impact will be temporary [2]. BDAC can promote sustainable development in many fields including the Internet of Things, social networks, sustainable development [1], and sustainable investment in supply chain [3].

Although literature suggests that big data and BDAC are new sustainable competitive advantages that enhance sustainability of organizational development, empirical research is still lacking and the limited research has provided conflicting results on the value of big data for sustainable innovation and organizational development. The issues of whether and under what conditions investments in big data can enhance firm performance remain unsettled [12]. There is very little empirical evidence to guide managers on how to leverage big data resources to achieve sustainability of organizational development and performance. We aim to fill the literature gap by investigating how big data relates to sustainability of innovation and organizational development and how big data analytics capability (BDAC) may moderate this relationship. Thus, four research questions are developed.

RQ1: Does big data have an inverted U-shaped relationship with sustainability of innovation and organizational development?

RQ2: Does BDAC strengthen or weakens the effects of big data on sustainability of innovation and organizational development?

To answer these research questions, we build on two of the most important theories of organizational development: the resource-based theory and the organizational inertia theory. We propose that big data may have an inverted U-shaped relationship with sustainability of innovation and organizational development. That is, a moderate level of big data resources is optimal, but a high level of big data resources actually decreases sustainability of innovation and organizational development. BDAC may help managers better leverage big data resources and overcome the inertia in resources and routines. As a result, BDAC has the potential to strengthen the positive effects of big data on sustainability of innovation and organizational development.

We test these hypotheses using data from 477 U.S. and 632 Chinese service innovation projects as part of a multi-year panel study of international innovation projects. From the database, we selected service innovation projects from each participating company in five industries (hotel, traveling, and tourism services; banking, insurances, securities, financial investments, and related activities; information and semiconductor; Internet-related services; and health care services) in two countries (the United States and China). The data reported in this study included only the projects with sales growth rate and project gross margin for first three years of the commercialization of the innovation projects.

There were three parts to our research design. First, to ensure the suitability of the research scales and methods for a cross-national comparative study of the theoretical model, we followed the procedures recommended by [13] and extended by [14]. Second, to test the causal relationships between variables, we collected data on big data resources and BDAC. Finally, after commercialization of each project, we collected track sales and gross profits data for three years.

This study makes several contributions to theory and practice in the fields of big data and sustainability of innovation and organizational development. First, this study enriches our knowledge of big data by revealing the linear and curvilinear relationships between big data and sustainability of innovation and organizational development. Second, by examining the moderating role of BDAC, this study responds to the recent call for uncovering the conditions under what big data enhances sustainability of innovation and organizational development. Third, the research findings of this study suggest how firms can realize the full potential of their big data resources.

Our study findings also provide some practical insights for innovation project leaders regarding big data resources and BDAC. It is suggested that project leaders in both U.S and Chinese countries

can increase innovation project performance by ensuring that the project teams have adequate big data resources on customer needs, user behavior, competitive intelligence (e.g., marketing, products, services, etc.), and technology development intelligence.

Although big data can indeed increase sustainable sales growth and sustainable gross margin, managers in U.S. service innovation teams must be aware of the limitations of their big data resources and understand that too much investment in big data may actually hurt sales growth. The innovation teams also need to build their BDAC to enable them to harness the potential of their big data resources. For U.S. innovation teams, devoting resources to BDAC is preferable to additional investment in big data. The stronger this capability is, the greater effect big data can have in improving sales growth and gross margin.

Managers of Chinese innovation projects should be aware that developing BDAC might have its downsides. When investment in big data is low, promoting BDAC can increase sales growth and gross margin. However, after a certain point, allocating further resources to BDAC may inhibit sustainability of innovation and organizational development. Chinese managers may also use the suggested tactics to increase big data analytics capability.

2. Theoretical Background and Hypotheses

2.1. Theoretical Model

Organizational inertia refers to how the patterns and processes that firms have established to maximize the efficiency of their business operations can create strong internal resistance against radical changes [15]. Gilbert [16] has proposed that there are two types of organizational inertia: resource and routine. The former refers to inertia in resource allocation patterns, and the latter pertains to inertia in the organizational processes that use the resources. Big data refers to high-volume, high-velocity, and/or high-variety information assets [17] that cannot be easily processed using traditional methods [17]. Service innovation projects' investments in big data involve the resources allocated to information assets and the organizational processes that use these assets, both of which are subject to inertial pressures.

The resource-based view suggests that resources are important for firms to gain sustainable competitive advantage [18]. Sirmon et al. [19] developed the resource management model suggesting that resource management contributes more to superior performance than simply owning resources. Following this logic, we argued that merely investing in big data resources is not enough; innovation projects must also possess the capability to process, manage, and deploy their big data resources—that is, they need BDAC [10]. As such, this dynamic business intelligence capability [20] has the potential to strengthen the impact of big data resources on sustainability of innovation and organizational development.

We drew from the resource-based view and organizational inertia theory to construct a theoretical model (shown in Figure 1) illustrating how big data resources influence sustainability of innovation and organizational development. Our framework also proposed that the effect of big data on sustainability of innovation and organizational development depends on the level of BDAC.

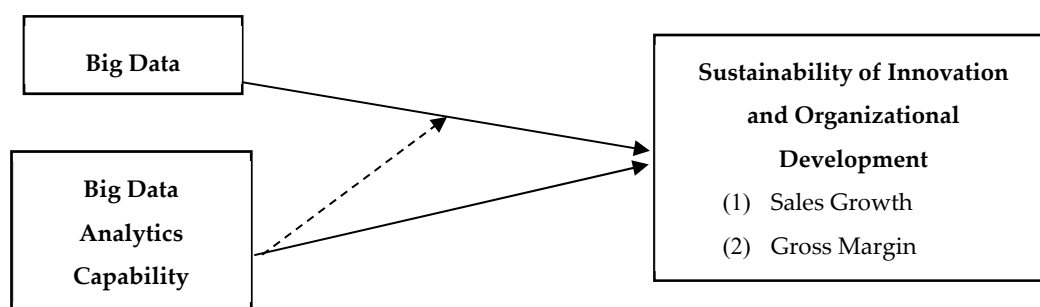


Figure 1. Theoretical framework.

2.2. Research Hypothesis 1: Big Data and Innovation Performance

The resource-based view explains how firms achieve sustainable competitive advantage through owning resources [18] and thereby hypothesizes that big data can provide firms with sustainable competitive advantage. Indeed, an emerging stream of research has highlighted that big data positively affects sustainability of innovation and organizational development. For example, White [21] suggested that big data enables new solutions that bring significant benefits to the business. McAfee et al. [12] found that firms that are more data driven have better performance outcomes. Gupta and George [10] found that big data helps managers generate knowledge and transform it into efficient decision making, resulting in improved performance. New projects, whose resources and external networks are limited [22], can extract important information from big data to identify profitable customers and reliable suppliers, detect quality problems, set appropriate prices, and determine optimal inventory levels. Big data enable new service development projects to attract more resources and develop external linkage, which are vital for new projects to grow and succeed.

However, we hypothesized that the positive effect of big data on sustainability of innovation and organizational development may decline after the level of big data reaches a critical point for two reasons. First, because of resource and routine inertia [16], when new projects increase their investment in big data, they may depend too much on the big data assets to gain information and knowledge and thereby fail to alter the way they allocate resources, overlook other resource allocation patterns, and inhibit their ability to adapt to changing environmental conditions. Second, new projects may eventually be constrained in garnering big data returns from simply increasing investments. These constraints involve diminishing and negative returns of increased big data resources. Excessive investment in big data may confront new projects with information overload and overstretch available resources. More importantly, new projects are often resource constrained, which will undermine the returns from big data and decrease service performance. Thus, we proposed the following hypothesis:

Hypothesis 1 (H1). *Big data has an inverted U-shaped relationship with sustainability of innovation and organizational development.*

2.3. Research Hypothesis 2: The Moderating Effect of Big Data Analytics Capability

BDAC refers to a firm's proficiency in managing and leveraging its big data resources to enhance performance [10,23]. We hypothesized that BDAC enhances the positive effect of big data on sustainability of innovation and organizational development. That is, as BDAC increases, the effect of big data resources on sustainability of innovation and organizational development increases. First, the resource management model asserts that resources alone do not guarantee sustainable competitive advantage [24,25]; rather, resources must be managed appropriately to derive benefit from them [26]. BDAC entails a firm's proficiency in utilization of big data to enhance sustainability of innovation and organizational development. When the relationship between big data and performance is positive, stronger BDAC enables NPD project teams to extract value from big data, thus leveraging the full potential of the big data resources. We posited above that with an increase in big data resources, the relationship between big data and service performance turns negative, but stronger BDAC can help innovation teams deal with information overload and thus mitigate the diminishing and negative returns of increased big data resources.

Second, BDAC allows innovation teams to integrate and reconfigure their information resources and business processes to adapt to rapidly changing environments. In so doing, BDAC can help innovation teams overcome their resource and routine inertia. Taken together, these arguments suggest that BDAC can play a facilitating role in the relationship between big data resources and innovation performance. Therefore, we proposed the following hypothesis:

Hypothesis 2 (H2). *BDAC strengthens the effects of big data on sustainability of innovation and organizational development.*

3. Materials and Methods

To test our model, we collected data as part of a multi-year panel study. From the database, we selected service innovation projects from each participating company in five industries (hotel, traveling, and tourism services; banking, insurances, securities, financial investments, and related activities; information and semiconductor; Internet-related services; and health care services) in two countries (China and the United States). The data reported in this study included only the projects with sales growth rate and project gross margin for the first three years after the commercialization.

3.1. Overall Research Design

There were three parts to our research design. First, to ensure the suitability of the research scales and methods for a cross-national comparative study of a theoretical model, we followed the procedures recommended by [13] and extended by [14]. Second, to test the causal relationships between variables, we collected data on big data resources and big data analytics capabilities. Finally, we tracked the projects over time and collected sales and gross profits data for the first three years after the commercialization.

3.2. Measurement Development Procedures

We developed the measurement scale using the procedures reported by Song and Parry [14]. We conducted in-depth case studies and focus group interviews with service innovation teams in China and the United States, follow-up interviews with team members, and consultations with academic experts from two national business administration schools. We used our findings to improve measurement scales from the literature on big data and big data analytics capabilities.

In the focus group interviews, we included open-ended and semi-structured questions. We examined the conceptual equivalence of the constructs with the first set of questions. Team members were requested to define the key constructs of big data and big data analytics capabilities [13,14]. We assessed the functional equivalence of the constructs with the second set of questions [13,14]. Team members evaluated how well our theoretical model describes their won service innovation experience [13,14]. We assessed the category equivalence of the constructs with the third set of questions, which addressed the team members' perceptions of the relevance and completeness of the scale items from the literature and previous case studies [13,14]. The results of these case studies and interviews indicate that several scales that have been used in academic research should be modified for cross-national comparative research on the service innovation process.

Following Churchill's procedures for developing better measures [27], we asked academic experts and service innovation project managers in two countries to review a list of our constructs and corresponding measurement items and provide critical assessments and recommendations for additional measurement items [14]. After a small modification and a second review of the items by the scholars and managers, we constructed the questionnaires.

3.3. Variable Measurement

Our study variables include big data, big data analytics capabilities, and project performance. Except for project performance, which was measured using objective data, the other variables were assessed using scales we adopted from the literature or developed based on our measurement development as described. Prior studies suggest that a 0–10 rating scale reflects the metric system in structure and is thus easier to understand in international surveys than 1–7 or 1–5 rating scales [14,15,28]. Therefore, we directed the participants in this study to respond to all survey measures using a 0–10 scale, as shown in the Appendix A.

Our measurement scale for big data resources includes five items that assess how adequate the project team has the big data resources on customer needs, user behavior, competitive intelligence (e.g., marketing, products, services, etc.), technology development intelligence, and product usages.

Project team leaders were requested to provide assessment of the team's big data resources in these areas (0 = not adequate for this development project; 10 = much more than adequate for this development project). These measures were developed based on the measurement development described earlier.

Our measurement scale for big data analytics capabilities included five items: (1) we have advanced tools (analytics and algorithms) to extract value from big data [10,29]; (2) we have the capability to discover relationships and dependencies from big data; (3) we have the capability to perform predictions of outcomes and behaviors from big data [10]; (4) we have the capability to discover new correlations from big data to spot market demand trends and predict user behavior [5,9]; and (5) our big data analytics staff has the right skills to accomplish their jobs successfully [10]. Project team leaders were requested to provide assessment of the team's capabilities in these areas (0 = no capability; 10 = very high level of capability).

The measure for the average sales growth rate was adopted from [30]. The scale for the average gross margin was adopted from [31]. These variables are calculated using following formulas:

Average Sales growth rate (%; adopted from [30])

$$= (\text{second year sales growth rate} + \text{third year sales growth rate})/2,$$

where,

$$\text{second year sales growth rate} = 100 \times (\text{second year sales} - \text{first year sales})/(\text{first year sales});$$

$$\text{third year sales growth rate} = 100 \times (\text{third year sales} - \text{second year sales})/(\text{second year sales}).$$

Average gross margin (in %; adopted from [31])

$$= (\text{first year gross margin} + \text{second year gross margin} + \text{third year gross margin})/3,$$

where

$$\text{gross margin} = 100 \times (\text{total sales} - \text{total variable costs})/(\text{total sales}).$$

3.4. Data Collection Procedures

To ensure that the participants in both countries would have the same understanding of the scale items, we used the double translation method to translate the questionnaire [14]. Two translators translated the English version into Chinese and then two other translators independently translated the Chinese version into English. The four translators discussed and resolved inconsistencies they found and also consulted with several case study participants to determine the appropriateness of the Chinese questionnaire. We then conducted two pretests of the questionnaire, one in which participants completed the questionnaire in our presence and raised questions about unclear wording and another in which a professionally drafted questionnaire surveyed participants in the case study [14]. As a result of this preliminary research, we made minor modifications to create the final survey.

We administered the survey by express mails and emails. We first sent a package/email that included a personalized letter, the survey, and a prepaid return envelope. Each participating company was requested to provide data on a recent service innovation project, a typical service innovation project, a successful, and a failure service innovation project. A week later, we mailed a follow-up letter/email to each company to encourage their participation. After that, we sent two follow-up letters/emails and made phone calls to nonresponding companies in an attempt to improve the response rate. The final data in each country is described below.

3.4.1. U.S. Data

We selected our U.S. sample of 1000 companies from the Dun and Bradstreet database. Over three years, we used the data collection procedures described earlier and ultimately obtained complete data

for 477 service innovation projects, with industry representation of 46 from hotel, traveling, and tourism services; 146 from banking, insurances, securities, financial investments, and related activities; 99 from information and semiconductor; 95 from Internet related-services; and 91 from health care services.

3.4.2. Chinese Data

The initial Chinese sample frame comprised all 524 initial public offerings (IPOs) from 2004 to 2014 in the Small and Medium Enterprise (SME) and Growth Enterprise Market (GEM) boards in the Shenzhen Stock Exchange, China. The sample was reduced to 482 IPOs due to missing data.

Using the procedures described earlier, we ultimately collected complete data for 632 service innovation projects from 278 companies. The industry representation of the projects included 40 from hotel, traveling, and tourism services; 217 from banking, insurances, securities, financial investments, and related activities, 120 from information and semiconductor; 91 from Internet-related services; and 164 from health care services.

4. Results

Cronbach's alpha was used to assess construct reliability for big data and BDAC. The alpha coefficients of big data and BDAC, respectively, are 0.891 and 0.803 for the U.S. sample and 0.814 and 0.767 for the Chinese sample. Therefore, our study measures for these two theoretical constructs have high reliability.

We presented the construct median, means, standard deviations, and correlations in Table 1. Exploratory factor analysis was performed to evaluate the construct validity. The factor loadings presented in Table 2 ranged from 0.746 to 0.901 in the U.S. sample and 0.656 to 0.838 in the China sample. These results suggested that the study measures correctly loaded onto the corresponding factor. All retained measures have no double loadings of greater than 0.40. Thus, the constructs possessed convergent validity.

Table 1. Descriptive Statistics.

U.S. Sample: N = 477	Median	Mean	S.D.	Average Sales Growth	Average Gross Margin	Big Data	(Big Data) ²	Big Data Analytics Capability
Average sales growth	30.500	52.400	55.408	NA				
Average gross margin	57.967	77.531	56.152	0.930 ***	NA			
Big data	6.250	5.345	2.697	0.435 ***	0.430 ***	0.891		
(Big data) ²	39.063	35.830	27.057	−0.198 ***	−0.182 ***	−0.395 ***	NA	
Big data analytics capability	6.250	6.044	2.056	0.382 ***	0.392 ***	0.347 ***	−0.169 ***	0.803
China Sample: N = 632	Median	Mean	S.D.	Average Sales Growth	Average Gross Margin	Big Data	(Big Data) ²	Big Data Analytics Capability
1. Average sales growth	26.250	43.894	43.678	NA				
2. Average gross margin	48.600	55.989	28.712	0.888 ***	NA			
3. Big data	5.000	4.685	2.471	0.441 ***	0.457 ***	0.814		
4. (Big data) ²	25.000	28.047	23.785	−0.036	−0.022	−0.013	NA	
5. Big data analytics capability	6.500	6.254	2.085	0.427 ***	0.451 ***	0.405 ***	−0.186 ***	0.767

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The Cronbach Alpha for each scale is on the diagonal in italics, the intercorrelations among the variables are on the off diagonal.

Table 2. Factor loadings from exploratory factor analyses.

	U.S. Sample (N = 477)		China Sample (N = 632)	
	Big Data	BDAC	Big Data	BDAC
Big data 1	0.857	0.133	0.838	0.126
Big data 2	0.894	0.079	0.754	0.189
Big data 3	0.759	0.327	0.777	0.143
Big data 4	0.901	0.101	0.765	0.173
BDAC 1	0.098	0.785	0.227	0.787
BDAC 2	0.099	0.824	0.016	0.782
BDAC 4	0.157	0.768	0.303	0.656
BDAC 5	0.176	0.746	0.139	0.762

Note: BDAC = Big data analytics capability. Bold numbers indicate items that load highly for the corresponding factor.

Our theoretical model in Figure 1 includes following two equations:

$$\begin{aligned} \text{Average sale growth over the first three years} = & \alpha + \beta_1 \times (\text{Big data}) + \beta_2 \times (\text{Big data})^2 + \\ & \beta_3 \times (\text{BDAC}) + \beta_4 \times (\text{Big data}) \times (\text{BDAC}) + \beta_5 \times (\text{Big data})^2 \times (\text{BDAC}) + \beta_6 \times \text{Industry1} + \\ & \beta_7 \times \text{Industry2} + \beta_8 \times \text{Industry3} + \beta_9 \times \text{Industry4} + \varepsilon. \end{aligned} \quad (1)$$

$$\begin{aligned} \text{Average gross-margin over the first three years} = & \alpha + \beta_1 \times (\text{Big data}) + \beta_2 \times (\text{Big data})^2 + \\ & \beta_3 \times (\text{BDAC}) + \beta_4 \times (\text{Big data}) \times (\text{BDAC}) + \beta_5 \times (\text{Big data})^2 \times (\text{BDAC}) + \beta_6 \times \text{Industry1} + \\ & \beta_7 \times \text{Industry2} + \beta_8 \times \text{Industry3} + \beta_9 \times \text{Industry4} + \varepsilon. \end{aligned} \quad (2)$$

To test the hypotheses, the above two equations are estimated for each country separately using Proc Reg in SAS 9.4 program. We performed multiple ordinary least square (OLS) regression analyses for each country using project performance measures: project's three-year average sales growth rate and gross margin data. We presented results for the three-year average growth rate in Table 3 and the three-year average gross margin in Table 4. We also estimated the same equations using data for each year and found the major conclusions do not change. Therefore, we reported only the results of the average sales growth and the average gross margin.

Table 3. Results of regression analyses: average sales growth as dependent variable.

	Dependent Variable: Three-year Average Sales Growth					
	U.S. Sample (N = 477)			China Sample (N = 632)		
	Parameter Estimate	Standard Error	Standardized Estimate	Parameter Estimate	Standard Error	Standardized Estimate
Intercept	56.046 ***	5.759	0.000	44.062 ***	3.336	0.000
Big data	5.853 ***	0.970	0.285	6.473 ***	0.694	0.366
(Big data) ²	−0.700 *	0.418	−0.074	−0.087	0.284	−0.011
BDAC	2.561	1.763	0.095	8.607 ***	1.065	0.411
(Big data) × BDAC	2.860 ***	0.535	0.244	0.419	0.308	0.052
(Big data) ² × BDAC	0.733 ***	0.237	0.222	−0.353 ***	0.107	−0.175
Industry 1	−2.297	8.517	−0.012	6.363	6.537	0.036
Industry 2	−3.133	6.276	−0.026	0.485	3.850	0.005
Industry 3	−4.241	6.828	−0.031	−6.847	4.459	−0.062
Industry 4	−10.438	6.909	−0.075	2.903	4.868	0.023
F-value		21.980 ***			28.420 ***	
R ²		0.298			0.291	
Adjusted R ²		0.284			0.281	

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. BDAC = Big data analytics capability; Industry 1 = Hotel, Traveling and Tourism Services; Industry 2 = Banking, Insurances, Securities, Financial Investments and Related Activities; Industry 3 = Information and Semiconductor; Industry 4 = Internet Related Services. Industry 5 = Health Care Services (the estimate for industry 5 is the intercept term).

Table 4. Results of regression analyses: average gross margin as dependent variable.

	Dependent Variable: Three-year Average Gross Margin					
	U.S. Sample (N = 477)			China Sample (N = 632)		
	Parameter Estimate	Standard Error	Standardized Estimate	Parameter Estimate	Standard Error	Standardized Estimate
Intercept	82.349 ***	5.802	0.000	53.837 ***	2.162	0.000
Big data	5.775 ***	0.977	0.277	4.194 ***	0.450	0.361
(Big data) ²	−0.522	0.421	−0.054	0.111	0.184	0.022
BDAC	2.482	1.776	0.091	5.607 ***	0.690	0.407
Big data × BDAC	2.940 ***	0.539	0.248	0.153	0.200	0.029
(Big data) ² × BDAC	0.831 ***	0.238	0.248	−0.171 **	0.069	−0.129
Industry 1	−5.101	8.581	−0.027	2.307	4.237	0.020
Industry 2	−4.867	6.323	−0.040	3.852	2.496	0.064
Industry 3	−7.397	6.879	−0.053	−2.981	2.890	−0.041
Industry 4	−15.922 **	6.961	−0.113	3.954	3.155	0.048
F-value		22.850 ***			31.200 ***	
R ²		0.306			0.311	
Adjusted R ²		0.292			0.301	

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. BDAC = Big data analytics capability. Industry 1 = Hotel, Traveling and Tourism Services; Industry 2 = Banking, Insurances, Securities, Financial Investments and Related Activities; Industry 3 = Information and Semiconductor; Industry 4 = Internet Related Services. Industry 5 = Health Care Services (the estimate for industry 5 is the intercept term).

4.1. Research Hypothesis 1: Direct Effects of Big Data and Innovation Performance

Hypothesis 1 predicts an inverted U-shaped relationship between big data and sustainability of innovation and organizational development. The results vary according to the performance measure being considered. For the average sales growth, Table 3 shows that for the U.S. sample, big data had a significant positive relationship ($b = 5.853$, $p < 0.01$) whereas (big data)² had a significant negative relationship ($b = -0.700$, $p < 0.10$). These results indicate that big data had an inverted U-shaped relationship with sustainability of innovation and organizational development as measured by the average sales growth in the U.S. sample, in support of Hypothesis 1.

For the Chinese sample, the relationship to big data was again significantly positive ($b = 6.473$, $p < 0.01$) whereas (big data)² had a negative but not significant relationship ($b = -0.087$, $p > 0.10$). In contrast to the U.S. sample, big data had a positive effect on the average sales growth, but did not have an inverted U-shaped relationship with the average sales growth. Thus, H1 was not supported by the Chinese sample.

For the three-year average gross margin, Table 4 shows that for the U.S. sample, big data had a significant positive relationship ($b = 5.775$, $p < 0.01$) whereas (big data)² had a negative but not significant relationship ($b = -0.522$, $p > 0.10$). For the Chinese sample, big data again had a significant positive relationship ($b = 4.194$, $p < 0.01$) and the relationship with (big data)² was also positive but not significant ($b = 0.111$, $p > 0.10$). Thus, counter to H1, these results indicate that for both U.S. and China, big data had a positive linear relationship with sustainability of innovation and organizational development as measured by the average gross margin.

4.2. Research Hypothesis 2: The Moderating Effects of BDAC

To test Hypothesis 2, we needed to evaluate the coefficient estimates of (big data) × BDAC and (big data)² × BDAC. Results in Tables 3 and 4 indicate that for the U.S. sample, both (big data) × BDAC and (big data)² × BDAC had a significant positive relationship with the average sales growth ($b = 2.860$, $p < 0.01$; $b = 0.733$, $p < 0.01$) and the average gross margin ($b = 2.940$, $p < 0.01$; $b = 0.831$, $p < 0.01$). For the Chinese sample, (big data) × BDAC had a positive, but not significant, relationship with both the average sales growth ($b = 0.419$, $p > 0.10$) and the average gross margin ($b = 0.153$, $p > 0.10$) whereas (big data)² × BDAC had a significant negative relationships with both the average sales growth ($b = -0.353$, $p < 0.01$) and the average gross margin ($b = -0.171$, $p < 0.05$). These results

thus indicate that in the U.S., with the interaction of BDAC, big data had a U-shaped relationship with sustainability of innovation and organizational development. Yet, for the Chinese sample, there was an inverted U-shaped relationship.

To further illustrate these interesting relationships, we presented Figures 2 and 3 to show the specific relationships. Figure 2a shows the relationship between big data and the average sales growth at different levels of BDAC for the U.S. service innovation projects. As we can see from Figure 2a, when BDAC was 0, big data had an inverted U-shaped effect on the average sales growth. When BDAC was 1 or higher, there was a half U-shaped relationship between big data and the average sales growth; further increases in big data increased the average sales growth. These findings provide empirical supports for Hypotheses 1 and 2.

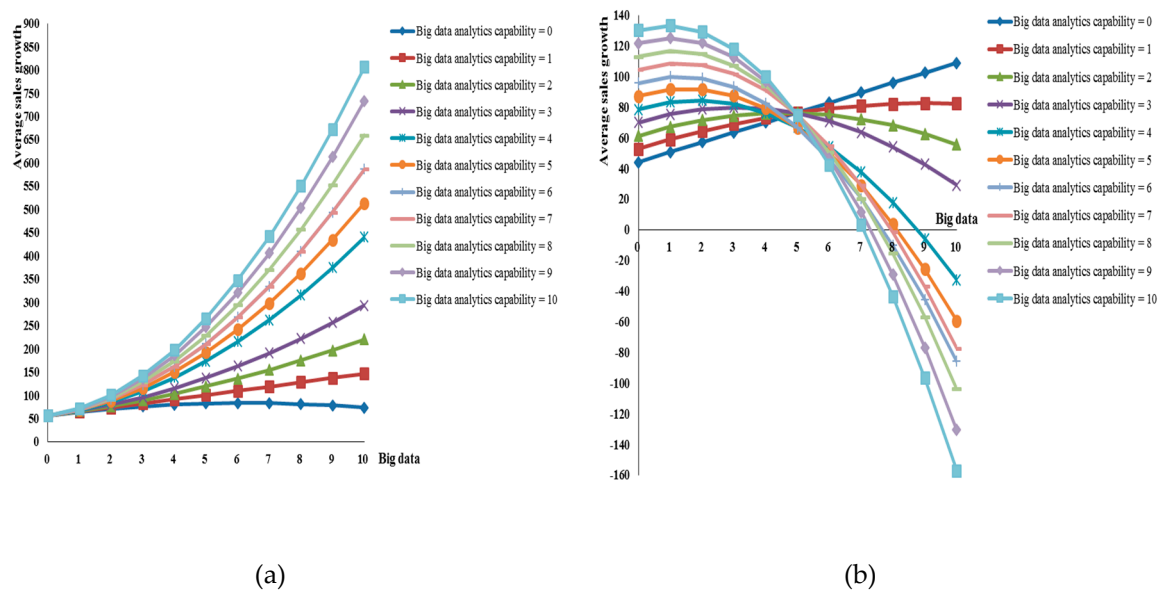


Figure 2. Effects of big data on average sale growth at different levels of big data analytics capability. (a) US and (b) China.

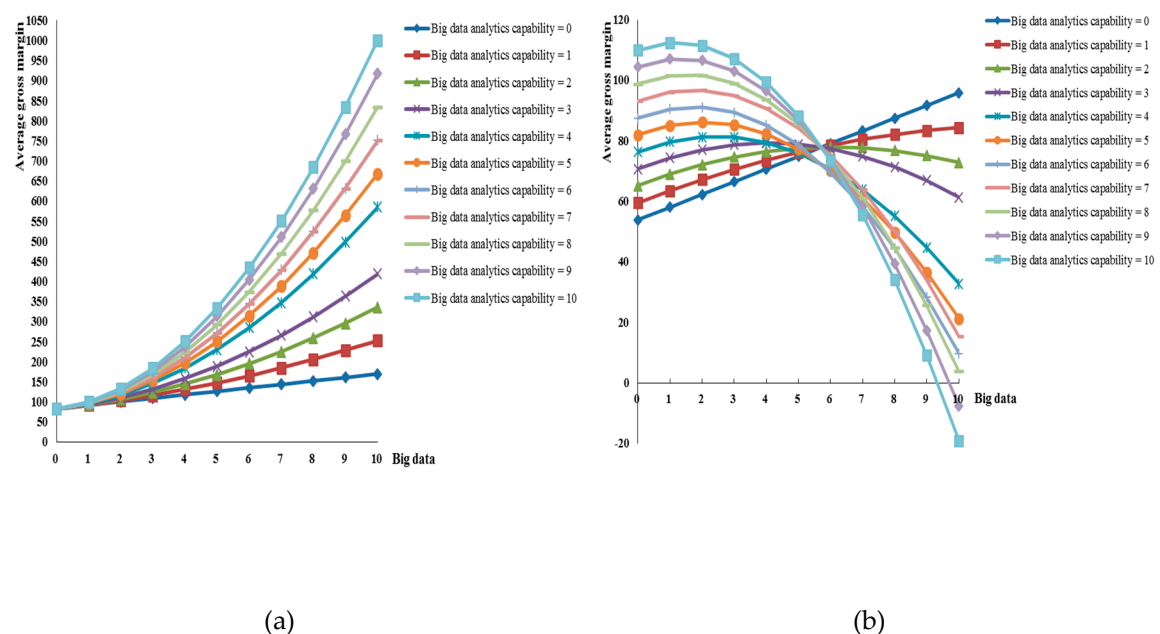


Figure 3. Effects of big data on average gross margin at different levels of big data analytics capability. (a) US and (b) China.

For Chinese service innovation projects, Figure 2b shows that when BDAC was 0, big data had a positive linear relationship with the average sales growth. When BDAC was 1 or higher, this relationship between big data and the average sales growth became an inverted U shape. Before a certain point (big data level of 5), BDAC strengthened the effect of big data on the average sales growth; after this point, BDAC exerted a negative moderating effect. These results provided partial support for Hypothesis 2.

Figure 3a shows the relationship between big data and average gross margin at different levels of BDAC for U.S. innovation projects. When BDAC was 0, big data had a positive linear relationship with gross margin, that is, increases in big data positively affected gross margin. When BDAC was 1 or higher, there was a U-shaped relationship between big data and gross margin; further increases in big data positively affected sale growth. These findings fully support Hypothesis 2.

Figure 3b shows that when BDAC was 0, big data had a positive linear relationship with gross margin. When BDAC was 1 or higher, the relationship between big data and gross margin became an inverted U shape. Before a certain point (big data level was 6), BDAC strengthened the effect of big data on gross margin: whereas after this point, BDAC exerted a negative moderating effect. These results provide partial support for Hypothesis 2.

5. Discussion

Building on the resource-based view and organizational inertia theory, we presented a model that specified how big data could be leveraged by innovation teams to achieve superior performance. Using samples from the United States and China, we empirically tested the model and found two interesting and surprising results.

First, our findings provided a more nuanced understanding of the impact of big data on sustainability of innovation and organizational development. Previous research has highlighted the role of big data in attaining sustainable competitive advantage. Consistent with this logic, we found that big data enhanced sales growth and gross margin. More interestingly, for U.S. service innovation projects, we found that big data had an inverted U-shaped relationship with sales growth: a moderate level of big data related to the highest degree of sales growth whereas a high level of big data actually inhibited sales growth (see Figure 2a). An explanation for this might be that resource and routine inertia [16] might cause the innovation teams to depend too much on big data resources to gain information and knowledge, thereby overlooking other resource allocation patterns necessary to increase sales growth and increase innovation successes. These findings contribute to the extant literature by demonstrating and explaining the limited benefit of increasing investment in big data.

Second, we extended the existing literature by presenting and confirming empirically that BDAC strengthened the positive impact of big data on the sustainability of innovation and organizational development. As shown in Figures 2a and 3a for the U.S. sample, increased big data investment was a more important precursor to innovation projects sales growth and gross margin if there was a moderate to high level of BDAC. Thus, innovation teams would gain more from their investment in big data if they also devote resources to building their BDAC. As shown in Figures 2b and 3b, the findings for the Chinese sample were more complicated: BDAC strengthened the effects of big data on sales growth and gross margin only before a certain point (big data level of 5 or 6), but beyond this point, BDAC exerted a negative moderating effect.

6. Managerial Insights

Our findings provided some practical insights for innovation project leaders regarding big data resources and BDAC. It is suggested that project leaders in both U.S and China can increase service innovation project performance by ensuring that the project teams have adequate big data resources on customer needs, user behavior, competitive intelligence (e.g., marketing, products, services, etc.), and technology development intelligence.

Although these big data resources can increase sales growth and gross margin, managers in U.S. service innovation teams must be aware of the limitations of their big data resources and understand that too much investment in big data may actually hurt sales growth of the service innovation projects. The innovation teams also need to build their BDAC to enable them to harness the potential of their big data resources.

For U.S. service innovation teams, devoting resources to BDAC is preferable to additional investment in big data. The stronger the BADC is, the greater effect big data can have in improving sales growth and gross margin. Using the variables in our data, U.S. innovation teams should do the following to achieve a high level of BDAC:

- (1) Use more advanced tools (analytics and algorithms) to extract values from big data;
- (2) Develop the capability to discover relationships and dependencies from big data;
- (3) Develop the capability to perform predictions of outcomes and behaviors from big data;
- (4) Develop the capability to discover new correlations from big data to spot market demand trends and predict user behavior; and
- (5) Employ and cultivate big data analytics staff who have the right skills to accomplish their jobs successfully.

Managers of Chinese service innovation projects should be aware that developing BDAC might have its downsides. When investment in big data is low, promoting BDAC can increase sales growth and gross margin. However, after a certain point, allocating further resources to BDAC may inhibit sustainability of innovation and organizational development. Chinese managers may also use the suggested tactics to increase big data analytics capability.

Our findings thus provided insights to U.S. and Chinese innovation teams on the subtle ways in which big data could lead to superior performance and might help them to maximize their use of big data.

7. Limitations and Future Research

Although this study added to our understanding of how big data and big data analytics capabilities affected sustainable innovation performance, there were at least two limitations that also offer future research opportunities. First, the adjusted R^2 of our models were below 0.33 because we focused on only the impact of big data and big data analytics capabilities on sustainable innovation performance. Past research suggests that other capabilities such as technology, marketing, market-linking, and information technology capabilities also affect innovation performance. Collecting new data and including these capabilities in the proposed model should increase R^2 and offer additional interesting findings. This constitutes an interesting avenue for future research. Second, there are some new and emerging capabilities such as AI capabilities and machine learning capabilities. It will be very interesting to conduct future research to study how these new capabilities together with big data and big data analytics capabilities affect sustainable innovation performance.

Author Contributions: All authors share equally for the first authorship and contribute equally to the development of this research. Conceptualization, S.H., H.Z., and M.S.; Data curation, H.Z. and M.S.; Methodology, M.S. and H.Z.; Writing—original draft, S.H. and H.Z., and M.S.; Writing—review & editing, S.H., H.Z., and M.S.

Funding: This research was funded by the China Natural Science Foundation, grant number 71772054; Natural Science Foundation of Shaanxi Province of China, grant number 2018JQ7003; and Shaanxi Province of China, Department of Education Scientific Research Plan Projects, grant number 18JK0367.

Acknowledgments: The authors thank assistant editor, Letty Zhang, and the two anonymous reviewers for their useful suggestions which improve the quality of this research.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Study Measures and Sources

Big Data (New Scale Developed Based on Filed Research):

Our big data on the following aspects are more than adequate for this development project:

- (1) Customer needs (0 = not adequate for this development project; 10 = much more than adequate for this development project).
- (2) User behavior (0 = not adequate for this development project; 10 = much more than adequate for this development project).
- (3) Competitive intelligence (e.g., marketing, products, services, etc.; 0 = not adequate for this development project; 10 = much more than adequate for this development project).
- (4) Technology development intelligence (0 = not adequate for this development project; 10 = much more than adequate for this development project).
- (5) Product usages (0 = not adequate for this development project; 10 = much more than adequate for this development project).

Big Data Analytics Capability:

The following is a set of possible big data analytics capabilities. Please evaluate how well or poorly you believe that your company possesses the specific capability.

- (1) We have advanced tools (analytics and algorithms) to extract values of the big data. (0 = no capability; 10 = very high level of capability; adapted from Dubey, Gunasekaran, and Childe [29]; Gupta and George [10]).
- (2) Our capability to discover relationships and dependencies from the big data is: (0 = no capability; 10 = very high level of capability; new scale developed based on filed research).
- (3) Our capability to perform predictions of outcomes and behaviors from the big data is: (0 = no capability; 10 = very high level of capability; adapted from Gupta and George [10]).
- (4) Our capability to discover new correlations from the big data to spot market demand trends and predict user behavior is: (0 = no capability; 10 = very high level of capability; adapted from Akter, Wamba, Gunasekaran, Dubey, and Childe [5]; Wamba, Gunasekaran, Akter, Ren, and Childe [9]).
- (5) Our big data analytics staff has the right skills to accomplish their jobs successfully. (0 = none; 10 = very high level of capability; adopted from Gupta and George [10]).

Average Sales growth rate (%; adopted from Hu, et al. [30])

$$= (\text{second year sales growth rate} + \text{third year sales growth rate})/2,$$

where

$$\text{second year sales growth rate} = (\text{second year sales} - \text{first year sales})/(\text{first year sales}) \times 100;$$

$$\text{third year sales growth rate} = (\text{third year sales} - \text{second year sales})/(\text{second year sales}) \times 100.$$

Average gross margin (in %; adopted from Zhao, Song and Storm [31])

$$= (\text{first year gross margin} + \text{second year gross margin} + \text{third year gross margin})/3,$$

where

$$\text{gross margin} = (\text{total sales} - \text{total variable costs})/\text{total sales}.$$

References

1. Hu, F.; Liu, W.; Tsai, S.B.; Gao, J.; Bin, N.; Chen, Q. An empirical study on visualizing the intellectual structure and hotspots of big data research from a sustainable perspective. *Sustainability* **2018**, *10*, 667. [\[CrossRef\]](#)
2. Lytras, M.D.; Visvizi, A. Big data and their social impact: Preliminary study. *Sustainability* **2019**, *11*, 5067. [\[CrossRef\]](#)
3. Oncioiu, I.; Bunget, O.C.; Türkeş, M.C.; Căpuşneanu, S.; Topor, D.I.; Tamaş, A.S.; Rakos, I.S.; Hint, M.S. The impact of big data analytics on company performance in supply chain management. *Sustainability* **2019**, *11*, 4864. [\[CrossRef\]](#)
4. Xu, L.; Gao, R.; Xie, Y.; Du, P. To be or not to be? Big data business investment decision-making in the supply chain. *Sustainability* **2019**, *11*, 2298. [\[CrossRef\]](#)
5. Akter, S.; Wamba, S.F.; Gunasekaran, A.; Dubey, R.; Childe, S.J. How to improve firm performance using big data analytics capability and business strategy alignment? *Int. J. Prod. Econ.* **2016**, *182*, 113–131. [\[CrossRef\]](#)
6. Mikalef, P.; Pappas, I.O.; Krogstie, J.; Giannakos, M. Big data analytics capabilities: A systematic literature review and research agenda. *Inf. Syst. E-Bus. Manag.* **2018**, *16*, 547–578. [\[CrossRef\]](#)
7. Syam, N.; Sharma, A. Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Ind. Market Manag.* **2018**, *69*, 135–146. [\[CrossRef\]](#)
8. Ferraris, A.; Mazzoleni, A.; Devalle, A.; Couturier, J. Big data analytics capabilities and knowledge management: Impact on firm performance. *Manag. Decis.* **2019**, *57*, 1923–1936. [\[CrossRef\]](#)
9. Wamba, S.F.; Gunasekaran, A.; Akter, S.; Ren, S.J.; Dubey, R.; Childe, S.J. Big data analytics and firm performance: Effect of dynamic capabilities. *J. Bus. Res.* **2016**, *70*, 356–365. [\[CrossRef\]](#)
10. Gupta, M.; George, J.F. Toward the development of a big data analytics capability. *Inf. Manag. Amst.* **2016**, *53*, 1049–1064. [\[CrossRef\]](#)
11. Jun, S. Bayesian structural time series and regression modeling for sustainable technology management. *Sustainability* **2019**, *11*, 4945. [\[CrossRef\]](#)
12. McAfee, A.; Brynjolfsson, E.; Davenport, T.H.; Patil, D.J.; Barton, D. Big data: The management revolution. *Harv. Bus. Rev.* **2012**, *90*, 61–67.
13. Douglas, S.P.; Craig, C.S. *International Marketing Research*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1983.
14. Song, X.M.; Parry, M.E. What separates Japanese new product winners from losers. *J. Prod. Innov. Manag.* **1996**, *13*, 422–439. [\[CrossRef\]](#)
15. Hannan, M.T.; Freeman, J.H. Structural inertia and organizational change. *Am. Sociol. Rev.* **1984**, *49*, 149–164. [\[CrossRef\]](#)
16. Gilbert, C.G. Unbundling the structure of inertia: Resource versus routine rigidity. *Acad. Manag. J.* **2005**, *8*, 741–763. [\[CrossRef\]](#)
17. Beyer, M.A.; Laney, D. The importance of ‘big data’: A definition. *Stamford CT Gart.* **2012**, *21*, 2014–2018.
18. Barney, J.B. Firm resources and sustained competitive advantage. *J. Manag.* **1991**, *17*, 99–120. [\[CrossRef\]](#)
19. Sirmon, D.G.; Gove, S.; Hitt, M.A. Resource management in dyadic competitive rivalry: The effects of resource bundling and deployment. *Acad. Manag. J.* **2008**, *51*, 918–935. [\[CrossRef\]](#)
20. Olszak, C.M. Towards an understanding business intelligence. A dynamic capability-based framework for business intelligence. In Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, Warsaw, Poland, 7–10 September 2014; Volume 2, pp. 1103–1110.
21. White, C. Using big data for smarter decision making. *BI Res.* **2011**, 1–11. [\[CrossRef\]](#)
22. Song, L.Z.; Song, M.; Di Benedetto, C.A. Resources, supplier investment, product launch advantages, and first product performance. *J. Oper. Manag.* **2011**, *29*, 86–104. [\[CrossRef\]](#)
23. Garmaki, M.; Boughzala, I.; Wamba, S.F. The effect of BDAC on firm performance. In Proceedings of the 20th Pacific Asia conference on information systems (PACIS), Brisbane, Australia, 7–11 July 2016.
24. Priem, R.L.; Butler, J.E. Is the resource-based view a useful perspective for strategic management research? *Acad. Manag. Rev.* **2001**, *26*, 22–40.
25. Barney, J.B.; Arian, A.M. The resource-based view: Origins and implications. In *Handbook of Strategic Management*; Wiley Online Library: Hoboken, NJ, USA, 2001; pp. 124–188.
26. Grant, R.M. The resource-based theory of competitive advantage: Implications for strategy formulation. *Calif. Manag. Rev.* **1991**, *33*, 114–135. [\[CrossRef\]](#)

27. Churchill, G.A. A paradigm for developing better measures of marketing construct. *J. Mark. Res.* **1979**, *6*, 64–73. [[CrossRef](#)]
28. Song, X.M.; Parry, M.E. A cross-national comparative study of new product development processes: Japan and the United States. *J. Mark.* **1997**, *61*, 1–18. [[CrossRef](#)]
29. Dubey, A.; Gunasekaran, A.; Childe, S.J. Big data analytics capability in supply chain agility: The moderating effect of organizational flexibility. *Manag. Decis.* **2018**. [[CrossRef](#)]
30. Hu, C.; Zhang, H.; Song, M.; Liang, D. Past Performance, Organizational Aspiration, and Organizational Performance: The Moderating Effect of Environmental Jolts. *Sustainability* **2019**, *11*, 4217. [[CrossRef](#)]
31. Zhao, Y.L.; Song, M.; Storm, G.L. Founding Team Capabilities and New Venture Performance: The Mediating Role of Strategic Positional Advantages. *Entrep. Theory Pract.* **2013**, *37*, 789–814. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).