DOI 10.3868/s070-008-023-0032-7

Big Data Capability, Knowledge Dynamic Capability, and Business Model Innovation: The Moderating Effect of Innovation Legitimacy*

> ZHANG Jichang[a], LONG Jing[a], CHEN Feng[b]

Abstract Based on dynamic capability theory and legitimacy theory, a theoretical model is constructed to examine how big data capability, through the mediation of knowledge dynamic capability, drive business model innovation under the moderation effect of innovation legitimacy. The analysis is conducted using regression analysis and fuzzy set qualitative comparative analysis (fsQCA) on survey data from 302 enterprises that have already implemented big data application practices. The study finds the following four conclusions. (1) Big data capability has a significant positive impact on business model innovation. (2) Dynamic knowledge capability partially mediates the relationship between big data capability and business model innovation. (3) Innovation legitimacy positively influences business model innovation and positively moderates the relationship between big data capability and business model innovation. (4) Through further qualitative comparative analysis, two causal paths that influence business model innovation are identified.

Keywords big data capability, dynamic knowledge capability, business model innovation, innovation legitimacy, fuzzy set qualitative comparative analysis (fsQCA)

1 Introduction

With the development of the internet economy and the advent of the big data era, data have become the fourth important asset, alongside land, capital, and labor. Big data is regarded as the "future new oil" or "diamond mine" by many countries. The Chinese government also attaches great importance to the role of big data in social change, economic development, and industrial upgrading. The *Report of the Fifth Plenary Session of the 18th Central Committee of the Communist Party of China* explicitly emphasizes the implementation of a national big data strategy, marking the formal elevation of such a strategy to a national level. Subsequently, the Chinese government has formulated and promulgated numerous policy documents regarding the development of big data, aiming to accelerate China's digitalization

* This paper is translated from *Economy and Management* 经济与管理, 2022, (5), 19–28.
a Business School, Nanjing University, Nanjing 210093, China
b School of Economics and Management, South China Normal University, Guangzhou 510006, China

process and build China into a strong country on data. These policies and measures greatly promote the implementation of the big data strategy, bringing unprecedented opportunities for upgrading of traditional industry and for business model innovation based on big data. Scholars have pointed out that business model innovation is deeply influenced by digital technology and network technology. However, many enterprises, especially small and medium-sized enterprises and traditional enterprises, still face many challenges in business model innovation in the context of big data. They lack theoretical guidance, and there is an urgent need to study the impact of the "big data era" on business model innovation (Wei et al., 2014). This is a research question of great significance (Li & Xia, 2013).

In recent years, scholars have made numerous valuable explorations regarding the relationship between big data and business model innovation. Early studies are primarily qualitative analyses. Li and Xia (2013) discuss the impact of big data resources and technologies on business models from a theoretical and speculative perspective. Hartmann et al. (2016) and Yi and Xu (2018) propose comprehensive analytical frameworks for big data-driven business model innovation. Zeng (2019) distinguish different types of big-data-driven business model innovation. In addition, there is much controversy in the academic community regarding the impact of big data application practice on enterprise innovation. Most scholars acknowledge the positive role of big data application practice in enterprise innovation, but a small number of scholars have questioned this positive impact. For example, Yang et al. (2017) argue that big data resources themselves cannot directly drive enterprise innovation. They suggest that it is necessary to use new big data processing technologies and tools to effectively mine and apply massive amounts of data, otherwise, big data will be useless. Ghasemaghaei and Calic (2020) find through empirical research that the macro-volume characteristics of big data do not have a significant positive impact on enterprise innovation. As research has deepened, a few scholars have moved beyond the perspective of resources and technologies and have extracted the concept of big data capability, conducting related empirical research. Xie et al. (2018) explore the direct relationship between big data capability and business model innovation, and their work is one of the earlier quantitative studies on the relationship between the two, domestically and internationally. Scholars whose work is based on the capability perspective have suggested that the value of big data itself is limited. Only by elevating big data resources to the level of capability and fully tapping into the business value and experiential knowledge of big data, can innovation activities of enterprises be more meaningful (Tang et al., 2020). This provides a research opportunity, to uncover the mechanism of the role of big data capability in business model innovation.

From the current research perspective, the exploration of the relationship between big data capability and business model innovation is just beginning, and future research can further investigate at least four aspects. Firstly, in terms of the mechanism, extant literature on the impact

mechanism of big data capability driving business model innovation is still in the preliminary stage of understanding, and there is a lack of empirical studies in this area. This obviously hinders the ability of enterprises to utilize big data for innovation activities (Zeng, 2019). Secondly, in terms of the action boundaries, there is still a lack of research on the application scenarios of big data capability, especially on the legitimacy issues faced by enterprises in conducting business model innovation through the collection and utilization of consumer behavioral data (Ding & Liu, 2018; Pan & Xie, 2021). However, little attention has been paid to this issue, leading to vague conclusions in research, and even to controversies, regarding the effectiveness of big data (Zhao et al., 2019). Thirdly, in terms of the evolution of big data capability, early studies regarded big data capability as dynamic capability, but recent research has noted that big data capability is a type of low-order operational capability that promotes the formation of high-order dynamic capability (Mikalef et al., 2019). However, there is still a lack of relevant empirical evidence. Fourthly, in terms of research methods, extant studies are mostly theoretical speculation. Qualitative research, with large-sample empirical data and mixed-method research are even rarer.

Big data capability refers to the ability of enterprises to deploy, manage, and apply internal and external big data resources and technologies. According to the perspective of dynamic capability theory, on the one hand, dynamic capability has hierarchical and evolutionary characteristics, and big data capability is a low-order operational capability that does not yet possess the traits of dynamic capability, serving only as the foundation for the formation of other higher-level dynamic capability (Mikalef et al., 2019). On the other hand, the key to enterprise dynamic capability lies in the integration and reconfiguration of resources, with knowledge resources being the core resources for the formation and maintenance of sustainable competitive advantages (Teece et al., 1997). If a company only possesses the ability to collect big data without the ability to extract information and process it into knowledge, it will be difficult to drive and lead business model innovation. In other words, enterprises need to elevate their operation capability, based on big data resources, to have knowledge-based dynamic capability. Therefore, knowledge dynamic capability provides an entry point for exploring how big data capability drive business model innovation. Additionally, many business model innovations triggered by big data face legitimacy issues (Ding & Liu, 2018). In recent years, there have been widespread social concern over cases such as excessive collection of user information, misuse of user data, and swindle acquaintances based on big data (Li et al., 2021). While utilizing big data capability for business model innovation, enterprises also need to comply with market validation and legal constraints (Pan & Xie, 2021). The legitimacy of innovation is an important boundary condition.

This paper comprehensively applies dynamic capability theory and legitimacy

theory to explore the intermediate mechanisms and boundary conditions of how big data capability influences business model innovation. It constructs a theoretical model with knowledge dynamic capability as the mediating variable and with innovation legitimacy as the moderating variable. The study utilizes regression analysis and fuzzy set qualitative comparative analysis (fsQCA) to test the model using survey data from 302 enterprises that have introduced big data applications.

The possible contributions of this paper are as follows. Firstly, it expands the research on the antecedents of enterprise business model innovation. In addition to finding a significant positive correlation between big data capability and business model innovation, this work also discovers the promoting effect of knowledge dynamic capability and innovation legitimacy on business model innovation. The synergistic linkage between different element combinations under the configurational perspective leads to similar outcomes in terms of business model innovation. Secondly, it finds that knowledge dynamic capability plays a mediating role between big data capability and business model innovation. It unveils the "black box" of the transformation process from big data capability to business model innovation from a dynamic capability perspective, responding to Teece's (2018) exploration of the relationship between dynamic capability and business models. It also focuses on the current booming digital economy in China and contribute to the further maturity and development of dynamic capability theory in the context of digitization, responding positively to the calls from researchers such as Jiao et al. (2021) and Zhang and Long (2022). Thirdly, extant research has already paid attention to the legitimacy issues of big data (Cheng, 2018; Li & Xia, 2013; Li et al., 2021), but fewer studies have examined innovation legitimacy as an important contextual variable in the relationship between big data capability and business model innovation. Based on legitimacy theory, this paper examines the moderation effect of innovation legitimacy on the relationship between big data capability and business model innovation, providing a new theoretical perspective for subsequent research on data-driven innovation. Fourthly, most extant research remains at the stage of descriptive case studies and qualitative analysis, lacking empirical research. This paper provides new large-scale empirical evidence on the question of "how big data capability influence business model innovation." It further employs the fsQCA method to explore different paths leading to high business model innovation from a configurational perspective, providing new research ideas for exploring this topic from a methodological standpoint.

2 Literature Review and Research Hypotheses

2.1 | Big Data Capability and Business Model Innovation

Cheng and Li (2014) first propose the term big data capability and define it as the ability of enterprises to collect, store, mine, and

utilize big data. Xie et al. (2018) then develop a measurement scale for big data capability. Subsequently, Lin and Kunnathur (2019), based on dynamic capability theory, propose the term big data capability and developed a measurement scale, defining it as the ability of enterprises to identify, collect, store, and analyze various rapidly flowing data to achieve strategic and operational goals. In the era of emerging digital technologies and the booming development of big data, the relationship between enterprises and big data has become increasingly close, with big data as a new technological tool and resource that continuously drives business model innovation. The business model is used to describe the logic and architecture through which enterprises achieve market transactions (Amit & Zott, 2012; Zott & Amit, 2007), and its process can be summarized as the creation, delivery, and capture of value. Each step in this process can serve as a trigger for business model innovation (Li & Xia, 2013). The impact of big data capability on business model innovation can be manifested in the following three ways.

Firstly, big data resources provide the foundation for enterprises to find new value propositions for their business models (Wamba et al., 2017), and hold the potential for business model transformation. When the process of product and service output is deconstructed and digitized, enterprises can discover new value propositions and find new business opportunities through value associations.

Secondly, enterprises can innovate the value creation path in their business models by integrating big data resources. Based on big data infrastructure and technology, enterprises can establish solid hardware and software foundations for the exchange and communication of information among various business departments, enabling the absorption of external big data resources. The data from all transaction parties permeate the entire business process, driving enterprises to redesign and reconstruct their production and service processes. As a result, traditional value creation processes are upgraded or even disrupted, and enterprises generate new ways of creating value (Li & Xia, 2013; Liu et al., 2020).

Thirdly, enterprises can change the value delivery method in their business models by utilizing the predictive capability of big data. Enterprises are decision-making systems, and the predictive capability of big data can provide strong support for precise decision-making. Traditionally, enterprises deliver value to consumers in a standardized manner. However, enterprises with big data predictive capability can effectively capture and understand consumers' latent and personalized needs by collecting, organizing, and analyzing massive consumer trajectory data. They can achieve a one-to-one value delivery when delivering products or services (Akter et al., 2016; Zeng, 2019).

Based on the above, the following hypothesis is proposed.

H1: Big data capability has a positive impact on business model innovation.

2.2 | The Mediating Role of Knowledge Dynamic Capability

Dynamic capability refers to the ability

of enterprises to integrate, build, and reconfigure internal and external resources to adapt to rapidly changing external environments (Teece et al., 1997). Therefore, resources are the core element influencing the formation and evolution of dynamic capability. Among various heterogeneous resources, scholars generally believe that the most valuable, scarce, non-substitutable, and difficult to imitate resource for enterprises is their knowledge resource. Various types of knowledge activities within enterprises are the sources that drive the formation of dynamic capability. In this context, Wang et al. (2007) focus on knowledge resources and propose the concept of knowledge dynamic capability, and define it as the abilities of enterprises to dynamically integrate, adjust, and apply their knowledge base to gain competitive advantages. Subsequent scholars have expanded and interpreted the dynamic capability theory from different perspectives, pointing out that dynamic capability have hierarchical and evolutionary characteristics (Winter, 2003; Zhang et al., 2022). Big data capability, as a low-order operational capability, can create competitive advantages for enterprises by driving the formation of higher-level dynamic capability (Chen et al., 2015).

Based on these viewpoints, this paper suggests that big data capability is a low-order operational capability that can be imitated by other enterprises, while knowledge dynamic capability is a higher-order heterogeneous capability developed on the foundation of big data capability and is difficult to be imitated by other enterprises. Firstly, big data capability enhances the ability of enterprises to acquire knowledge. In the knowledge acquisition stage, big data capability significantly improves the access to valid information and knowledge in the current digital context, ensuring the timeliness and accuracy of knowledge reserves, and facilitating the improvement of internal knowledge structures.

Secondly, big data capability enhances the abilities of enterprises to integrate new and old knowledge. In the knowledge integration stage, big data capability possesses attributes of integration, utilization, and sharing. It efficiently integrates the acquired knowledge with existing knowledge reserves, enabling internal personnel to better learn and master content that aligns with the development of the enterprise, thereby assisting in enhancing the overall knowledge integration capability.

Finally, big data capability improves the abilities of enterprises to create knowledge. In the knowledge reconfiguration stage, big data capability enables the reanalysis of integrated knowledge, effectively classifying knowledge and creating knowledge that is more in line with current trends and market development needs. Chen et al. (2015) point out that big data capability helps enterprises scan and interpret the external environment, providing them with regular knowledge and enabling them to predict the environment in advance, thus effectively stimulating knowledge creation capability.

Based on the hierarchical perspective of dynamic capability theory, knowledge

dynamic capability is a higher-order dynamic capability that helps enterprises break path dependence and core rigidity by dynamically adjusting their knowledge base and structure. It enables them to identify and capture innovation opportunities. From a knowledge perspective, enterprise innovation activities involve the recombination of old and new knowledge. Therefore, the dynamic management capability of knowledge influences how enterprises allocate resources, explore opportunities (Han & Li, 2015), and ultimately how they drive business model innovation.

Firstly, knowledge acquisition capability expands the depth and breadth of enterprises' knowledge sources, enabling them to quickly identify windows of opportunity for business model innovation and transform innovation opportunities. In the digital era, the boundaries of business model innovation have become increasingly blurred. It is difficult to predefine the sources of knowledge, as stakeholders such as customers, competitors, suppliers, and governments can all become sources of knowledge for enterprises. Enterprises with knowledge acquisition capability can obtain diverse knowledge from various sources through an open approach, providing knowledge resource support for business model innovation.

Secondly, knowledge integration capability helps enterprises to integrate internal and external knowledge, to build unique and valuable knowledge systems, and thereby to influence business model innovation. When innovating business models, enterprises need to integrate and reshape existing knowledge elements. They combine externally acquired knowledge from multiple channels with their existing internal knowledge structures, thereby exploring new development opportunities.

Finally, knowledge reconfiguration capability is the ability of enterprises to iteratively optimize the integrated knowledge system. Knowledge dynamic capability is embedded in enterprise processes and its effectiveness is influenced by path dependence and developmental opportunities. Enterprises can continuously update their existing knowledge structures through knowledge reconfiguration, enabling their capability to evolve. This helps them to break free from path dependence traps, to capture new market and technological opportunities, and ultimately to drive business model innovation.

The above analyses indicate that enterprises can quickly convert data resources into knowledge resources and generate knowledge dynamic capability through the use of big data capability, thereby promoting business model innovation. Many domestic and foreign research outcomes also support this observation to varying degrees. For example, Yang et al. (2017) believe that enterprises need to effectively mine massive data, extract knowledge and information, and make effective use of big data to effectively drive enterprise innovation. Mikalef et al. (2019) find that in driving incremental and radical innovation with big data capability,

dynamic capability has a completely mediation effect. Feng (2021) finds through empirical research that big data analysis drives process innovation in enterprises by positively influencing knowledge dynamic capability. Tang et al. (2020) note through a case study of Ant Financial that Ant Financial's business model innovation is driven by big data capability, with knowledge creation playing a bridging role between the two.

Based on above, the following hypothesis is proposed.

H2: Knowledge dynamic capability has a mediating effect on the relationship between big data capability and business model innovation.

2.3 | Legitimacy of Innovation and Business Model Innovation

Legitimacy refers to the extent to which the actions of an entity are considered acceptable, appropriate, or agreeable within the normative, value, and standard construction of society (Suchman, 1995). The acquisition of legitimacy helps enhance an organization's ability to attract and acquire resources, and it has a direct and significant impact on the survival and development of the organization (Suchman, 1995). The legitimacy of innovation refers to the degree to which an organization's innovative behavior and outcomes are accepted by stakeholders within the social ethics, norms, and legal framework (Wang & Zhang, 2016). The impact of legitimacy on business model innovators has the following two manifestations (Wang & Zhang, 2016). Firstly, recognition by stakeholders is itself a critical resource, and its importance for business model innovation is no less than those of human resources, capital, technology, and other resources. Secondly, the high acceptance by stakeholders of an organization's innovative behavior creates a favorable external environment for business model innovation. The legitimacy of business model innovation requires that the business model aligns with existing laws, regulations, industry standards, social value systems, etc., and can enhance stakeholders' recognition and understanding of the organization's business model innovation. This creates a favorable external environment for business model innovation and improves its efficiency (Pan & Xie, 2021). Based on this, the following hypothesis is proposed.

H3: The legitimacy of innovation positively influences business model innovation.

2.4 | The Moderation Effect of Innovation Legitimacy

As a novel enterprise capability, big data capability often exhibits foresight and novelty in driving business model innovation, which may go beyond the existing acceptance range of stakeholders. Innovation legitimacy can enhance public awareness of the practical application of big data and attract more wide-ranging and higher-level resources and recognition for business model innovation. The moderation effect of innovation legitimacy on the relationship between big data capability and business model innovation can be observed in two ways. Firstly, from the perspective of acquiring big data resources, the understanding and support of novelty

directly relate to their ability to mobilize resources. Legitimacy is of paramount importance for mobilizing resources for the survival and development of new technologies (Ren, 2017). Legitimacy is an "invisible threshold" that enterprises must cross to mobilize various resources for achieving business model innovation. It is a determining factor for enterprises in a competitive environment and is also known as a resource within resources. Big data is an essential resource for enterprise management and development, and the legitimacy of business model innovation is the key to the effective utilization of big data capability. Secondly, from the perspective of big data application, the abundance of resources is one of the influencing factors for successful transformation of business models. However, social norms should not be overlooked. According to the legitimacy theory, business model innovation is essentially an institutional innovation that needs to be based on legitimacy and needs to obtain acceptance and support from internal and external stakeholders of the enterprise. Under the socialized institutional framework, the practical application of big data by enterprises must comply with the requirements of institutional norms and meet the social expectations and cognition of relevant stakeholders. This will enable shareholders, customers, employees, and other stakeholders to acknowledge the form, impact, and results of big data capability, thereby effectively driving business model innovation. Based on the above, this hypothesis is proposed.

H4: The innovation legitimacy positively moderates the relationship between big data capability and business model innovation.

In summary, the research model of this study is illustrated in Figure 1.

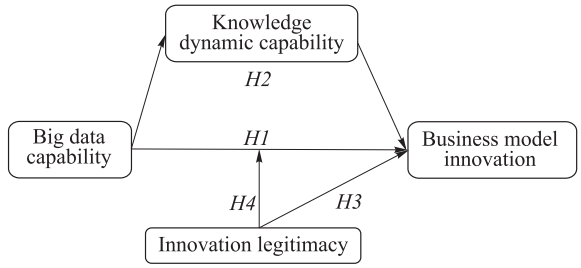

Figure 1 Research Model

3 Research Design

3.1 | Research Methods

This study utilizes both hierarchical regression analysis and fsQCA for data analysis. On the one hand, hierarchical regression analysis only relies on constructing fixed variable paths based on theory for simple statistical validation. Hierarchical regression analysis cannot analyze the complex mechanisms of the interaction between multiple independent variables on business model innovation, such as discovering configurational relationships and the equifinality effect among various factors. On the other hand, fsQCA, based on configurational thinking, conducts complex causal analysis of each independent variable, providing an important supplement to the net effect thinking in mainstream quantitative research (Fiss, 2011). Therefore, this study uses fsQCA to conduct confirmatory

analysis of the relationship between various independent variables and business model innovation, aiming to enhance the credibility and robustness of the research findings and obtain more practical research conclusions.

3.2 | Sample Selection and Data Collection

This study conducts a questionnaire survey on enterprises that have already used big data application, and targets founders and mid-senior-level managers. A total of 500 questionnaires are distributed in both electronic and paper formats. A total of 373 questionnaires are collected, and each collected questionnaire is carefully examined. 71 questionnaires are excluded from further analysis due to incomplete responses, obvious patterns in the selection of options, or consistently choosing the same option, which make them invalid for subsequent analysis. Finally, 302 questionnaires are valid for analysis, resulting in an effective response rate of 60.4%. Table 1 presents the basic characteristics of the sample.

Table 1 Descriptive Statistics of Sample Basic Information

Statistical variable	Feature	Frequency	Percentage (%)
Gender	male	182	60.265
	female	120	39.735
Age	25 years old and under	1	0.331
	26–30 years old	48	15.894
	31–35 years old	148	49.007
	36–40 years old	79	26.156
	41 years old and above	26	8.612
Position	Middle manager	183	60.596
	Top management	119	39.404
Academic degree	High school and below	2	0.662
	Professional training college qualification	5	1.656
	Bachelor level qualification	178	58.940
	Master level qualification	100	33.113
	Doctorate	17	5.629
Enterprise age	Less than 5 years	35	11.589
	5–10 years	103	34.106
	11–20 years	105	34.768
	More than 20 years	59	19.537

(*To be continued*)

(*Continued*)

Statistical variable	Feature	Frequency	Percentage (%)
Enterprise size	Less than 10 million yuan	34	11.258
	10–50 million yuan	69	22.848
	50–100 million yuan	80	26.490
	100–400 million yuan	68	22.517
	More than 400 million yuan	51	16.887
Enterprise nature	State-owned enterprise	49	16.225
	Private enterprise	222	73.510
	Fully foreign-owned enterprise	21	6.954
	Other	10	3.311
Industry type	Information transmission, software and information technology services	201	66.556
	Manufacturing industry	33	10.927
	Wholesale and retail trade	15	4.967
	Financial industry	24	7.947
	Scientific research and technical service industry	9	2.980
	Other	20	6.623

3.3 | Variable Measurement

In this study, a Likert five-point scale is used to measure the variables in the questionnaire. The measurement of big data capability is adapted from the studies of Lin and Kunnathur (2019) and Xie et al. (2018), with a total of 11 items. The measurement of knowledge dynamic capability is based on the scales proposed by Wang et al. (2007) and Han and Li (2015), with a total of 9 items. The measurement of innovation legitimacy is adapted from the scales used by Suchman (1995) and Wang and Zhang (2016) in their research, with a total of 9 items. The measurement of business model innovation is based on the scales used by Zott and Amit (2007), with a total of 12 items. Additionally, this study controls for the basic characteristics of enterprises, including their age, size, ownership nature, and industry affiliation.

4 Data Analysis and Hypothesis Testing

4.1 | Reliability and Validity Testing

Table 2 reports the results of reliability testing for the scales used in this study. The results show that the Cronbach's alpha coefficients and composite reliabilities (CR) for all scales are above the critical value of 0.7, indicating high reliability of the scales used.

Table 2 Reliability Testing of Scales

Variable	Cronbach's alpha coefficient	CR
Big data capability	0.708	0.790
Knowledge dynamic ability	0.771	0.831
Innovation legitimacy	0.751	0.818
Business model innovation	0.835	0.869

Validity testing includes three aspects: content validity, convergent validity, and discriminant validity.

4.1.1 *Content Validity*

The scales used in this study are based on existing mature scales, and suggestions for improvement in item wording and question quality are obtained from relevant researchers and practitioners in the field. The final version of the questionnaire is revised based on these suggestions. Therefore, the scales used in this study have high content validity.

4.1.2 *Convergent Validity*

The factor loadings of the variables, including big data capability, knowledge dynamic capability, innovation legitimacy, and business model innovation, range from 0.5 to 0.7, indicating good convergent validity of the scales used in this study.

4.1.3 *Discriminant Validity*

A four-factor baseline model and three comparative models are constructed, and their goodness of fit is compared. Table 3 shows that the four-factor baseline model has better fit indices than the other three comparative models, indicating good discriminant validity of the scales used in this study.

4.2 | Common Method Bias Testing

To mitigate and measure the potential common method bias, this study follows

Table 3 Comparison of Fit Indices for Different Models

Model	χ^2	*df*	χ^2/df	RMSEA	SRMR	CFI	IFI
Benchmark model: BDA, KDA, IL, BMI	1049.232	740	1.418	0.037	0.0513	0.903	0.905
Three-factor comparison model: BDA+KDA, IL, BMI	1205.951	770	1.566	0.043	0.0548	0.863	0.865
Two-factor comparison model: BDA+KDA, IL, BMI	1302.734	780	1.670	0.047	0.0580	0.836	0.838
Single-factor comparison model: BDA+KDA, IL, BMI	1347.191	782	1.723	0.049	0.0579	0.823	0.825

Note. BDA represents big data capability, KDA represents knowledge dynamic capability, IL represents innovation legitimacy, BMI represents business model innovation, "+" represents combining multiple factors into one factor.

the suggestions of Podsakoff et al. (2003) by implementing both procedural controls and statistical tests.

In terms of procedural controls, the following measures are taken in advance. For English language scales, the questionnaire is translated and back-translated, with careful comparison to provide clear and understandable wording for the respondents. To minimize potential psychological interference during questionnaire completion, the order of the scales is randomized, the purpose of data collection is explained to the respondents, and anonymity is ensured. For post-test examination, the following methods are employed.

Harman's single-factor test is conducted. Principal component factor analysis with varimax rotation is performed on the item data of all variables, resulting in a Kaiser-Meyer-Olkin (KMO) value of 0.911, χ^2 value of 3817.625, and p value of 0.000. The first factor accounts for the maximum variance contribution rate of 26.059%. Correlation analysis of latent variables is conducted. Table 4 shows that the correlation coefficients among the latent variables in this study do not exceed 0.9. This indicates that the questionnaire data in this study do not suffer from serious common method bias issues.

4.3 | Correlation Analysis

Table 4 suggests that there is a significant positive correlation between big data capability and knowledge dynamic capability (r=0.782, p<0.001), as well as between big data capability and business model innovation (r=0.727, p<0.001). There is also a significant positive correlation between knowledge dynamic capability and business model innovation (r=0.759, p<0.001). Furthermore, there is a significant positive correlation between innovation legitimacy and business model innovation (r=0.622, p<0.001). These findings provide a foundation for further empirical testing.

4.4 | Hypothesis Testing

This study conducts hypothesis testing on the direct relationship between big data capability and business innovation, as well as the direct relationship between innovation legitimacy and business model innovation. The empirical test results are shown in Table 5. Model 1 is the baseline model that includes control variables only. When the independent variable of big data capability is included in Model 1, Model 2 is obtained, which shows that big data capability has a significant positive impact on business model innovation (β=0.692, p<0.01), supporting *H1*. Model 4 shows that innovation legitimacy has a significant positive impact on business model innovation (β=0.588, p<0.01), indicating that *H3* is supported.

Furthermore, this study employs the sequential method proposed by Wen et al. (2004) to test the mediation effect of knowledge dynamic capability between big data application capability and business model innovation. In the first step, the influence of the independent variable on the dependent variable is analyzed. Model 2 shows that big data capability has a significant positive impact on

Table 4 Correlation Coefficient Analysis

Variable	1	2	3	4	5	6	7	8
Enterprise age	1							
Enterprise size	0.280***	1						
Enterprise nature	−0.210***	−0.176**	1					
Industry type	0.110	−0.075	0.290***	1				
Big data capability	0.024	0.113	0.159**	0.205***	1			
Knowledge dynamic capability	0.011	0.119*	0.123*	0.168**	0.782***	1		
Legitimacy of innovation	−0.042	0.057	0.129*	0.158**	0.749***	0.745***	1	
Business model innovation	0.078	0.143*	0.151**	0.244***	0.727***	0.759***	0.622***	1
Mean	0.660	0.390	0.740	0.670	4.129	4.052	4.231	3.961
Standard deviation	0.475	0.489	0.442	0.473	0.410	0.499	0.456	0.552
Tolerance	0.886	0.885	0.856	0.885	0.318	0.329	0.369	NA
VIF	1.139	1.130	1.168	1.130	3.145	3.145	2.708	NA

Note. *** indicates $p<0.01$, ** indicates $p<0.05$, * indicates $p<0.1$ (two-tailed test); NA indicates not applicable.

business model innovation (β=0.692, $p<0.01$). In the second step, the influence of the independent variable on the mediating variable is examined. Model 5 demonstrates that big data capability has a significant positive influence on knowledge dynamic capability (β=0.776, $p<0.01$). In the third step, both the independent variable and the mediating variable are tested for their effects on the dependent variable. Model 3 reveals that big data capability significantly affects the dependent variable, which is business model innovation (β=0.314, $p < 0.01$), and knowledge dynamic capability also significantly influences business model innovation (β=0.487, $p<0.01$). However, the regression coefficient of big data capability decreases (0.314<0.692). This indicates that the introduction of knowledge dynamic capability weakens the promoting effect of big data capability on business model innovation, suggesting that big data capability partially mediates between the two variables. *H2* is supported.

Table 5 Hypothesis Testing on the Relationships between Big Data Capability, Innovation Legitimacy, and Business Model Innovation

Variable	Business model innovation				Knowledge dynamic capability
	Model 1	Model 2	Model 3	Model 4	Model 5
Enterprise age	0.086	0.063	0.071	0.105**	–0.017
Enterprise size	0.159**	0.062	0.044	0.104**	0.036
Enterprise nature	0.132*	0.035	0.035*	0.073	–0.001
Industry category	0.227***	0.103*	0.099	0.149**	0.010
Big data capability		0.692***	0.314***		0.776***
Knowledge dynamic capability			0.487***		
Innovation legitimacy				0.588***	
R^2	0.105	0.547	0.639	0.436	0.613
Adj. R^2	0.093	0.540	0.632	0.427	0.606
F	8.689	71.615***	87.091***	45.782***	93.628***

Note. *** indicates $p<0.01$, ** indicates $p<0.05$, * indicates $p<0.1$ (two-tailed test).

Finally, this study utilizes the Process program in SPSS 23.0 software to conduct the moderation effect test of innovation legitimacy. The Bootstrap method is employed with a sample size of 5,000 and a confidence level set at 95%. Model 1 is selected for the analysis, and the results are presented in Table 6. Under different levels of innovation legitimacy (low, medium, and high), the direct effect values of the independent variable on the dependent variable increase accordingly. Additionally, the confidence intervals for different levels of innovation legitimacy do not include 0. Therefore, *H4* is supported.

To visually demonstrate the moderation effect, this study has created a moderation effect plot of innovation legitimacy, as shown in Figure 2.

Table 6 Moderation Effect Test of Innovation Legitimacy

Moderator	Direct effect	Standard error	95% confidence interval	
Low innovation legitimacy	0.718***	0.083	0.554	0.881
Mid innovation legitimacy	0.752***	0.081	0.592	0.911
High innovation legitimacy	0.785***	0.094	0.601	0.970

Note. High innovation legitimacy represents the mean value plus one standard deviation, medium innovation legitimacy represents the mean value, and low innovation legitimacy represents the mean value minus one standard deviation.

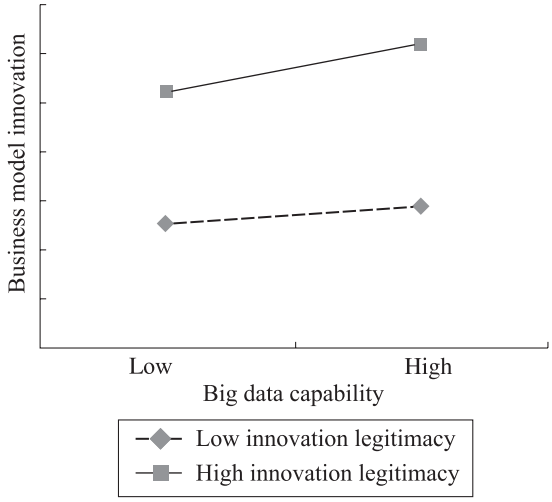

Figure 2 Moderation Effect of Innovation Legitimacy

4.5 | FsQCA

4.5.1 *Data Calibration*

Calibration refers to the process of assigning fuzzy set membership scores to cases and conditions. Following Ragin's (2008) suggestion, data are transformed into fuzzy set membership scores using the critical values of 0.95, 0.5, and 0.05. The calibration values for each variable are shown in Table 7.

4.5.2 *Analysis of the Necessary Individual Condition*

The analysis of the necessity of individual conditions examines whether a specific antecedent condition is a necessary condition for the occurrence of the outcome (Ragin, 2008). The results of the analysis are shown in Table 8. The criterion for determining whether an individual antecedent condition is a necessary condition for the outcome is whether the consistency level is greater

Table 7 Data Calibration

Standard	Condition			
	Big data capacity	Knowledge dynamic capacity	Innovation legitimacy	Business model innovation
Full affiliation	4.636	4.667	4.778	4.583
Partial affiliation	4.182	4.111	4.333	4.000
No affiliation at all	3.364	3.111	3.444	2.921

Table 8 Analysis of Necessary Conditions

Antecedent condition	High business model innovation	
	Consistency	Coverage
Big data capability	0.840	0.874
~ Big data capability	0.452	0.547
Knowledge dynamic capability	0.853	0.886
~ Knowledge dynamic capability	0.453	0.549
Innovation legitimacy	0.782	0.843
~ Innovation legitimacy	0.484	0.563

Note. ~ represents the absence of a condition.

than 0.9. Table 8 shows that the consistency levels of all three antecedent conditions are below the threshold of 0.9. Therefore, none of the three antecedent conditions constitute a necessary condition for high business model innovation.

4.5.3 *Analysis of the Sufficiency of Configurational Conditions*

The analysis of the sufficiency of configurational conditions is based on different combinations of antecedent conditions to determine whether the real-world configurations would lead to the occurrence of the outcome condition. Following the authoritative fsQCA research (Fiss, 2011), this study sets the consistency threshold at 0.84 and the frequency threshold at 6. The fsQCA procedure produces three different types of configurations: complex, parsimonious, and intermediate configurations. The results of the configurational analysis are shown in Table 9, which identifies two paths of configurations that influence high business model innovation. The table reports the intermediate configurations, where the core conditions are elements that appear in both intermediate and parsimonious configurations (Fiss, 2011).

Based on Table 9, the overall consistency of the solution is 0.862, which is higher than the acceptable minimum standard of 0.75. This is similar to the R^2 in empirical research, indicating a good overall consistency of the configurations. The coverage is 0.857, which means that the two paths can explain 85.7% of the cases. For Configuration 1, the antecedent conditions are big data capability and dynamic knowledge capability. These two conditions are the core conditions that lead to high business model innovation and provide further support for *H1*. The coexistence of big data capability and dynamic knowledge capability further validates *H2*. For

Table 9 Configurational Analysis

Conditional variable	High business model innovation	
	Configuration 1	Configuration 2
Big data capability	●	●
Knowledge dynamic capability	●	
Innovation legitimacy		●
Consistency	0.934	0.912
Original coverage	0.776	0.722
Unique coverage	0.037	0.010
Consistency of overall solutions	0.862	
Coverage of overall solution	0.857	

Note. ● represents a core condition, while a blank space indicates that the condition may or may not be present.

Configuration 2, the antecedent conditions are big data capability and innovation legitimacy. These two conditions are the core conditions that lead to high business model innovation and further support *H3* and *H4*.

4.6 | Robustness Test

In terms of the robustness test of the regression analysis, this study considers the institutional environment as an important contextual factor affecting the effectiveness of organizational management practices (Zhao et al., 2019). A good institutional environment helps enterprises utilize resources more effectively for business model innovation. Therefore, this study adopts the marketization index compiled by Wang et al. (2019) to measure the institutional environment in different regions and includes it as a control variable in the regression equation to retest the conclusions. Additionally, after adding the institutional environment as a control variable, the Bootstrap bootstrapping method is used to retest the hypotheses from the previous analysis. Due to space limitations, the results of this robustness test are not reported in the paper but are available upon request. Regarding the robustness test of the configurational analysis, this study follows the suggestion of Ragin (2008) and adjusts the data calibration method. The data are transformed into fuzzy set membership scores using the critical values of "5, mean value, 1." The results are shown in Table 10. The robustness test results, mentioned above, once again support the conclusions presented in the previous analysis.

Table 10 Robustness Test of Configurational Analysis

Conditional variable	High business model innovation	
	Configuration 1	Configuration 2
Big data capability	●	●
Knowledge dynamic capability	●	
Innovation legitimacy		●
Consistency	0.954	0.939
Original coverage	0.887	0.880
Unique coverage	0.016	0.011
Consistency of overall solutions	0.914	
Coverage of overall solution	0.924	

Note. ● indicates core conditions, while a blank space indicates that the condition may be either present or absent.

5 Conclusions and Implications

5.1 | Conclusions

This paper empirically tests the impact of big data capability on business model innovation using regression analysis and fsQCA. It explores the mediation effect of knowledge dynamic capability and the moderation effect of innovation legitimacy. The main conclusions are as follows.

First, big data capability has a significant positive impact on business model innovation and positively promote knowledge dynamic capability. This study reconfirms that big data capability is a key factor in driving business model innovation, which is consistent with previous research (Xie et al., 2018). Big data capability has a significant positive impact on knowledge dynamic capability, which aligns with the research on the relationship between big data and knowledge-related capability (Feng, 2021; Tang et al., 2020).

Second, knowledge dynamic capability partially mediates the relationship between big data capability and business model innovation. This finding effectively answers the question of how big data capability drive business model innovation, further clarifying the role and importance of knowledge dynamic capability in the process of big data capability influencing business model innovation. It is a productive attempt to open the "black box" of business model innovation for organizations.

Third, innovation legitimacy has a positive impact on business model innovation and positively moderates the relationship between big data capability and business model innovation. This finding validates the research conducted by Wang and Zhang (2016). In addition, Li and Xia (2013) have systematically explained the benefits of big data-driven business model innovation and pointed out the dark side of big data, such as information security and privacy protection. This provides a new direction for exploring the relationship between big data and business model innovation, and the present study contributes to the understanding of this research direction from the perspective of innovation legitimacy.

Fourth, through further qualitative comparative analysis, two configurational paths influencing business model innovation are identified. The configuration matching of big data capability, knowledge dynamic capability, and innovation legitimacy can promote business model innovation. Configuration 1, formed by the synergy of big data capability and knowledge dynamic capability, and Configuration 2, formed by the synergy of big data capability and innovation legitimacy, can explain more than 70% of the cases, indicating the significant role of synergistic linkages among these elements in driving business model innovation.

5.2 | Implications

First, enterprises should prioritize data resources and cultivate big data capabilities. With the continuous progress of informatization, various aspects of economic activities are filled with a large amount of data, and enterprises also accumulate many data resources in their operations. The massive and complex unstructured data presents opportunities and challenges

for enterprises to implement precise and effective business analysis strategies. Managers need to elevate data resources to the level of capability. Only by possessing effective management capabilities for big data resources, technology, and talents can enterprises accurately identify and match the resources needed for business model innovation, more reasonably change their value proposition, value creation, and value delivery, and improve the targeting and success rate of big data-driven business model innovation, thereby truly unleashing the commercial value of big data.

Second, enterprises should build knowledge-based organizations and cultivate knowledge dynamics capability. The development of the digital economy and the internet economy is closely related to the knowledge economy. Managers need to distinguish between operational capabilities that ensure the survival of the company and higher-order capabilities that help the enterprise gain sustained competitive advantage. In the process of transforming business models using big data capabilities, the bridging role of knowledge dynamics capabilities should be leveraged from a strategic perspective. Particularly in the context of the big data era, the intensification of business activities triggered by big data is evident, and the importance of big data capabilities is self-evident. However, enterprises should not only focus on the accumulation of big data resources and technologies but also pay attention to long-term development and shape knowledge-based dynamic capabilities. The integration and application capabilities of knowledge have higher strategic value for enterprises.

Third, enterprises should pay attention to the legitimacy of business model innovation, as it relates to their capability to obtain big data resources in a reasonable, legitimate, and sustainable manner. While the application of big data empowers enterprises in business model innovation, it has also led to many legally controversial issues, such as swindle acquaintances based on big data, services of locating and recording customer locations and dwell time, and tracking customer browsing history. This requires enterprises to seek a balance between business model innovation and customer rights. Otherwise, enterprises may lose their survival space due to insufficient legitimacy in innovation, and naturally, they cannot leverage big data capabilities to enhance business model innovation.

Acknowledgments This paper is supported by general project (No. 71672080, 72072086) of the National Natural Science Foundation of China.

References

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, (8), 113–131.

Amit, R., & Zott, C. (2012). Creating value through business model innovation. *MIT Sloan Management Review,* (3), 41–49.

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems*, (4), 4–39.

Cheng, G., & Li, M. (2014). Research on

mechanisms for cultivating big data capability in enterprises. *Journal of Modern Information*, (3), 7–11 (程刚, 李敏 (2014). 企业大数据能力培育机制研究. 现代情报, (3), 7–11).

Cheng, X. (2018). The personal data rights in the era of big data. *Social Sciences in China*, (3), 102–122 (程啸 (2018). 论大数据时代的个人数据权利. 中国社会科学, (3), 102–122).

Ding, X. & Liu, X. (2018). The analytical framework for enterprise innovation management and transformation in the big data era. *Science Research Management*, (12), 1–9 (丁雪辰, 柳卸林 (2018). 大数据时代企业创新管理变革的分析框架. 科研管理, (12), 1–9).

Feng, J. (2021). The value chain strategy of big data analysis on e-commerce platforms from the perspective of knowledge dynamic capability. *China Business and Market,* (2), 37–48 (丰佳栋 (2021). 知识动态能力视角的电商平台大数据分析价值链战略. 中国流通经济, (2), 37–48).

Fiss, P. C. (2011). Building better casual theories, a fuzzy set approach to typologies in organizational research. *Academy of Management Journal*, (2), 393–420.

Ghasemaghaei, M., & Calic, G. (2020). Assessing the impact of Big data on firm innovation performance: Big data is not always better data. *Journal of Business Research,* 108, 147–162.

Han, Y. Q., & Li, D. Y. (2015). Effects of intellectual capital on innovative performance: The role of knowledge-based dynamic capability. *Management Decision,* (1), 40–56.

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely A (2016). Capturing value from big data–a taxonomy of data-driven business models used by start-up firms. *International Journal of Operations & Production Management*, (10), 1382–1406.

Jiao, H., Yang, J., & Ying, Y. (2021). A review of dynamic capability research and suggestions for conducting contextualized research in China. *Journal of Management World*, (5), 191–210 (焦豪, 杨季枫, 应瑛 (2021). 动态能力研究述评及开展中国情境化研究的建议. 管理世界, (5), 191–210).

Li, S., Wu, Y., & Bao, R. (2021). Big data, personal information protection, and price discrimination: An analysis based on the vertical differentiated duopoly model. *Economic Research Journal*, (1), 43–57 (李三希, 武玙璠, 鲍仁杰 (2021). 大数据、个人信息保护和价格歧视——基于垂直差异化双寡头模型的分析. 经济研究, (1), 43–57).

Li, W., & Xia, J. (2013). Business model innovation based on "Big Data." *China Industrial Economics*, (5), 83–95 (李文莲, 夏健明 (2013). 基于"大数据"的商业模式创新. 中国工业经济, (5), 83–95).

Lin, C., & Kunnathur, A. (2019). Strategic orientations, developmental culture, and big data capability. *Journal of Business Research*, 105, 49–60.

Liu, Y., Dong, J., & Wei, J. (2020). Digital innovation management: Theoretical framework and future research. *Journal of Management World,* (7), 198–217, 219 (刘洋, 董久钰, 魏江 (2020). 数字创新管理：理论框架与未来研究. 管理世界, (7), 198–217, 219).

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics capability and innovation: The mediating role of dynamic capability and moderation effect of the environment. *British Journal of Management*, (2), 272–298.

Pan, D., & Xie, H. (2021). The evolutionary game between government supervision and e-commerce enterprises' swindle acquaintances behavior in the digital economy. *Economy and Management*, (1), 77–84 (潘定, 谢菡 (2021). 数字经济下政府监管与电商企业"杀熟"行为的演化博弈. 经济与管理, (1), 77–84).

Podsakoff, P. M., Mackenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, (5), 879–903.

Ragin, C. C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond*. University of Chicago Press.

Ren, M. (2017). The key to success of technology application: An explanation from the intra-organizational legitimacy perspective. *Sociological Studies*, (3), 169–192, 245 (任敏 (2017). 技术应用何以成功?——一个组织合法性框架的解释. 社会学研究, (3), 169–192, 245).

Suchman, M. C. (1995). Managing legitimacy: Strategic management and institutional approaches. *Academy of Management Review*, (3), 571–610.

Tang, B., Lu, Y. & Ye, Y. (2020). Research on platform enterprise knowledge creation model from the perspective of big data capability. *Information Studies: Theory & Application*, (7), 123–129 (唐彬, 卢艳秋, 叶英平 (2020). 大数据能力视角下平台企业知识创造模型研究. 情报理论与实践, (7), 123–129).

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capability and strategic management. *Strategic Management Journal*, (7), 509–533.

Teece, D. J. (2018). Business models and dynamic capability. *Long Range Planning,* (1), 40–49.

Wamba, S. F., Gunasekaran, A., Akter, S., & Krogstie, J. (2017). Big data analytics and firm performance: Effects of dynamic capability. *Journal of Business Research,* (8), 356–365.

Wang, E., Klein, G., & Jiang, J. J. (2007). IT support in manufacturing firms for a knowledge management dynamic capability link to performance. *International Journal of Production Research*, (11), 2419–2434.

Wang, B., & Zhang, S. (2016). Business model innovation, employee absorptive capacity and innovation legitimacy: A cross-level empirical analysis. *Science Research Management*, (11), 1–10 (王炳成, 张士强 (2016). 商业模式创新、员工吸收能力和创新合法性——跨层次的实证分析. 科研管理, (11), 1–10).

Wang, X., Fan, G., & Hu, L. (2019). *Report on marketization index by provinces in China* (2018). Social Sciences Academic Press (China) (王小鲁, 樊纲, 胡李鹏 (2019). 中国分省份市场化指数报告 (2018). 社会科学文献出版社).

Wei, J., Wu, A., & Peng, X. (2014). Chinese strategic management research: situational issues and theoretical frontiers. *Journal of Management World*, (12), 167–171 (魏江, 邬爱其, 彭雪蓉 (2014). 中国战略管理研究：情境问题与理论前沿. 管理世界, (12), 167–171).

Wen, Z., Zhang, L., Hou, J., & Liu, H. (2004). Testing and application of the mediating effects. *Acta Psychologica Sinica,* (5), 614–620 (温忠麟, 张雷, 侯杰泰, 刘红云 (2004). 中介效应检验程序及其应用. 心理学报, (5), 614–620).

Winter, S. G. (2003). Understanding dynamic capability. *Strategic Management Journal*, (10), 991–995.

Xie, W., Li, Z., Su, F. & Wang, Y. (2018). Top management support, big data capability and business model innovation. *R&D Management*, (4), 152–162 (谢卫红, 李忠顺, 苏芳, 王永健 (2018). 高管支持、大数据能力与商业模式创新. 研究与发展管理, (4), 152–162).

Yang, C., Huang, Q., Li, Z., Liu, K., & Hu, F. (2017). Big data and cloud computing: Innovation opportunities and challenges. *International Journal of Digital Earth,* (1), 13–53.

Yi, J., & Xu, D. (2018). The research on the impact mechanism of big data on business model innovation: An Analytical Framework. *Science & Technology Progress and Policy*, (3), 15–21 (易加斌, 徐迪 (2018). 大数据对商业模式创新的影响机理——一个分析框架. 科技进步与对策, (3), 15–21).

Zeng, Q. (2019). The research on business model innovation driven by big data: An analysis unit. *Studies in Science of Science*, (6), 1142–1152 (曾锵 (2019). 大数据驱动的商业模式创新研究. 科学学研究, (6), 1142–1152).

Zhang, J., & Long, J. (2022). Digital transformation, dynamic capability, and firm innovation performance: Empirical evidence from high-tech listed enterprises. *Economy and Management*, (3), 74–83 (张吉昌, 龙静 (2022). 数字化转型、动态能力与企业创新绩效——来自高新技术上市企业的经验证据. 经济与管理, (3), 74–83).

Zhang, J., Long, J., Sun, K. & Guo, L. (2022). Research on the mechanism of new product development advantage formation in technology enterprises from the perspective of dynamic capability: A hybrid method based on NCA and SEM. *Science and Technology Management Research*, (7), 126–136 (张吉昌, 龙静, 孙珂, 吳林义 (2022). 动态能力视角下科技企业新产品开发优势形成机制研究——基于NCA与SEM的混合方法. 科技管理研究, (7), 126–136).

Zhao, Y., Zhang, Z., Feng, T., & Tao, K. (2019). Development of big data, institutional environment, and government governance efficiency. *Journal of Management World*, (11), 119–132 (赵云辉, 张哲, 冯泰文, 陶克涛 (2019). 大数据发展、制度环境与政府治理效率. 管理世界, (11), 119–132).

Zott, C., & Amit, R. (2007). Business model design and the performance of entrepreneurial firms. *Organization Science*, (2), 181–199.