ARTICLE IN PRESS

Information & Management xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Information & Management

journal homepage: www.elsevier.com/locate/im

Data analytics and firm performance: An empirical study in an online B2C platform

Peijian Song[a], Chengde Zheng[a], Cheng Zhang[b], Xiaofeng Yu[a,*]

[a] School of Business, Nanjing University, 22, Hankou Road, Nanjing, 210093, China
[b] School of Management, Fudan University, 670, Guoshun Road, Shanghai, 200433, China

A R T I C L E I N F O

Keywords:
Data analytics
Business value
Product variety
Competitive intensity
Online platform

A B S T R A C T

Data analytics has become an increasingly eye-catching practice in both the academic and the business communities. The importance of data analytics has also prompted growing literature to focus on the design of data analytics. However, the boundary conditions for data analytics in creating value have been largely overlooked in the literature. The objective of this article therefore is to examine the business value of data analytics usage and explore how such value differs in different market conditions. We rely on an online B2C platform as our empirical setting and obtain several important insights. First, both demand-side and supply-side data analytics usage has a positive effect on the performance of merchants. Second, when merchants' product variety is high, the influence of usage toward demand-side data on performance is strengthened, whereas such impact is weakened for supply-side data analytics. Third, when competitive intensity is high, the performance implication of demand-side data analytics usage is strengthened, whereas such impact is not strengthened for supply-side data analytics. This study contributes by advancing the overall understanding of business value of data analytics.

1. Introduction

With dramatic advancement in data collecting, storage, and processing technologies in recent decades, organizations worldwide are exploring new data-enabled ways to compete and win—transforming themselves to take advantage of the vast array of available data to improve decision-making and performance [1–3]. These new opportunities have led many managers to rely less on business decision-making process and more on data itself in their decision-making [4,5]. Data analytics, therefore, has become an increasingly eye-catching practice in both academic and the business communities, and industry studies have highlighted this significant development. For example, on the basis of a survey of over 4000 information technology (IT) professionals from 93 countries and 25 industries, the IBM Tech Trends Report [6] identified business analytics as one of the four major technology trends in the 2010s. In a survey of the state of data analytics by Bloomberg Businessweek [7], 97% of companies with revenues exceeding US$100 million were found to use some form of data analytics [8].

The importance of data analytics has also prompted growing literature to focus on the design of data analytics for generating knowledge and intelligence to support decision-making and strategic objectives [3,9,10]. For example, Bardhan et al. [11] demonstrated that using health IT-related data could help hospitals save millions of dollars by avoiding costly readmission-related penalties. Martens et al. [9] examined the use of massive, fine-grained data on consumer behavior to improve predictive models for targeted marketing. Guo et al. [3] proposed a system framework for extracting representative information from intra-organizational blogging platforms.

The puzzle, however, is why not all firms have implemented this practice? The stream of IT business value research points that value creation process of technological innovations in a firm cannot depart from how to use them [12,13]. For example, Devaraj and Kohli [14] posited and proved that the actual usage of technology is a key variable to explain the impact of IT application on organizational performance. Scholars reveal that IT usage breadth and depth [13], post-adoptive extended use of IT [15], and deployment on usage breadth and depth [16] are critical to associate technological applications and firm performance. Even though an innovation is well developed, it may still depend on different complementary adjustments in a firm, thus leading to variation in firms' performance [17]. Furthermore, the boundary conditions for data analytics in creating value have been largely overlooked in the literature. Changes in some market conditions can enable or inhibit the performance advantages associated with usage toward data analytics.

The objective of this article therefore is to examine the business value of data analytics usage in an online B2C platform context and

* Corresponding author.
E-mail addresses: songpeijian@nju.edu.cn (P. Song), cdzheng@nju.edu.cn (C. Zheng), zhangche@fudan.edu.cn (C. Zhang), xiaofengyu@nju.edu.cn (X. Yu).

https://doi.org/10.1016/j.im.2018.01.004
Received 15 May 2017; Received in revised form 7 January 2018; Accepted 17 January 2018

Table 1
Categorization of Data Analytics Usage.

	Demand-side data analytics usage	Supply-side data analytics usage
Definition	Using data analytics to see buyers' purchases and interactions and listen to buyers' unique wants and needs	Using data analytics to see their internal operational activities and improve their operational efficiencies
Focus	Downstream customers	Internal operation process and upstream suppliers
Area in Organization	Sales; Promotion; Pricing; Customer service/support; Brand or market management	Operations/Supply chain management; Product development; Strategy/ planning; Human resource; Finance
Reference	[7,8,22]	[27,7,24]

explore how such value differs in different market conditions. Online B2C platform (e.g., eBay or eHarmony) refers to an intermediary that facilitates economic transactions between buyers (i.e., demand-side) and merchants (i.e., supply-side) by enabling them to search for feasible contracts [18]. We focus on two critical aspects of data analytics usage: demand-side data and supply-side data. Demand-side data analytics usage refers to using data analytics to see buyers' purchases and interactions and listen to buyers' unique wants and needs. Supply-side data analytics usage relates to merchants' use of data analytics to see their internal operational activities and improve their operational efficiencies. For market conditions, we focus on two attributes: product variety and competitive intensity. Product variety refers to the number of distinct products for each merchant, whereas competitive intensity relates to the extent of competition among merchants in the platform.

We rely on one of the world's largest online B2C platform (i.e., Tmall) as our empirical setting. In particular, a survey was conducted among merchants in the Tmall platform to test our hypotheses. Through the regression analysis, we obtained several important insights. First, usage toward both demand-side and supply-side data analytics has a positive effect on the performance of merchants. Second, the effect of usage toward data analytics on performance is contingent upon the merchants' product variety. Compared with low product variety, the influence of usage toward demand-side data analytics on performance is stronger for high product variety. By contrast, the impact of usage toward supply-side data analytics on performance is weaker for high product variety. Third, we find that the magnitude of performance implications of data analytics depends on competitive intensity. When competitive intensity is high, the value of demand-side data analytics usage is strengthened, whereas the performance implication of supply-side data analytics usage is not weakened and remains constant.

These findings accordingly contribute to extant literature. First, our study contributes by addressing the business value of system usage in data analytics context, which has not yet been explored in the information systems (IS) literature (e.g., [19–21]). Second, we theorize about and test the value of data analytics in different market conditions. Our theoretical framework and empirical results suggest that "a more is better" view is not accurate, rather assessing the business value of data analytics usage requires noting their unique characteristics and fit with market conditions. Finally, this study contributes to platform governance research and generates important implications regarding how the platform may improve its management.

2. Theoretical background

2.1. Data analytics: from insights to value

For the last 2 or 3 years, the field of "big data" has emerged as the new frontier in the wide spectrum of IT-enabled innovations and opportunities allowed by the information revolution [2]. The ever-increasing creation of massive amounts of data has prompted organizations to direct their attention to how to harness and analyze big data [1,22]. Some organizations are developing better ways to collect, incorporate, and act on data analytics effectively, so that they can begin to optimize their organizations. Web-based companies such as Amazon, eBay, and Google test factors that drive performance—from where to place buttons on a Web page to the sequence of content displayed—to determine what will increase sales and user engagement.

To compete in a globally integrated economy, organizations need a comprehensive understanding of markets, consumers, products, regulations, competitors, suppliers, employees, and more. This understanding demands the effective use of data analytics. Data analytics is being used to optimize business processes—to identify the best consumers, select the ideal price, calculate the best supply chain routing, or pick the best person to hire [23,24]. For example, Tesco gathers transaction data on its 10 million consumers through a loyalty card program. It then uses the information to analyze new business opportunities and to inform decisions on pricing, promotions, and shelf allocation. The online grocery FreshDirect shrinks reaction times even further. It adjusts prices and promotions daily or even more frequently on the basis of data feeds from online transactions, visits by consumers to its web site, and consumer service interactions.

Technology for capturing and analyzing information is widely available at ever lower price points. The available data analytics makes it possible for decision makers to completely see their consumers' purchases, payments, and interactions. Businesses will be able to listen to consumers' unique wants and needs about channel and product preferences [25,26]. Many companies use data to support rigorous, constant business experimentation that guides decisions and to test new products, business models, and innovations in consumer experience [23]. The use of business information and analytics differentiates them within their industry [24].

Data analytics has been used in different areas within an organization to address internal or external business initiatives. Along supply chain perspective, we divide data analytics usage into demand-side and supply-side usage (see Table 1). Demand-side data analytics usage refers to the use of data analytics to see buyers' purchases and interactions and listen to buyers' unique wants and needs for market responsiveness. The typical usage areas in organizations are sales, promotion, pricing, consumer service/support, and brand or market management [3,24,28]. Supply-side data analytics usage refers to the use of data analytics to see their internal operational activities and improve their operational efficiencies [29–31]. The typical usage areas in an organization are supply chain management, strategic planning, human resource, finance, and product development [7,32]. Thus, data analytics usage at supply side and at demand side, however, may have different effects and cannot be treated as the same thing [33].

2.2. B2C platform

Online B2C platform provides necessary services to facilitate transactions between merchants and buyers. To facilitate transactions between merchants and buyers, online platforms usually provide data analytics services for merchants. For example, eBay uses data about the behavior of its millions of buyers and launches the Listing Analytics to help merchants get closer to its buyers. Merchants can review key metrics of their listings, see how their listings are performing, and identify best and worst performers. Through the help of Listing Analytics, merchants can get insights into the visibility, clicks, and sales their listings are generating. After learning about best and worst performers, merchants will revise their listings to maximize sales. Amazon also launched the analytics to help merchants compete in the Amazon marketplace. By maximizing all of the opportunities that a platform provides, analytics in Amazon helps merchants make better merchandising and pricing decisions so that they capture as much profit as possible.

The online B2C platform provides a unique context to investigate data analytics value for two reasons. First, merchants on the platform share the same technical infrastructure and analytics functions provided by the platform, thus largely avoiding the variation of technology availability and functionality that may affect their performance. In other words, merchants' data analytics usage on one platform is more comparable than the situation in which they apply different data analytics technologies. Second, merchants share the same customer base in one B2C platform in which merchants' data source and targets are the same, causing their data analytics performance more comparable than those applied in different platforms. The uniqueness of research context enhances the validation of the results in this study.

3. Hypotheses development

3.1. Value of data analytics for merchants

The performance of merchants depends upon how they use the data analytics provided by the platform. Merchants can use demand-side data analytics to boost competitive advantage from displays and for marketing, customer service, and customer experience management [23,24]. If merchants extensively use the demand-side data analytics, they can see their consumers' purchases, payments, and interactions in depth and can listen to consumers' unique wants and needs about product preferences [26]. As a result, merchants can make better decisions about how to choose and promote the products [23,26]. Hence, the launched products can better match the consumers' preferences and needs. The probability that the merchants will have the products that consumers want should be enhanced.

If merchants use the demand-side data analytics more extensively, they can more accurately forecast consumers' demand. By forecasting accurate demand, merchants can improve their supply chain efficiency [34]. The extent of negative effect of consumer demand on logistics service, order fulfillment, and inventory cost should be decreased [35,36]. Therefore, the performance of merchants should be enhanced. We hypothesize the following:

H1. Usage toward demand-side data analytics has a positive effect on the performance of merchants.

Merchants have troves of raw data combined with powerful data analytics to gain insights that can improve operational performance. If merchants extensively use the supply-side data analytics, they can improve their operational efficiency. The errors due to employees or suppliers, shelving, or other causes should not be increased with customer demand [36,37]. With better operation management, the inventory errors that come from customer demand can be decreased. The quality of logistics service and order fulfillment should be enhanced, and inventory cost should be decreased [35,36]. Therefore, the performance of merchants should be improved. We develop the following hypotheses:

H2. Usage toward supply-side data analytics has a positive effect on the performance of merchants.

3.2. Moderating role of product variety

A merchant's high product variety can heighten the effect of usage toward demand-side data analytics on performance by stimulating more volume and more in-depth user behavioral information to leverage data analytics efficiency. First, an increase in product variety implies more searches by consumers trying to find better matches [38]. Consumers can actively perform specific searches that take consumers directly to the product page that displays the product being searched for and can browse more product details through links on the page, which results in more volume of available demand-side data [39]. Much of these added available data are in the form of clicks, images, text, or signals of various sorts and can be recorded down by the system for analysis. Therefore, higher product variety can stimulate more behavioral data that leverage merchants' usage efficiency toward demand-side data analytics than lower product variety.

Second, higher product variety increases the choice set for the consumers and the decision complexity [40] and thus increases both variety of user behaviors such as performing more non-specific searches on product information and more intensive comparisons with other relevant products [39]. These actions provide in-depth information such as browsing paths to explore and discover customer preference and the origin of customer demand [41]. Specifically, path data may contain information about a user's goals, knowledge, and interests [42]. With path analysis techniques that encode the sequence of events, merchants can also predict consumers' future movements in the store and purchase conversion [43].

Furthermore, higher product variety can satisfy a greater range of customer needs [44], increase consumers' purchase probability [45,46], and thus stimulate more consumers' in-depth behaviors to search, browse, and compare products. With in-depth demand-side information, merchants can see consumers' purchases and interactions and listen to consumers' unique wants and needs about product preferences with more depth by applying data analytics [47–49].

Considering that product variety can enhance volume and depth of demand-side information to leverage data analytics efficiency, merchants can make better decisions and create new market opportunities according to the demand-side data analytics in a higher product variety situation than that in the lower one. We thus expect that the product variety positively moderates the enhancing effect of usage toward demand-side data analytics on performance, and we hypothesize the following:

H3. Compared with low product variety, the influence of usage toward demand-side data analytics on the performance is stronger for high product variety.

For the supply side, increasing product variety may strengthen the enhancing effect of usage toward supply-side data analytics on performance in two ways. First, when the product variety is high, merchants can capture more data from supply chains, equipment, and internal processes [37]. Merchants' operational decisions can be based on more evidence, experiments, and more accurate forecasts [4]. These additional data can help merchants in predictive maintenance, process management, procurement, and logistics planning [50]. With more available data to optimize the operations and supply chains, the enhancing effect of usage toward supply-side data analytics should be higher. Second, when the product variety is high, the operational complexity of merchants should be increased [51]. In the context with high operational complexity, the operation management plays a more important role in the performance of merchants. The situation offers more opportunities for data-driven operation optimization and increases the importance of supply-side data analytics to help improve the firm's performance [17]. As a result, the enhancing effect of usage toward supply-side data analytics should be higher. We develop the following hypotheses accordingly:

H4a. Compared with low product variety, the influence of usage toward supply-side data analytics on the performance is stronger for high product variety.

In juxtaposition to H4a, we also consider some mechanisms that make product variety weaken the enhancing effect of usage toward supply-side data analytics. Despite the benefits of high product variety, the product variety may also generate counterproductive effect on logistics service, order fulfillment, and inventory cost [52,34,53]. For merchants, there are significant logistics and order fulfillment complications from product proliferation. As a merchant's product variety increases, the errors in demand forecasting should be increased [35,36], which in turn causes logistics service metrics such as fill rates to decline even as inventory levels increase. A high product variety also contributes to inventory errors. "Phantom products" due to a product variety strategy may arise from input errors by employees or suppliers [36,37]. These negative influences of product variety reduce logistics efficiency and decrease fill rates, which may in turn reduce the enhancing effect of usage toward supply-side data analytics on performance. We develop the following hypothesis accordingly:

H4b. Compared with low product variety, the influence of usage toward supply-side data analytics on the performance is weaker for high product variety.

3.3. Moderating role of market competitive intensity

Promoting competition among merchants is one mechanism that a platform uses to make its ecosystem more competitive and to enhance its overall competitive position [54–56]. A platform can influence the degree of competition in a variety of ways including licensing policies and the use of soft power inducements [57]. Armstrong [54] suggested that a platform can boost its performance by encouraging increasing levels of competition on the supply side. This would effectively stimulate the supply of a larger variety of products and, by force of indirect network effects, increase consumer adoption. A larger installed base of consumers creates a larger market for merchants, which, in turn, creates a cascade of reinforcing effects [58]. Turner et al. [59] found that even when the market for supply side is concentrated, the threat of competition from other merchants will induce incumbents to respond to rivals and introduce new products.

Competitive intensity influences survival and growth of competitors in an industry. When the level of competitive intensity is high, there are more merchants in exactly the same category, and the market shows greater crowding [56]. Intensely competitive markets generate more choices for consumers, which results in less-competitive products being rejected. Merchants must provide high quality and differentiated offerings [60]. At this time, listening to consumers' unique wants and needs about product preferences is more important. When merchants increase their usage toward demand-side data analytics, the performance implications for merchants should be higher. In contrast, when the competitive intensity is low, there are fewer competitors. As a result, the influence of usage toward demand-side data analytics on the performance should be lower (Fig. 1). We hypothesize the following:

H5. Compared with low competitive intensity, the influence of usage toward demand-side data analytics on the performance is stronger for high competitive intensity.

Competitively intense environments are characterized by a cutthroat rivalry, price and promotion wars, and a stiff competition where merchants have the ability to match the competitors' moves [61]. Consequently, the result of a merchant's behavior will no longer be deterministic but stochastic as the behavior is heavily influenced by the actions and contingencies undertaken by competitors [62,63]. Thus, under intensive competition, the predictability diminishes. This requires firms to engage simultaneously in supply-side operation improvement. Merchants must improve their operation management to face greater numbers of competing merchants. Therefore, when merchants increase their usage toward supply-side data analytics, the performance implications for merchants should be higher. In contrast, when the competitive intensity is low, there are fewer competitors. The influence of usage toward supply-side data analytics on the performance should be lower. Therefore, we develop the following hypothesis:

H6. Compared with low competitive intensity, the influence of usage toward supply-side data analytics on the performance is stronger for high competitive intensity.

4. Research methodology

4.1. Research context

Our research context is Tmall (https://www.tmall.com/), a B2C platform operated by Alibaba Group in China. Launched in the year 2012, Tmall has become the largest B2C platform in China, which hosts more than 100,000 brands including Apple, Zara, Bose, Estee Lauder, P &G, and Unilever and more than 200 million consumers. The gross merchandise volume transacted on Tmall increased from RMB847 billion in fiscal year 2015 to RMB1215 billion (US$ 190 billion) in fiscal year 2016 [64]. Merchants on Tmall pay commissions based on a predetermined percentage of transaction value that varies by product category and typically ranges from 0.4% to 5.0%.

Tmall provides value-added services and management tools for merchants. Data analytics service is one of the most important management tools provided by the Tmall platform. Since the beginning of the platform, Tmall has provided a data analytics service called data cube. In December 2015, Tmall launched a new data analytics service named business adviser (https://sycm.taobao.com/) to replace the data cube. In the year 2016, business adviser had served more than 5 million merchants. For merchants with annual sales size larger than 300 thousands, 90% merchants have adopted and used the business adviser.

Through the use of the data analytics, merchants can make better decisions in both demand side and supply side. Through the aggregation of consumer transactions data, the analytics tool provides crucial business data including industry size and trends, online store statistics, and buyer demographics (e.g., age, gender, geo-locations, and shopping behaviors). Merchants can use the tools to analyze industry keyword rankings, spot keywords, and diagnose product description names so that merchants could keep product descriptions on trend and optimize their online store traffic driven by name. In addition; the data analytics can generate report to help improve online store operation

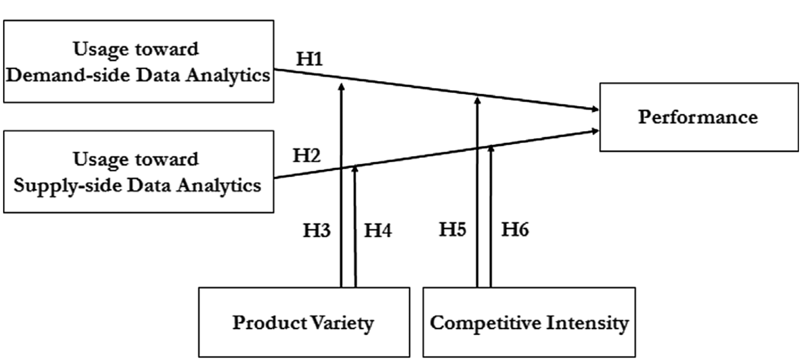

Fig. 1. Research Model.

Table 2
Sample Profile (n = 309).

Demographic Variables	Category	Frequency (Percent)
Occupation	General Manger	206 (66.7%)
	Sales	103 (33.3%)
Years	< 1 year	74 (24.2%)
	1–3 years	129 (41.6%)
	3–5 years	62 (20.0%)
	> 5 years	44 (14.2%)
Number of Employees	< 5	177 (56.5%)
	6–10	54 (17.4%)
	10–20	24 (7.7%)
	> 20	54 (17.4%)
Size of Sales (RMB)	< 100 thousands	111 (36.1%)
	100 thousands–1 million	102 (32.9%)
	> 1 million	96 (31.0%)
Product Category	Clothing and Luggage	76 (24.5%)
	Electronic Products	40 (12.9%)
	Household Items	40 (12.9%)
	Books and Music	15 (4.8%)
	Others	138 (44.8%)

management. By using the data analytics; merchants can identify; track; and analyze competitors in the platform. After comparison and learning from the competitors; merchants will adjust their competitive actions.

4.2. Data collection

A survey was conducted among merchants in the Tmall platform to test our hypotheses. Merchants were randomly chosen from different product categories. The questionnaire was administrated to the merchants in the Tmall platform by means of Skype-like communication device. Tmall created a Skype-like communication service that allows buyers and merchants to fully communicate and exchange information to facilitate transactions. Respondents voluntarily participated without any compensation. Students majoring in electronic business and marketing conducted the interviews. They attended detailed training sessions before engaging in the survey. Interviewers obtained respondents' names and telephone number for validation purposes. We randomly contacted 5% of the respondents to confirm that the interviews were completed as planned. The data were collected over 6 months. A total of 1000 merchants were contacted. However, some merchants refused to participate the survey regardless of our invitation, whereas some merchants agreed to participate but did not return their questionnaires. As a result, 332 merchants agreed to participate the survey and returned the questionnaires (33.2% response rate). After discarding responses with incomplete data and inconsistent demographic information, we collected 309 valid responses.

Table 2 reflects the demographics of the merchants. To test nonresponse bias, the procedure recommended by Armstrong and Overton [65] was adopted. This procedure assumes that late respondents in a sample are similar to theoretical non-respondents. Using *t*-tests to compare early and late respondents according to their demographic characteristics, we found no significant differences between the two groups. A Mann–Whitney test also revealed no significant difference in their demographic characteristics.

4.3. Measurement

Theoretical constructs were operationalized using validated items from prior research. Minor changes in the wordings were made so as to fit them into the current investigation context (see Appendix A in Supplementary material). Measures adapted from Devaraj and Kohli [14] and Kim and Malhotra [66] were used to measure usage toward demand-side data analytics. A three-item, seven-point Likert scale was developed to determine respondents' evaluations of the degree of merchants' use of data analytics to see their consumers' purchases, payments, and interactions and listen to consumers' unique wants and needs about product preferences. The coefficient alpha for this measure is 0.79. For measuring usage toward supply-side data analytics, we used a three-item, seven-point Likert scale as used in the studies by Subramani [31] and Kim and Malhotra [66]. Specially, the scale items ask respondents about the extent and frequency of merchants' use of data analytics to analyze their internal supply-side operational efficiency and result in a coefficient alpha of 0.81.

On the basis of previous studies on product variety (e.g., [67,36]), product variety was measured by a three-item, seven-point Likert scale. Specially, the scale items ask merchants about the extent of their variety and length of their product line and result in a coefficient alpha of 0.84. The measure of competitive intensity, which uses a three-item, seven-point semantic differential scale, was adapted from Jaworski and Kohli [61] and Yeniaras and Unver [63]. The scale items ask respondents about the extent of competitive intensity that merchants perceive and result in a coefficient alpha of 0.87. Measures adapted from Li and Atuahene-Gima [68] were used to measure the performance of merchants. A three-item, seven-point Likert scale was developed to determine respondents' evaluations of the merchants' performance. The coefficient alpha for this measure is 0.77. The control variables are age of merchants, size of merchants, extent of differentiation with other merchants, and product category. Table 3 reflects the correlations between variables.

The convergent and discriminant validities of the constructs were examined with confirmatory factor analysis. First, all items load on their respective constructs, and each loading is large and significant at the 0.01 level, showing a satisfactory convergent validity [69]. In addition, all pairs of constructs pass [78] test of discriminant validity. That is, the amount of variance extracted by each construct (considering measurement error) is greater than the squared correlation between the two constructs.

4.4. Model specification

To evaluate the main effect of usage toward demand-side data analytics and supply-side data analytics on performance, we first estimate the following model:

$$\text{Performance}_i = \beta_0+\beta_1\text{Demand-side data}_t+\beta_2\text{Supply-side data}_i+\beta_3\text{Product variety}_i +\beta_4\text{Competitive intensity}_i + \text{Control variables}_i+\varepsilon_i \quad (1)$$

where, i is the index of the merchants. The dependent variable, Performance_i, is the performance of each merchant. The control variables are age of merchants, size of merchants, extent of differentiation with other merchants, and product category. For Model 1, we used regression with ordinary least squares for estimation. In Model 1, the coefficients of $\text{Demand-side data}_i$ and $\text{Supply-side data}_i$ capture the main effect of usage toward demand-side and supply-side data. In addition, to further capture the moderating effect of product variety and competitive intensity, we created one additional specification by adding interaction terms in Model 1:

$$\text{Performance}_i = \beta_0+\beta_1\text{Demand-side data}_t+\beta_2\text{Supply-side data}_i+\beta_3\text{Product variety}_i+\beta_4\text{Competitive intensity}_i+\beta_5\text{Demand-side data}_i \times \text{Product variety}_i + \beta_6\text{Supply-side data}_i \times \text{Product variety}_i+\beta_7\text{Demand-side data}_i \times \text{Competitive intensity}_i+\beta_8\text{Supply-side data}_i\times \text{Competitive intensity}_i + \text{Control variables}_i+\varepsilon_i \quad (2)$$

Table 3
Correlation Matrix and Descriptive Statistics.

	Variables	Mean	S.D.	Correlation Matrix								
				1	2	3	4	5	6	7	8	9
1.	Performance	4.223	1.083	1								
2.	Demand-side data analytics	4.122	1.306	0.389**	1							
3.	Supply-side data analytics	4.019	1.013	0.324**	0.475**	1						
4.	Product variety	5.311	1.384	0.318**	0.253**	0.261**	1					
5.	Competitive intensity	5.708	1.361	0.058	0.220*	0.216*	0.148	1				
6.	Age of merchants	2.293	1.096	0.141	0.166	0.112	0.061	0.133	1			
7.	Size of merchants	3.381	1.452	0.491**	0.192*	0.130	0.107	0.069	0.437**	1		
8.	Extent of differentiation	4.167	1.445	0.331**	0.353**	0.328**	0.355**	0.059	0.217*	0.238*	1	
9.	Product category	3.685	2.152	0.024	0.005	0.063	−0.031	−0.228*	0.061	−0.041	0.002	1

** $p < 0.01$.
* $p < 0.05$.

4.5. Common method variance

As both dependent and independent variable data were collected from a single informant around the same time, common method variance was a concern. We performed several statistical analyses to assess the severity of this potential issue. First, following Podsakoff and Organ [70], we used the Harman's one-factor test to examine the extent of the bias. The results of principal component analyses indicated that common-method variance is not a problem because several factors with eigenvalues greater than 1 were identified and because no single factor accounted for almost all the variances. Second, following Bentler and Bonnett [71] and Podsakoff et al. [72], we compared the chi-square values of the following three estimations: the null model that has no underlying factor, a common-factor measurement model in which all items have one underlying factor, and our measurement model. The comparison shows that our measurement model fits the data better than a single-factor model; we contend that common method bias is unlikely a serious concern in this study.

Third, following Podsakoff et al. [72] and Liang et al. [73], we included in the model a common-method factor whose indicators included all the principal constructs' indicators and calculated each indicator's variances substantively explained by the principal construct and by the method. The results demonstrate that the average substantively explained variance of the indicators is 0.691, while the average method-based variance is 0.015. The ratio of substantive variance to method variance is about 46:1. In addition, most method factor loadings are not significant. Given the small magnitude and insignificance of method variance, we contend that the method is unlikely to be a serious concern for this study. Finally, to further alleviate the potential of common method bias, we collected objective data about the performance of merchants. We used the objective performance data as the alternative measurement to verify the insights of our analysis. The details are provided in the following robustness check.

5. Data analysis and results

5.1. Hypotheses testing

Table 4 shows the results of the regression analysis and predicts performance. In our conceptual model, we explicitly predicted that the performance of merchants would be related to usage toward demand-side and supply-side data analytics, and this relationship would change with the product variety and competitive intensity. H1 predicts that usage toward demand-side data analytics has a positive effect on the performance of merchants. The coefficient for the usage toward demand-side data analytics in model 1 is positive and significant ($\beta = 0.112$, $p < 0.01$). Therefore, H1 is supported. H2 predicts that usage toward supply-side data analytics has a positive effect on the performance of merchants. H2 is also supported, as the coefficient for the usage toward supply-side data analytics is positive and significant ($\beta = 0.072$, $p < 0.05$).

Table 4
Data Analysis Results.

	Model 1 Parameter Est. (SE)	Model 2 Parameter Est. (SE)
Main effects		
Demand-side data analytics	0.112 (0.040)**	0.161 (0.042)**
Supply-side data analytics	0.072 (0.039)*	0.074 (0.040)*
Product variety	0.172 (0.052)**	0.240 (0.052)**
Competitive intensity	−0.100 (0.044)*	−0.143 (0.050)**
Interactions		
Demand-side data analytics × Product variety		0.114 (0.045)**
Supply-side data analytics × Product variety		−0.120 (0.041)**
Demand-side data analytics × Competitive intensity		0.049 (0.024)*
Supply-side data analytics × Competitive intensity		0.001 (0.031)
Controls		
Age of merchants	0.116 (0.047)**	0.156 (0.048)**
Size of merchants	0.214 (0.038)**	0.242 (0.039)**
Extent of differentiation	0.281 (0.045)**	0.274 (0.045)**
Product category	0.014 (0.021)	0.026 (0.022)
Adjusted R Square	0.472	0.485
F	31.616	23.301
N	309	309

** $p < 0.01$.
* $p < 0.05$.

H3 predicts that the influence of usage toward demand-side data analytics on the performance of merchants is stronger for high product variety, and therefore, H3 is supported. The coefficient for the interaction term between usage toward demand-side data analytics and competitive intensity in model 2 is significant ($\beta = 0.114$, $p < 0.05$). The results of simple slope analysis provide more insights into this moderating effect (see Fig. 2, Panel A). When product variety is low (one standard deviation below mean), usage toward demand-side data analytics has a non-significant effect on performance ($\beta = 0.042$, n.s); when product variety is high (one standard deviation above mean), this effect becomes positive and significant ($\beta = 0.280$, $p < 0.01$).

H4 a predicts that the influence of usage toward supply-side data analytics on the performance of merchants is weaker for high product variety. H4b is supported, as the coefficient for the interaction term between usage toward supply-side data analytics and product variety is

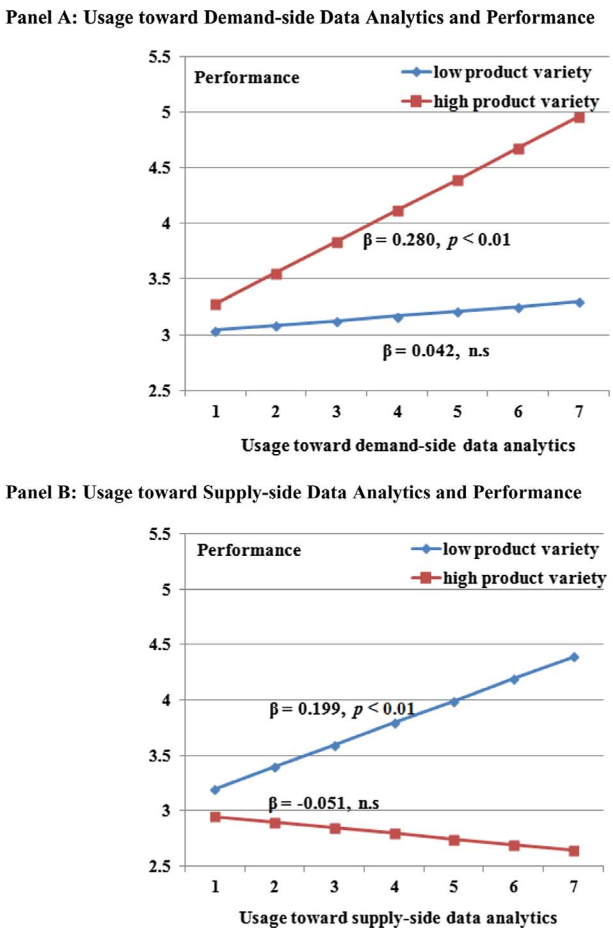

Fig. 2. Simple Slope Analyses for the Moderating Effect of Product Variety.

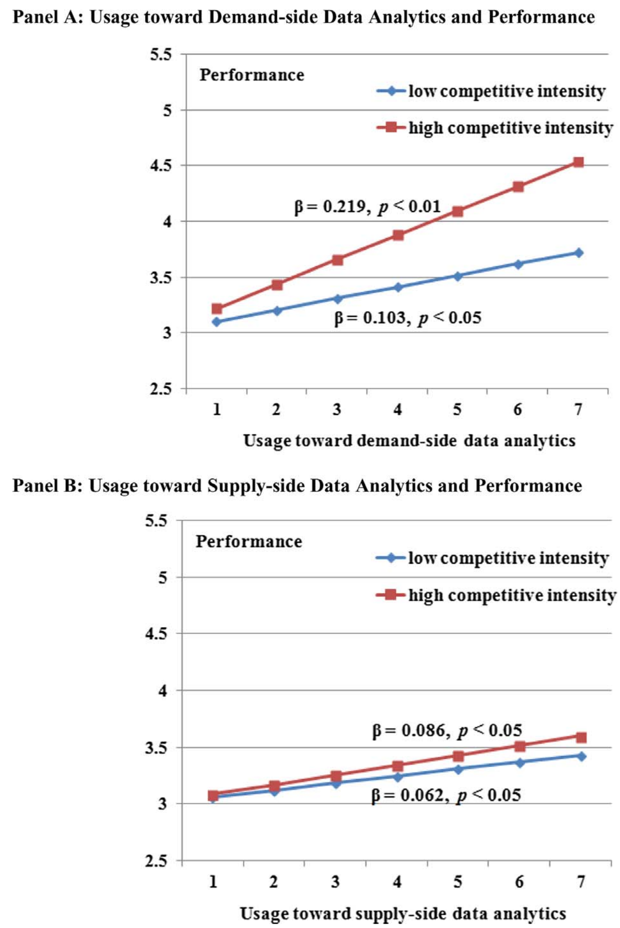

Fig. 3. Simple Slope Analyses for the Moderating Effect of Competitive Intensity.

negative and significant ($\beta = -0.120$, $p < 0.01$). The simple slope analysis implies that when product variety is low (one standard deviation below mean), the coefficient for the main effect of usage toward supply-side data analytics is positive and significant ($\beta = 0.199$, $p < 0.01$); when product variety is high (one standard deviation above mean), this positive effect turns non-significant ($\beta = -0.051$, n.s.) (see Fig. 2, Panel B).

H5 predicts that the influence of usage toward demand-side data analytics on the performance of merchants is stronger for high competitive intensity, and therefore, H5 is also supported. The coefficient for the interaction term between usage toward demand-side data analytics and competitive intensity in model 2 is significant ($\beta = 0.049$, $p < 0.05$). The simple slope analysis shows that when competitive intensity is low (one standard deviation below mean), the coefficient for the usage toward demand-side data analytics is positive and significant ($\beta = 0.103$, $p < 0.05$). Such coefficient becomes higher ($\beta = 0.219$, $p < 0.01$) when competitive intensity is high (one standard deviation above mean) (see Fig. 3, Panel A).

H6 predicts that the influence of usage toward supply-side data analytics on the performance of merchants is stronger for high competitive intensity. H6 is not supported, as the coefficient for the interaction term between usage toward supply-side data analytics and competitive intensity is non-significant ($\beta = 0.001$, n.s.). The simple slope analysis implies that when competitive intensity is low (one standard deviation below mean), the coefficient for the usage toward demand-side data analytics is positive and significant ($\beta = 0.062$, $p < 0.05$). Such coefficient is positive and significant ($\beta = 0.086$, $p < 0.05$) when competitive intensity is high (one standard deviation above mean) (see Fig. 3, Panel B).

5.2. Robustness check

We conducted several additional analyses to verify the robustness of our results. First, we collected objective sales data and used monthly number of sold products as the alternative measurement for the performance of merchants. We log-transformed this variable to address the skewness in distribution. The results, as summarized in Appendix B (in Supplementary material), are consistent with the main results. Specifically, both demand-side and supply-side data analytics usage has a positive effect on the performance of merchants. When product variety is high, the influence of usage toward demand-side data on performance is strengthened, whereas such impact is weakened for supply-side data analytics. When competitive intensity is high, the performance implication of demand-side data analytics usage is strengthened, whereas such impact is not strengthened for supply-side data analytics usage.

Second, to verify the robustness of our analysis, we used the structural equation modeling as the alternative estimation approach. Following the recommendation of Anderson and Gerbing [69], we have tested the measurement model in the above measurement section. After the establishment of the measurement model, the structural model was assessed. The model fitting was satisfactory (RMSEA = 0.035, CFI = 0.97, RFI = 0.96, NFI = 0.95, and GFI = 0.93). The results, as summarized in Appendix C (in Supplementary material), are also consistent with the main results. Most hypothesized paths in the research model were found to be statistically significant. Both demand-side and supply-side data analytics usage has a positive effect on the performance. When product variety is high, the enhancing effect of usage toward demand-side data is strengthened, whereas such impact is weakened for supply-side data analytics. When competitive intensity is high, the enhancing effect of demand-side data analytics usage is strengthened, whereas such strengthened impact is not significant for supply-side data analytics usage.

Finally, data analytics may be more useful in some product categories. We conducted one additional analysis to verify the robustness of our results. We split the sample into different categories. The results based on different categories, as summarized in Appendix D (in Supplementary material), are consistent with the main results. Both demand-side and supply-side data analytics usage has a positive effect on the performance. The enhancing effect of usage toward data analytics on performance is contingent upon the merchants' product variety and competitive intensity.

6. Discussion

With this article, we have sought to examine the effects of data analytics usage on performance and explore how these effects differ in different market conditions. Although both demand-side and supply-side data analytics usage has a positive effect on the performance, such effect is contingent upon the merchants' product variety and competitive intensity. Contrary to our expectation, the influence of usage toward supply-side data analytics on performance is not strengthened by competitive intensity. This may be due to the fact that greater numbers of competing merchants diminish the performance implications of any supply-side operation improvement [56]. Because the market structure becomes more competitive, the performance enhancement due to operational efficiency improvement will be counteracted by the greater numbers of competing merchants. To overcome the negative performance implications of increased competition, merchants must use the supply-side data analytics to improve their operational efficiency to a greater extent. Therefore, the moderating role of competitive intensity in the enhancing effect of usage toward supply-side data analytics is non-significant.

6.1. Theoretical implications

Our study has significant implications for future research. First, this article advances our understanding of business value of data analytics. The value of data has been a long-standing area of inquiry in the economics of IS (e.g., [74]). As organizations increasingly treat data as a primary asset, quantifying the business value of data analytics has become a major discussion topic [27]. However, existing studies mainly focus on the main effect of data analytics usage rather than the factors that can inhibit or enable the performance advantages associated with data analytics usage. We have demonstrated that the value of data analytics usage is contingent upon the merchants' product variety and competitive intensity. We thus clarify the boundary conditions for the business value of data analytics usage and document how this value occurs in tandem with other market conditions.

Second, this study contributes to the implications of large number of merchants in platform markets. Platforms such as eBay and Tmall operate as two-sided markets to attract both merchants and consumers. Platform owners often pursue strategies for aggressively attracting merchants (e.g., subsides and marketing promotions) because a growing number of merchants make the platform more attractive to consumers (and vice versa), thereby giving rise to network effects and winner-take-all outcomes [58,75,76]. As this study shows, when competitive intensity between merchants is high, the performance implication of demand-side data analytics becomes higher. Therefore, a large number of merchants not only give rise to network effects but also increase the value of demand-side data analytics.

Third, our research advances our understanding of the value of product variety in data analytics context. In recent decades, product variety has increased dramatically in most industries. The number of products available in large supermarkets has increased from the order of 1000 in the 1950s to 30,000 in a modern supermarket. Similar developments can be found in most other industries such as the automotive, computer hardware, software, and telecommunication. Although the relationship between a merchant's product variety and its sales is considered to be positive, there are several trade-offs inherent with a product variety strategy [47,49]. High product variety is also likely to reduce merchant performance. Too much variety may lead to consumer confusion or frustration [48,40,37]. We have demonstrated that high product variety can strengthen the enhancing effect of demand-side data analytics usage but weaken the enhancing effect of supply-side data analytics usage. Therefore, our research develops a better understanding of the trade-offs inherent with high product variety.

6.2. Managerial implications

Our study has several important implications for business practice. As a "more is better" view is inaccurate, merchants should adjust their data analytics usage according to their market conditions. Our findings suggest that high product variety can strengthen the enhancing effect of usage toward demand-side data analytics and weaken the enhancing effect of usage toward supply-side data analytics. Therefore, when the product variety is high, merchants should increase their demand-side data analytics usage and decrease their supply-side data analytics usage. In this condition, using data analytics to see consumers' purchases and interactions and listen to consumers' unique wants and needs can generate higher business value. We also find that high competitive intensity can strengthen the enhancing effect of usage toward demand-side data analytics, whereas it does not influence the enhancing effect of usage toward supply-side data analytics. As a result, when the competitive intensity is high, merchants should increase their demand-side data analytics usage and keep the extent of their supply-side data analytics usage. In this situation, using the data analytics to see merchants' internal operational activities and improve their operational efficiencies can generate constant business value.

In addition, platform owners can use our findings to learn how to balance providing demand-side and supply-side data analytics. A firm's fundamental data, IT infrastructure, and know-hows accumulate through the firm's continuous capital, labor, and time inputs in long term and have become firm-specific assets and fundamental determinants of the firm's performance. With accumulated data assets within a transaction platform, many platforms adopt a "freemium" strategy in which they do not charge transaction commissions but rather rely on providing data analytics services [77]. Our empirical evidence indicates that supply-side data analytics is always valuable for merchants regardless of the competitive intensity within a platform. In contrast, demand-side data analytics has higher performance implications for merchants when competitive intensity is high. Therefore, platform owners should provide supply-side data analytics aggressively regardless of the competitive intensity, whereas they should increase the provision demand-side data analytics when the competitive intensity is high.

Finally, enhancing a platform's competitiveness requires a fundamental and complex balancing act between a platform's incentives to keep attracting larger number of merchants and its existing merchants' incentives to improve their performance through data analytics usage. A platform must facilitate and preserve a market environment that stimulates merchants' innovation and the provision of valuable products with better efficiency [56]. This suggests a new focus might be necessary. Instead of crediting platform performance to the size of the merchants alone, it may be prudent to shift our theoretical inquiry to the level of platform market competitiveness. This would reflect the quality dimension for merchants (not just size) and the underlying conflict between incentives for the platform and incentives for its merchants.

6.3. Limitations and further research directions

This study is subject to several limitations and also could be extended in several ways. Our field data are about merchants in a specific platform. Therefore, caution is required before generalizing the findings to other platforms. Contextual factors such as platform-specific characteristics may shape the performance implications of data analytics usage. Future research may extend this study to other contexts and collect contextually rich data to help generalize the insights about data analytics to other B2C platforms (e.g., Amazon) and non-transactional platforms (e.g., Airbnb and Uber). Second, measures of all constructs in this study were collected at one time. However, the investigated constructs were not static. Performance of merchants may have included feedback effects on the data analytics usage. Thus, this cross-sectional study may not fully capture the complexity of the phenomenon, limiting the extent to which causality can be inferred. Future research should use longitudinal studies to investigate the dynamics of data analytics usage.

In conclusion, this study has attempted to reveal the effects of data analytics usage on performance and how they differ in different market conditions. The outcomes are clear: both toward demand-side and supply-side data analytics usage has a positive effect on the performance of merchants. When merchants' product variety is high, the influence of usage toward demand-side data on performance is strengthened, whereas the impact of usage toward supply-side data analytics is weakened. When competitive intensity is high, the performance implication of demand-side data analytics usage is strengthened, whereas such impact is not strengthened for supply-side data analytics.

Acknowledgements

The authors gratefully acknowledge the constructive feedback from the review team and editor, and the financial support of the National Natural Science Foundation of China (# 71472086, 71472083, 71490721 and 71372112). The corresponding author of this publication is Xiaofeng Yu.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.im.2018.01.004.

References

[1] R. Agarwal, V. Dhar, Editorial—big data, data science, and analytics: the opportunity and challenge for IS research, Inf. Syst. Res. 25 (3) (2014) 443–448.
[2] P. Goes, Editor's comments: big data and IS research, MIS Q. 38 (3) (2014) iii–viii.
[3] X. Guo, Q. Wei, G. Chen, J. Zhang, D. Qiao, Extracting representative information on intra-organizational blogging platforms, MIS Q. 41 (4) (2017) 1105–1127.
[4] E. Brynjolfsson, K. McElheran, Data in Action: Data-Driven Decision Making in U.S. Manufacturing, (2016) US Census Bureau Center for Economic Studies Paper No. CES-WP-16-06.
[5] Y.K. Lin, H. Chen, R.A. Brown, S.H. Li, H.J. Yang, Healthcare predictive analytics for risk profiling in chronic care: a Bayesian multitask learning approach, MIS Q. 41 (2) (2017) 473–495.
[6] IBM, The 2011 IBM Tech Trends Report: The Clouds Are Rolling In...Is Your Business Ready? (2011) November 15. http://www.ibm.com/developerworks/techntrendsreport . (Accessed 4 August 2016).
[7] Bloomberg Businessweek, The Current State of Business Analytics: Where Do We Go from Here? Bloomberg Business-week Research Services, 2011, http://www.sas.com/resources/asset/busanalyticsstudy_wp_08232011.pdf.
[8] H. Chen, R. Chiang, V. Storey, Business intelligence and analytics: from big data to big impact, MIS Q. 36 (4) (2012) 1165–1188.
[9] D. Martens, F. Provost, J. Clark, E.J. Fortuny, Mining massive fine-grained behavior data to improve predictive analytics, MIS Q. 40 (4) (2016) 869–888.
[10] A.R. Saboo, V. Kumar, I. Park, Using big data to model time-varying effects for marketing resource (Re)Allocation, MIS Q. 40 (4) (2016) 911–939.
[11] I. Bardhan, C. Oh, E. Zheng, K. Kirksey, Predictive analytics for readmission of patients with congestive heart failure: analysis across multiple hospitals, Inf. Syst. Res. 26 (1) (2015) 19–39.
[12] W.J. Orlikowski, J.A. Yates, K. Okamura, M. Fujimoto, Shaping electronic communication: the metastructuring of technology in the context of use, Organ. Sci. 6 (4) (1995) 423–444.
[13] K. Zhu, K. Kraemer, Post-adoption variations in usage and value of e-business by organizations: cross-country evidence from the retail industry, Inf. Syst. Res. 16 (1) (2005) 61–84.
[14] S. Devaraj, R. Kohli, Performance impacts of information technology: is actual usage the missing link? Manage. Sci. 49 (3) (2003) 273–289.
[15] J.J. Hsieh, A. Rai, X.S. Xu, Extracting business value from IT: a sensemaking perspective of post-adoptive use, Manage. Sci. 57 (11) (2011) 2018–2039.
[16] C. Zhang, L. Xue, J. Dhaliwal, Alignments between the depth and Breadth of interorganizational systems deployment and their impact on firm performance, Inf. Manage. 53 (1) (2016) 79–90.
[17] E. Brynjolfsson, K. Mcelheran, The rapid adoption of data-driven decision-making, Am. Econ. Rev. 106 (5) (2016) 133–139.
[18] S. Sriram, P. Manchanda, M. Esteban-Bravo, J. Chu, L. Ma, M. Song, S. Shriver, U. Subramanian, Platforms: a multiplicity of research opportunities, Market. Lett. 26 (2) (2015) 141–152.
[19] A. Burton-Jones, D.W. Straub, Reconceptualizing system usage: an approach and empirical test, Inf. Syst. Res. 17 (3) (2006) 228–246.
[20] F.D. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology, MIS Q. 13 (3) (1989) 318–339.
[21] V. Venkatesh, M.G. Morris, G.B. Davis, F.D. Davis, User acceptance of information technology: toward a unified view, MIS Q. 27 (3) (2003) 425–478.
[22] V. Dhar, Data science and prediction, Commun. ACM 56 (12) (2013) 64–73.
[23] J. Bughin, J.M. Chui, J. Manyika, Clouds, Big Data, and Smart Assets: Ten Tech-Enabled Business Trends to Watch, McKinsey Quarterly, 2010, pp. 1–14.
[24] S. LaValle, E. Lesser, R. Shockley, M.S. Hopkins, N. Kruschwitz, Big data, analytics and the path from insights to value, MIT Sloan Manage. Rev. 52 (2) (2011) 21–32.
[25] E. Brynjolfsson, T. Geva, S. Reichman, Crowd-squared: amplifying the predictive power of search trend data, MIS Q. 40 (4) (2016) 941–961.
[26] T.H. Davenport, J.G. Harris, R. Morison, Analytics at Work: Smarter Decisions, Better Results, Harvard Business Press, Cambridge, MA, 2010.
[27] A. Abbasi, S. Sarker, R.H.L. Chiang, Editoria—big data research in information systems: toward an inclusive research agenda, J. Assoc. Inf. Syst. 17 (2) (2016) i–xxxii.
[28] M. Wade, J. Hulland, Review: the resource-based view and information systems research: review, extension, and suggestions for future research, MIS Q. 28 (1) (2004) 75–105.
[29] H. Liu, W. Ke, K.K. Wei, H. Chen, The role of institutional pressures and organizational culture in the firm's intention to adopt internet-enabled supply chain management systems, J. Oper. Manage. 28 (5) (2010) 372–384.
[30] H. Liu, S. Wei, W. Ke, K.K. Wei, Z. Hua, The configuration between supply chain integration and information technology competency: a resource orchestration perspective, J. Oper. Manage. 44 (2016) 13–29.
[31] M. Subramani, How do suppliers benefit from IT use in supply chain relationships, MIS Q. 28 (1) (2004) 45–74.
[32] C. Zhang, J. Dhaliwal, An investigation of resource-based and institutional theoretic factors in technology adoption for operations and supply chain management, Int. J. Prod. Econ. 120 (1) (2009) 252–269.
[33] B. Huo, C. Zhang, X. Zhao, The effect of IT and relationship commitment on supply chain coordination: a contingency and configuration approach, Inf. Manage. 52 (6) (2015) 728–740.
[34] U.W. Thonemann, J.R. Bradley, The effect of product variety on supply-chain performance, Eur. J. Oper. Res. 143 (3) (2002) 548–569.
[35] T. Randall, K. Ulrich, Product variety, supply chain structure, and firm performance: analysis of the U. S. bicycle industry, Manage. Sci. 47 (12) (2001) 1588–1604.
[36] Z. Ton, A. Raman, The effect of product variety and inventory levels on retail store sales: a longitudinal study, Prod. Oper. Manage. 19 (5) (2010) 546–560.
[37] X. Wan, M.E. Dresner, P.T. Evers, Assessing the dimensions of product variety on performance: the value of product line and pack size, J. Bus. Logist. 35 (3) (2014) 213–224.
[38] S.P. Anderson, R. Renault, Pricing, product diversity, and search costs: a Bertrand-Chamberlin-Diamond model, RAND J. Econ. (1999) 719–735.
[39] W.W. Moe, Buying, searching, or browsing: differentiating between online shoppers using in-store navigational clickstream, J. Consum. Psychol. 13 (1–2) (2003) 29–39.
[40] C. Huffman, B.E. Kahn, Variety for sale: mass customization or mass confusion? J. Retail. 74 (4) (1998) 491–513.
[41] A. Lambrecht, C. Tucker, Can Big Data Protect a Firm from Competition? (2015) Available at SSRN: https://ssrn.com/abstract=2705530 or https://doi.org/10.2139/ssrn.2705530.
[42] A.L. Montgomery, S. Li, K. Srinivasan, J.C. Liechty, Modeling online browsing and path analysis using clickstream data, Market. Sci. 23 (4) (2004) 579–595.
[43] R.E. Bucklin, C. Sismeiro, A model of web site browsing behavior estimated on clickstream data, J. Market. Res. 40 (3) (2003) 249–267.
[44] P.Y. Chen, L.M. Hitt, Measuring switching costs and the determinants of customer retention in internet-enabled businesses: a study of the online brokerage industry, Inf. Syst. Res. 13 (3) (2002) 255–274.
[45] S.A. Smith, N. Agrawal, Management of multi-item retail inventory systems with demand substitution, Oper. Res. 48 (1) (2000) 50–64.
[46] G. van Ryzin, S. Mahajan, On the relationship between inventory costs and variety benefits in retail assortments, Manage. Sci. 45 (11) (1999) 1496–1509.
[47] P. Boatwright, J.C. Nunes, Reducing assortment: an attribute based approach, J. Market. 65 (3) (2001) 50–63.
[48] S.M. Broniarczyk, W.D. Hoyer, L. McAlister, Consumers' perception of the assortment offered in a grocery category: the impact of item reduction, J. Market. Res. 35 (2) (1998) 166–176.

[49] L. Sloot, D. Fok, P.C. Verhoef, The short- and long-term impact of an assortment reduction on category sales, J. Market. Res. 43 (4) (2006) 536–548.
[50] McKinsey & Company, The Age of Analytics: Competing in a Data-Driven World, (2016) https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world . (Accessed 4 October 2017).
[51] Y. Wu, G. Frizelle, J. Efstathiou, A study on the cost of operational complexity in customer-supplier systems, Int. J. Prod. Econ. 106 (1) (2007) 217–229.
[52] B.L. Bayus, W.P. Putsis, Product proliferation: an empirical analysis of product line determinants and market outcomes, Market. Sci. 18 (2) (1999) 137–153.
[53] X. Wan, P.T. Evers, M.E. Dresner, Too much of a good thing: the impact of product variety on operations and sales performance, J. Oper. Manage. 30 (4) (2012) 316–324.
[54] M. Armstrong, Competition in two-Sided markets, Rand J. Econ. 37 (3) (2006) 668–691.
[55] M. Armstrong, J. Wright, Two-sided markets, competitive bottle-necks and exclusive contracts, Econ. Theory 32 (2) (2007) 353–380.
[56] C. Cennamo, J. Santalo, Platform competition: strategic trade-offs in platform markets, Strateg. Manage. J. 34 (11) (2013) 1331–1350.
[57] D.B. Yoffie, M. Kwak, With friends like these: the art of managing complementors, Harv. Bus. Rev. 84 (9) (2006) 88–98.
[58] C.W.L. Hill, Establishing a standard: competitive strategy and technological standards in winner-take-all industries, Acad. Manage. Exec. 11 (2) (1997) 7–25.
[59] S.F. Turner, W. Mitchell, R.A. Bettis, Responding to rivals and complements: how market concentration shapes generational product innovation strategy, Organ. Sci. 21 (4) (2010) 854–872.
[60] K.J. Boudreau, L.B. Jeppesen, Unpaid crowd complementors: the platform network effect mirage, Strateg. Manage. J. 36 (12) (2015) 1761–1777.
[61] B.J. Jaworski, A.K. Kohli, Market orientation: antecedents and consequences, J. Market. 57 (3) (1993) 53–70.
[62] S. Auh, B. Menguc, Balancing exploration and exploitation: the moderating role of competitive intensity, J. Bus. Res. 58 (12) (2005) 1652–1661.
[63] V. Yeniaras, S. Unver, Revisiting the mediating effect of entrepreneurial behaviour on proactiveness–performance relationship: the role of business ties and competitive intensity, Eur. Manage. Rev. 13 (4) (2016) 291–306.
[64] Alibaba, Alibaba Group's Financial Report in Fiscal Year 2016, (2016) http://www.alibabagroup.com/en/ir/pdf/form20F_160525.pdf . (Accessed 8 March 2017).
[65] S.J. Armstrong, T.S. Overton, Estimating nonresponse bias in mail surveys, J. Market. Res. 14 (3) (1977) 396–402.
[66] S.S. Kim, N.K. Malhotra, A longitudinal model of continued IS use: an integrative view of four mechanisms underlying post-adoption phenomena, Manage. Sci. 51 (5) (2005) 741–755.
[67] K. Lancaster, The economics of product variety: a survey, Market. Sci. 9 (3) (1990) 189–206.
[68] H. Li, K. Atuahene-Gima, Product innovation strategy and the performance of new technology ventures in China, Acad. Manage. J. 44 (6) (2001) 1123–1134.
[69] J.C. Anderson, D.W. Gerbing, Structural equation modeling in practice: a review and recommended two-step approach, Psychol. Bull. 103 (3) (1988) 411–423.
[70] P.M. Podsakoff, D.W. Organ, Self-reported in organizational research: problems and prospects, J. Manage. 12 (4) (1986) 531–544.
[71] P.M. Bentler, D.G. Bonnett, Significance tests and goodness of fit in the analysis of covariance structures, Psychol. Bull. 88 (3) (1980) 588–606.
[72] P.M. Podsakoff, S.B. Mackenzie, J.Y. Lee, N.P. Podskaoff, Common method biased in behavioural research: a critical review of the literature and recommended remedies, J. Appl. Psychol. 88 (5) (2003) 879–903.
[73] H. Liang, N. Saraf, Q. Hu, Y. Xue, Assimilation of enterprise systems: the effect of institutional pressures and the mediating role of top management, MIS Q. 31 (1) (2007) 59–87.
[74] R.D. Banker, R.J. Kauffman, 50th anniversary article: the evolution of research on information systems: a fiftieth-year survey of the literature in management science, Manage. Sci. 50 (3) (2004) 281–298.
[75] K.J. Boudreau, Let a thousand flowers bloom? An early look at large numbers of software app developers and patterns of innovation, Organ. Sci. 23 (5) (2012) 1409–1427.
[76] P. Song, L. Xue, A. Rai, C. Zhang, The ecosystem of software platform: a study of asymmetric cross-side network effects and platform governance, MIS Q. 42 (1) (2017) 121–142.
[77] E. Fang, X. Li, M. Huang, R.W. Palmatier, Direct and indirect effects of buyers and merchants on search advertising revenues in business-to-business electronic platforms, J. Market. Res. 52 (3) (2015) 407–422.
[78] C. Fornell, D.F. Larcker, Evaluating structural models with unobserved variables and measurement errors, J. Marketing Res. 18 (1) (1981) 39–50.

Peijian Song is an associate professor at the Department of Marketing and Electronic Business, School of Business, Nanjing University. He received his doctoral degree in management science from Fudan University. His research interests include technology innovation adoption and diffusion, and platform ecosystems. His work has been published or accepted in journals such as MIS Quarterly, Decision Support Systems, IEEE Transactions on Engineering Management, and Journal of International Marketing.

Chengde Zheng is a professor at the Department of Marketing and Electronic Business, School of Business, Nanjing University. He received his doctoral degree in management science from Nanjing University of Science and Technology. His research interests include e-commerce, and supply chain management.

Zhang, Cheng is a professor at Department of Information Management & Information Systems, Fudan University, China. His current research interests include information strategy, electronic markets and IT diffusion. His works have been published by journals including MIS Quarterly, Journal of Management Information Systems, INFORMS Journal on Computing, Journal of Marketing, Marketing Science, Information & Management, Decision Support Systems, IEEE Transactions on Engineering Management, IEEE Transactions on Professional Communication, International Journal of Electronic Commerce, International Journal of Production Economics, International Journal of Production Research, Journal of the American Society for Information Science and Technology, Journal of Global Information Management etc. He is serving in the editorial review board of Journal of Global Information Management and IEEE Transactions on Professional Communication.

Xiaofeng Yu is an associate professor at the Department of Marketing and Electronic Business, School of Business, Nanjing University. He received his doctoral degree in computer science from Nanjing University. His research interests include economics of e-commerce market, and online consumer behavior.