A Three-essay Dissertation on Big Data Analytics Value Creation for Organizations and Their Supply Chains

by

Suning Zhu

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 15, 2018

Keywords: big data analytics initiative, big data analytics capability,
supply chain management, construct development, system generalized method of moments

Approved by

Casey Cegielski, Chair, J.W. Woodruff Sr. Professor in Information Systems Management
David Paradice, Harbert Eminent Scholar in Business Analytics
Robert "Glenn" Richey, Harbert Eminent Scholar and Professor in Supply Chain Management
Kang Bok Lee, Assistant Professor in Business Analytics
Pei Xu, Assistant Professor in Information Systems Management
James H. Long, Associate Professor in Accounting

Abstract

This dissertation consists of three essays. The first essay conceptualizes BDA capability in supply chain context based on two aspects: the level of analytics and the operational functions of supply chain. Content analysis technique was adopted to analyze existing academic and practical articles concerning BDA in supply chain management (SCM). A rigorous inductive approach was employed to synthesize the 129 articles and develop the data structure of BDA capability in SCM. The proposed data structure includes four aggregate BDA capability dimensions, twenty-two BDA capability constructs, and the measures of each construct. The findings of this study expand the current view of BDA in supply chain context and ground new empirical research in this field.

Following the construct development and validation procedures proposed by Mckenzie, Podsakoff, and Podsakoff (2011), the second essay focuses on developing and validating a comprehensive instrument for measuring BDA capability in the supply chain domain. Building on the results from the first essay, BDA capability in SCM was developed into 22 first-order constructs that formed 4 second-order constructs. Measurement items were created to measure each first-order construct. After conducting face validity and content validity check, a set of data (n=137) was collected from supply chain practitioners to evaluate scale properties and refine measurement items. This study provides a comprehensive and detailed conceptualization of an instrument for BDA capability in SCM that can serve as a springboard for future empirical research to understand the antecedents and impacts of BDA capability on supply chains. Industry practitioners may adopt this instrument to evaluate their BDA capabilities and identify the capabilities they lack.

The third essay is a longitudinal study on how organizations' business analytics initiatives influence operational efficiency and business growth. Drawing upon dynamic capability and contingency theory, I conceptualize organizational BDA initiatives as a dynamic information processing capability which will bring competitive advantage to organizations. Additionally, industry factors (i.e. dynamism, munificence, and complexity) will moderate the relationship between BA initiatives and organizational performance. To test the research model, I collected secondary data from Lexis/Nexus and COMPUSTAT databases and constructed two dynamic panel data models. Using system generalized method of moments (system GMM), I found that: organizational BDA initiatives enhance operational efficiency and facilitate business growth; at lower level of industry dynamism and munificence, BDA initiatives have a greater impact on operational efficiency; at higher level of industry dynamism and complexity, BDA initiatives are associated with greater increase in business growth. These findings provide a theory-based understanding about the economic benefits of BDA and also offer guidance regarding what practitioners can expect from BDA initiatives and how firms can realize value from BDA given the characteristics of industries they are operating in.

Acknowledgments

I am very grateful to all the following individuals for the help and support they have given me to get through numerous difficult moments to achieve my dreams:

To my parents, Tiefeng Zhu and Youli Zhang, who valued education, respected my opinions, and offered me freedom and assistance so that I could pursue my dreams without any concerns. Without your love and support, none of this would have been possible. I love you very much!

To my best friend, Drew Borders, who helped me proofread and modify papers and job application documents and prepare for job interviews. I am very lucky to meet you in a foreign country and I love you very much!

To my family members and my friends, who encouraged me, mentored me, believed in me, and cheered me on every step of the way.

To my dissertation chair, Dr. Casey Cegielski, by whose wise and patient guidance I was able to navigate in the sea of the Doctoral Program and achieve my academic dreams. I am also immensely grateful to my committee members, Dr. David Paradice, Dr. Glenn Richey, Dr. Kang Lee, and Dr. Pei Xu whose professional guidance was essential in completing this dissertation. Also, I would like to thank Dr. James Long for serving as my committee's Outside Reader.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BDA	Big Data Analytics
BDBA	Big Data Business Analytics
SCM	Supply Chain Management
DPB	Data Science, Predictive Analytics, and Big Data
PLM	Product Lifecycle Management
RFID	Radio Frequency Identification
IT	Information Technology
IS	Information Systems
IoT	Internet of Things
LSCM	Logistics and Supply Chain Management
SM-SCM	Service and Manufacturing Supply Chain Management
SC	Supply Chain
IVC	Information Value Chain
GPS	Global Positioning System
AVE	Average Variance Extracted
TC	Tracking Capability
AC	Analytical Capability
PC	Predictive Capability
DSC	Decision Support Capability
ACD	Analytical Capability in Distribution

ACPC	Analytical Capability in Procurement
ACPDD	Analytical Capability in Product Design and Development
ACPD	Analytical Capability in Production
ACT	Analytical Capability in Transportation
DSCDM	Decision Support Capability in Demand Management
DSCD	Decision Support Capability in Distribution
DSCPDD	Decision Support Capability in Product Design and Development
DSCPC	Decision Support Capability in Procurement
DSCPD	Decision Support Capability in Production
DSCSCN	Decision Support Capability in Supply Chain Design and Development
DSCT	Decision Support Capability in Transportation
PCDM	Predictive Capability in Demand Management
PCD	Predictive Capability in Distribution
PCPDD	Predictive Capability in Product Design and Development
PCPC	Predictive Capability in Procurement
PCPD	Predictive Capability in Production
PCT	Predictive Capability in Transportation
TCD	Tracking Capability in Distribution
TCPC	Tracking Capability in Procurement
TCPD	Tracking Capability in Production
TCT	Tracking Capability in Transportation
VIF	Variance Inflation Factor
SFE	Stochastic Frontier Estimation

DDP	Dynamic Data Panel
LSDV	Least Square Dummy Variables
GMM	Generalized Method of Moments

Essay I

Understanding Big Data Analytics Capability in Supply Chain Management

1. Introduction

The widespread deployment of digital technologies (e.g. RFID, Internet of Things, etc.) at the periphery of enterprise supply chain networks has dramatically enhanced the volume, variety, and velocity of data, often described via the term “big data”. The emergence of big data poses a challenge for organizations because it complicates the identification and extraction of useful insights for managing the supply chain (Kache and Seuring, 2017). Big data analytics (BDA) emerges in this respect to provide organizations with better means for obtaining insights from massive amounts of data (Chen, Chiang, and Storey, 2012). It appears all along the supply chain decision spectrum–from assessing supplier risks to optimizing supply chain inventories to enabling more accurate demand forecast (Sanders, 2016; Wang, Gunasekaran, Ngai, and Papadopoulos, 2016). As a valuable decision-making asset, BDA can deliver competitive advantages by improving operational efficiency and maintenance, supply chain visibility and transparency, supply chain responsiveness, integration and collaboration, and much more (Kache and Seuring, 2017).

Due to the perceived value of BDA in supply chain, organizations make great investment in BDA resources in the forms of tangible resources (e.g. data and advanced technologies), human resources (e.g. big data-skilled employees), and intangible resources (e.g. data-driven culture) (Gupta and George, 2016). However, not much is known about how to generate business value from the invested BDA resources (Fosso Wamba, Gunasekaran, Akter, Coltman, and Ngai, 2015). Information Technology (IT) capability building view has emphasized that, in order to transform IT resources into valuable outputs, organizations have to build capabilities by selecting and deploying resources and assembling them into synergetic combinations (Karimi, Somers, and Bhattacherjee, 2007; Weill and Vitale, 2002). Therefore, the first step for exploring BDA value creation process for supply chains is to understand what unique and idiosyncratic BDA capabilities can be created through deploying and implementing BDA resources across the entire supply chain decision spectrum.

Extant research on BDA in supply chain management (SCM) has primarily focused on conducting systematic literature reviews (Wang et al., 2016), investigating BDA applications in SCM (Hahn and Packowski, 2015), proposing BDA methodologies that can be applied to supply chain issues (Zhong, Xu, Chen, and Huang, 2017), assessing business value of BDA (Brinch, 2018), and identifying barriers and opportunities of BDA (Kache and Seuring, 2017; Schoenherr and Speier-Pero, 2015). Little is known about the profile of BDA capability in supply chain domain. The purpose of this study is to enhance the understanding of BDA in SCM by answering the following research question: *How to conceptualize big data analytics capability in supply chain management?* The conceptualization of BDA capability is based on the level of analytics and operational functions of supply chain. Content analysis technique was employed to existing BDA and SCM academic and practitioner articles. A rigorous inductive coding procedure was adopted to synthesize the 129 articles to identify the pertinent BDA capability constructs and assimilated measures.

Following this introduction, the article continues with a literature review section to identify the research gaps in BDA in SCM and define BDA from a capability perspective. Section 3 describes the methods used to collect and analyze qualitative data. Section 4 presents the profile of BDA capability in SCM, including its constructs and associated measures. Section

5 discusses the theoretical and practical contributions of this study. Finally, section 6 points out several limitations of this research and proposes future research directions.

2. Literature Review

2.1 Current Research on Big Data Analytics in Supply Chain Management

Research on BDA in SCM is still in its infancy (Wang et al., 2016; Kache and Seuring, 2017). A review of literature (see Table 1) at the nexus of BDA and SCM reveals five research streams.

The first research stream focuses on exploratory studies regarding BDA adoption (Kache and Seuring, 2017; Richey, Morgan, Lindsey-Hall, and Adams, 2016; Schoenherr and Speier-Pero, 2015) and implementation (Sanders, 2016) in supply chain context. In this stream, researchers used exploratory approaches (e.g. Delphi research technique, case studies, etc.) to identify opportunities and challenges relating to BDA adoption in SCM (Kache and Seruing, 2017) as well as to propose a framework about how to implement BDA to drive a firm's supply chain (Sanders, 2016).

The second research stream is regarding the application areas of BDA in SCM (Hahn and Packowski, 2015; Li et al., 2015; Matthias, Fouweather, Gregory, and Vernon, 2017; Souza, 2014; Zhong, Newman, Huang, and Lan, 2016b). For example, Souza (2014) discussed the applications of advanced analytics in SCM along two dimensions–the level of analytics and operational functions of supply chain. Hahn and Packowski (2015) proposed a comprehensive framework that categorizes in-memory analytics in SCM along use cases and function domain.

Studies of the third research stream perform literature reviews of available BDA research in supply chain (Addo-Tenkorang and Helo, 2016; Nguyen, Zhou, Spiegler, and Ieromonachou, and Lin, 2017; Wang et al., 2016), identify research gaps, and propose new research directions in this area (Hazen, Boone, Ezell, and Jones-Farmer, 2014; Hazen, Skipper, Ezell, and Boone, 2016; Waller and Fawcett, 2013). For instance, Nguyen et al. (2017) classified literature on BDA in SCM based on the operational function of supply chain, level of analytics, BDA model used, and BDA techniques used to develop BDA models. Through the classification, they revealed research gaps and suggested future directions for research development in BDA in SCM. Wang, Gunasekaran, Ngai, and Papadopoulos (2016) reviewed and classified literature on big data business analytics (BDBA) in logistics and supply chain management (LSCM) along two aspects: the level of analytics and the focus of LSCM. Moreover, they proposed a maturity framework of BDBA in LSCM to assess the extent to which BDBA is applied within LSCM.

The fourth research stream is concerned with specific BDA methodologies applied to SCM. Studies in this stream primarily aim to develop and test a big data methodology that leverages big data captured from internal and external sources to solve supply chain issues. Existing works have developed BDA methodologies to address issues in new product development (Tan, Xhan, Ji, Ye, and Chang, 2015), logistics management on the shop floors (Zhong, Huang, Lan, Dai, Chen, and Zhang, 2015; Zhong, Lan, Xu, Dai, and Huang, 2016a; Zhong et al., 2017), inventory management (Kartal, Oztekin, Gunasekaran, and Cebi, 2016), and transport logistics (Kim, Kim, and Park, 2016).

The last research stream falls in BDA value creation in supply chain context. Studies in this stream conceptualize BDA in SCM and examine the relationships between BDA dimensions and performance outcomes (Chae, Olson, and Sheu, 2014a; Chae, Yang, Olson, and Sheu, 2014b; Chen, Preston, and Swink 2015; De Oliveira, McCormack, and Trkman 2012; Trkman et al., 2010; Zhu, Song, Hazen, Lee, and Cegielski, 2018). For example, Chen, Preston, and Swink (2015) studied the underlying value creation mechanism of BDA usage in SCM by identifying the antecedents of BDA usage and testing the role of BDA usage on firm performance. Drawing on the resource-based view, Chae, Olson, and Sheu (2014) conceptualized supply chain analytics as composed of three resources (data management resources, IT-enabled planning resources, and performance management resources) and investigated the impact of the three resources on operational performance.

Although studies in the aforementioned research streams all link BDA with supply chain context, few research (except Arunachalam, Kumar, and Kawalek (2017)) discusses BDA capability in supply chain context. Most academic studies explore BDA in SCM from an application/usage perspective (the first, second, and fifth research streams) or a technical perspective (the fourth research stream), but little is known about the capability profile of BDA in SCM. Information Technology (IT) literature has emphasized the importance of IT capability. The capability profile of IT can lead to improved organizational capabilities (Muller, et al., 2010; Pavlou and ElSawy, 2010) which finally generate business values (Muller et al., 2010; Pavlou and ElSawy, 2010; Rai, Pavlou, Im, and Du, 2012). The findings of IT capability indicate that it is the capability of BDA in SCM that transforms data into business value. Therefore, a comprehensive investigation is required to understand BDA capability profile in SCM and further to formulate proper strategies for developing BDA capability in SCM which eventually creates business values.

Table 1. Prior Studies on Big Data Analytics in Supply Chain Management

Research Stream	Author (Year)	Study Type	Main Contributions
Stream I: Opportunities and challenges of BDA adoption and implementation.	Kache and Seuring (2017)	Empirical (Exploratory approach)	Identified 43 opportunities and challenges related to BDA from a corporate and a supply chain perspective using Delphi research technique.
	Richey et al. (2016)	Empirical (Exploratory Approach)	Adopted the native category approach to • Develop an industry grounded definition of Big Data in the supply chain setting

			along four dimensions: volume, velocity, variety, and veracity. • Uncover Big Data key success factors • Identify barriers to developing the potential of Big Data
	Sanders (2016)	Empirical (Exploratory Approach)	• Uncovered the distinct differences of today's BDA capability • Identified the characteristics of successful BDA implementation in supply chain domains from cases of leading companies. • Provided a framework of how to proceed BDA implementation to drive a firm's supply chain.
	Schoenherr and Speier-Pero (2015)	Empirical (Exploratory Approach)	• Provided an assessment of the current state of SCM predictive analytics adoption. • Identified the motivations to use SCM predictive analytics. • Identified the benefits of and barriers to the use of SCM predictive analytics. • Provided insights into desired skills for successful data scientists and how research universities should train next-generation data scientists.
Stream II: Application areas of BDA in SCM	Hahn and Packowski (2015)	Empirical (Exploratory Approach)	Provided a comprehensive framework of emerging in-memory analytics applications in SCM along two dimensions: use cases (i.e. monitor-and-navigate, sense-and-respond, predict-and-act, and plan-and-optimize), and functional domain (i.e. operations management, sales management, and integrated business management).
	Li et al. (2015)	Conceptual	Provided a comprehensive and systematic framework of existing applications of Big Data in product lifecycle management (PLM) as well as potential applications of Big Data in PLM.
	Matthias et al. (2017)	Empirical (Exploratory Approach)	• Developed a framework for categorizing application areas of big data in operations management. • Empirically demonstrated how the use of big data in two of the application areas helps improve operational performance.
	Souza (2014)	Conceptual	Described the applications of advanced

			analytics techniques in supply chain management along two dimensions: level of analytics (i.e. descriptive, predictive, and prescriptive analytics), and key operational functions of supply chain (i.e. plan, source, make, deliver, and return).
	Zhong et al. (2016b)	Empirical (Exploratory Approach)	• Presented representative examples of big data applications in service and manufacturing supply chain management (SM-SCM). • Reviewed big data technologies and models used for decision-making in SM-SCM. • Reviewed the current movements world-wide on big data in SM-SCM. • Highlighted the challenges, opportunities, and future perspectives on big data in SM-SCM.
Stream III: Systematic literature review of available research on BDA in SCM.	Addo-Tenkorang and Helo (2016)	Conceptual (Literature Review)	• Classified available research on BDA in operations/supply chain management in terms of five main attributes of big data (i.e. variety, velocity, volume, veracity, and value). • Proposed “big data II” (Internet of Things – Value-adding) framework.
	Arunachalam et al. (2017)	Conceptual (Literature Review)	• Classified literature on BDA in supply chain from a capability perspective into five dimensions: data generation capability, data integration and management capability, advanced analytics capability, data visualization capability, and data-driven culture. • Provided a maturity model for BDA capabilities in the supply chain context. • Identified organizational and technical challenges of practicing BDA in supply chain.
	Hazen et al. (2014)	Conceptual	• Reviewed literature on data quality from the perspective of data science, predictive analytics, and big data (DPB) in SCM. • Introduced the application of statistical process control as a means to monitor and control data quality in the context of DPB in SCM. • Proposed three theories that can be

			leveraged to explore research in the context of DPB in SCM.
	Hazen et al. (2016)	Conceptual	Proposed a theoretical-driven research agenda that can be leveraged to inform future research on how BDA impacts sustainable SCM outcomes.
	Nguyen et al. (2017)	Conceptual (Literature Review)	• Classified literature on BDA in SCM based on (1) the operational functions of SC in which BDA is applied, (2) the level of analytics, (3) the types of BDA models used in SCM, and (4) the BDA techniques used to develop BDA models. • Revealed research gaps and suggested future directions for the research development of BDA in SCM.
	Waller and Fawcett (2013)	Conceptual	Proposed possible research opportunities at the nexus of supply chain management and data science, predictive analytics, and big data.
	Wang et al. (2016)	Conceptual (Literature Review)	• Classified research on big data business analytics in logistics and supply chain management (LSCM) based on the level of analytics (i.e. descriptive, predictive, and prescriptive analytics) and the focus of LSCM (i.e. operations and strategy). • Proposed a maturity framework of BDBA in LSCM.
Stream IV: Specific BDA methodologies applied to SCM.	Kartal et al. (2016)	Empirical	Developed a hybrid methodology that combines machine learning algorithms with multi-criteria decision-making techniques to effectively address multi-attribute inventory classification problems.
	Kim et al. (2016)	Empirical	Proposed a data-driven method for early or real-time detection of vessel delays using the combination of real-time vessel tracking data and historical shipping data.
	Tan et al. (2015)	Empirical	Developed and tested an analytic infrastructure that combines data mining and deduction graph techniques to harvest big data and suggest the optimal expanding process of incorporating a firm’s competence sets with others.
	Ou Cheng, Chen, and Perng (2016)	Empirical	Established a dynamic cost model to accurately forecast manufacturing costs of finished products and to help decision makers adjust purchasing and pricing strategy.

	Zhong et al. (2015)	Empirical	Proposed a big data approach to excavate frequent trajectory from RFID-enabled logistics big data collected from production shop floors for decision-making such as logistics planning and scheduling.
	Zhong et al. (2016a)	Empirical	• Introduced an approach to integrate Internet of Things (IoT) and Cloud manufacturing to enable intelligent manufacturing environment. • Proposed a RFID-Cuboid model to address the visualization of complex and abstract radio frequency identification (RFID) data generated on production shop floors. • Reported a case study to show the feasibility and practicability of the proposed visualization approach.
	Zhong et al. (2017)	Empirical	Developed a BDA framework for processing RFID big data collected from physical Internet-based production shop floors to visualize logistics trajectory and to assess logistics operators and operations based on defined behaviors and key performance indicators.
Stream V: BDA value creation in SCM.	Chae and Olson (2013)	Conceptual	Proposed a framework of business analytics for supply chain (i.e. supply chain analytics) as IT-enabled, analytical dynamic capabilities composed of data management capability, analytical supply chain process capability, and supply chain performance management capability.
	Chae et al. (2014a)	Empirical (Survey-based Approach)	• Defined the architecture of supply chain analytics as composed of three sets of resources: data management resources, IT-enabled planning resources, and performance management resources. • Provided a better understanding of the effect of supply chain analytics on supply chain planning satisfaction and operational performance.
	Chae et al. (2014b)	Empirical (Survey-based Approach)	Provided a better understanding of the impact of two specific business analytics resources (i.e. accurate manufacturing data and advanced analytics) on firms' operational performance.
	Chen et al.	Empirical	Provided a better understanding of the

	(2015)	(Survey-based Approach)	underlying mechanisms of organizations' big data analytics usage in supply chain management. Specifically, • the role of organizational BDA usage in SCM in value creation • the antecedents of enterprise-level BDA usage.
	De Oliveira et al. (2012)	Empirical (Survey-based Approach)	Provided a better understanding of the impact of business analytics on supply chain performance at different process maturity levels.
	Trkman et al. (2010)	Empirical (Survey-based Approach)	• Defined business analytics in supply chain as an integration of analytical capabilities in key supply chain operational functions (i.e. plan, source, make, and deliver) • Provided a better understanding of the relationship between business analytics capabilities in key operational functions of supply chain and supply chain performance.
	Zhu et al. (2018)	Empirical (Survey-based approach)	Provided a better understanding of analytics capability in SC on operational supply chain transparency under different levels of supply uncertainty.

2.2 A Capability-Set Perspective of Big Data Analytics in Supply Chain Management

There is no unified definition on big data analytics. Holsapple, Lee-Post, and Pakath (2014) assembled the published views of analytics in the business field and organized its definitions into six categories–a movement, a collection of practices and technologies, a transformation process, a capability set, specific activities, and a decisional paradigm. In this study, I adopted the capability-set perspective that views big data analytics as a set of capabilities and correspondingly named it big data analytics capability in supply chain management. IT literature has suggested that the capability profile of IT can lead to improved organizational capabilities (Muller, et al., 2010; Pavlou and ElSawy, 2010) which finally generate business values (Muller et al., 2010; Pavlou and ElSawy, 2010; Rai et al., 2012). Studies on BDA also argued that it is the capability set of BDA that influences an organization's process transformation which eventually results in superior organizational performance (Wamba et al., 2017; Wang and Hajli, 2017; Wang, Kung, Wang, and Cegielski, 2017). An investigation of big data analytics capability profile in the supply chain domain will allow practitioners and researchers to clarify how BDA transforms supply chain to realize its economic potential.

There is no definition on big data analytics capability specific to the supply chain management context. However, broad definitions of big data analytics capability have been developed from different views. In general, big data analytics capability refers to the ability to manage large quantities of disparate data to allow users to conduct data analysis and reactions (Hurwitz, Nugent, and Hapler, 2013). Wixom, Yen, and Relich (2013) emphasized that big data analytics capabilities for maximizing organizations' business value need to include (1) speed to insights which is the ability to expeditiously transform raw data into usable information and (2) pervasive use which is the extent to which business analytics is used across the enterprise. Through the lens of big data analytics adoption, LaValle, Lesser, Shockley, Hopkins, and Kruschwitz (2011) classified analytics capability into three levels: aspirational, experienced, and transformed. The first level focuses on efficiency of existing processes and cost management. The second level centers on preparing for the optimization of organizations. The third level aims to drive customer profitability and make targeted investments in niche analytics. Drawing on resource-based view, Wamba et al. (2017) defined big data analytics capability as the ability to provide business insights using data management, infrastructure flexibility, and personnel capability to transform business into a competitive force. Based on the same view, Gupta and George (2016) described big data analytics capability as an organization's ability to assemble, integrate, and deploy its big data-specific resources. Besides the broad definitions of big data analytics capability, Wang et al. (2017) adopted the information lifecycle management view and defined BDA capability in the context of healthcare as the ability to acquire, store, process, and analyze large amounts of health data in various forms, and provide meaningful information for users to discover insights in a timely fashion.

In this study, BDA capability is defined through the view of information value chain (IVC). Abbasi, Sarker, and Chiang (2016) described IVC as a cyclical set of activities necessary to convert data into information and, subsequently, to transform information into knowledge that decision makers can use to make decisions and take actions. IVC encompasses two groups of activities: knowledge derivation and decision making. The concept of IVC can help us understand the role of big data analytics in facilitating organizations' knowledge derivation and decision making (Goes, 2014; Sharma, Mithas, and Kankanhalli, 2014). As such, through the IVC view, BDA capability in SCM is defined as

the ability of an organization to convert high volume, variety, and velocity of data generated along supply chain networks into meaningful knowledge, which helps decision makers make decisions and take actions to solve supply chain issues in a timely fashion.

3. Research Method

To achieve the goals of this study, I conduct a content-analysis based review of existing academic and practitioner publications. Big data analytics in supply chain context has undergone a transition from a new topic to a growing research and application area, with an increasing number of articles being published in academic journals and practitioner periodicals (Mishra, Gunasekaran, Papadopoulos, and Childe, 2016; Schoenherr and Speier-Pero, 2015). Therefore, there are a considerable number of articles upon which to ground a content analysis to extract big data analytics capabilities in supply chain management. The article collection, selection, and analysis procedures are described in the following subsections.

3.1 Data Collection and Selection

Academic journal databases (Web of Science and ABI/INFORM complete) as well as major supply chain trading periodicals (e.g. Supply Chain Management Review, Mckinsey Quarterly, Logistics Management) were explored to capture articles at the nexus of BDA and SCM published between 2007 and 2017. The following keywords were used both separately and in combination (using "AND"/ "OR") for the first-round data selection: big data, analytics, supply chain, logistics, operations management, demand management, supply management, procurement, sourcing, production, manufacturing, distribution, inventory, warehouse, and transportation. The inclusion criteria were that all the articles should be written in English and available for full-text download. This search generated 1,578 articles from academic journals and 1,800 articles from trading periodicals. Next, I removed the duplicates, read the titles and abstracts, and kept the ones that focus on BDA within supply chain domain. The second-round selection lead to 408 and 110 articles from trading periodicals and academic journals respectively. Since BDA capability in SCM is reflected through organizations gaining knowledge from supply chain data collected from heterogeneous systems distributed across organizational boundaries, selected articles for content analysis should fully/partially discuss how BDA enables the knowledge derivation which further benefits the supply chain. Thus, in the third-round selection, I carefully read the full text of each article and kept the ones having the entire article or specific sections concerning BDA value creation process for supply chain. Irrelevant articles such as pure BDA techniques and methodologies were excluded. The final dataset contained 19 academic articles and 110 practitioner papers.

3.2 Data Analysis Procedures

Content analysis technique is an empirically grounded method for extracting themes and topics from text (Krippendorff, 2012). It can help with the identification of key aspects or attributes of a construct's domain required in the conceptualization step (Mackenzie, Podsakoff, and Podsakoff, 2011). To ensure a better understanding of BDA capability in SCM, a three-phase content analysis (i.e. preparation, organizing, and reporting) proposed by Elo and Kyngäs (2008) was conducted for each selected article. Moreover, this study adopted inductive content analysis because there is not enough former knowledge on the application of BDA within SCM (Fosso Wamba, Angappa, Papadopoulos, and Ngai, 2018; Wang et al., 2016).

3.2.1 Preparation Phase

The goal of the preparation phase is to understand the coding process, in terms of the selection of unit of analysis, the level of analysis, and the purpose of evaluation (Elo and Kyngäs, 2008). "Themes" were selected as the unit of analysis, which primarily express an idea that can be sentences, paragraphs, or a portion of a page (Krippendorff, 2012). The level of analysis is related to BDA within the scope of SCM. The purpose of evaluation is to identify specific aspects of BDA capability in SCM.

3.2.2 Organizing Phase

The second phase is to organize qualitative data. This is the phase through which all the aspects of BDA capability are identified and conceptualized (Elo and Kyngäs, 2008). I initially read each article several times and highlighted statements related to how analytics can transform big data into useful insights to solve supply chain problems. A total of 428 statements were obtained and copied and pasted into the spreadsheet, which served as the basis for further analysis. The selected statements were then sent to the second coder, a senior professor whose research focuses on the intersection of supply chain and big data. In order to process the highlighted statements, Strauss and Corbin's (1998) open, axial, and selective coding procedures were followed to identify conceptually similar aspects. Open coding refers to "the analytical process through which concepts are identified and their properties and dimensions are discovered in data" (Strauss and Corbin, 1998, p.101). Axial coding is "the process of relating categories to their subcategories, termed 'axial' because coding occurs around the axis of a category, linking categories at the level of properties and dimensions" (Strauss and Corbin, 1998, p.123). Selective coding is "the process of integrating and refining the theory" (Strauss and Corbin, 1998, p.143). Following this approach, both coders (i.e. the senior professor and I) independently coded each statement and organized the codes in a Microsoft Excel spreadsheet.

In open coding, all the 428 statements were analyzed using Strauss and Corbin's (1998) line-by-line analysis. Statements found to express similar topics were grouped together and summarized into concepts (i.e. categories). By "concepts", I mean "a more general, less well-specified notion capturing qualities that describe or explain a phenomenon of theoretical interest" (Gioia, Corley, and Hamilton, 2013). Both coders classified statements expressing similar topics into concepts that describe various aspects of BDA capability in supply chain domain. A comprehensive compendium of concepts were developed in this process. In axial coding, the coders started seeking similarities and differences among the many concepts. Similar concepts were organized into a theory-centric theme (i.e. construct) that suggests the content of each category and moves beyond description to a higher level of abstraction (Urquhart, Lehmann, and Myers, 2010). The axial coding process eventually reduced the sheer number of categories into a more manageable number of themes. As the example displayed in Appendix A, a passage from Briest, Dilda, and Somers (2014) described that big data analytics allows chemical manufacturers to model the effect of various operating conditions on production yields through advanced mathematical modeling. This passage was coded as “predict the effects of variations in operating conditions on production performance” in the open coding. Then, the theme “Predictive Capability in the Production Process” was created during axial coding to move the concept to a more abstract level. In the selective coding, the coders focused on finalizing all the concepts and themes through comparing, integrating, and refining similar codes emerged during open and axial coding.

Once the coders completed all the coding procedures, they met and compared the coding outcomes. The coding reliability was calculated using the reliability index developed by Perreault and Leigh (1989). The interrater reliability was 0.88, which was higher than the convention of 0.7 (Ryan and Bernard, 2000). Most disagreements occurred between the two coders were on where to put the concepts that describe how BDA solves issues in inventory management, material handling, warehousing, and transportation. A third researcher with more than 15 years of research experience in IT and supply chain management was invited to facilitate the discussion of the disagreement in order to reach an agreement. The three researchers re-analyzed each disputed statement and eventually reached a consensus: combine concepts describing BDA in inventory management, material handling, and warehousing to formulate BDA capability in distribution constructs (i.e. tracking capability in distribution, predictive capability in distribution, analytical capability in distribution, and decision support capability in distribution); relate the concepts describing BDA application in transportation to BDA capability in transportation construct (i.e. tracking capability in transportation, predictive capability in transportation, analytical capability in transportation, and decision support capability in transportation). In the end, 22 theory-centric themes (i.e. constructs) were obtained to constitute BDA capability in SCM.

After finalizing a set of concepts and themes, I investigated whether it is possible to distill the emergent themes even further into aggregate dimensions. Constructs sharing similar essential characteristics should be theoretically abstracted to a higher level (Hoehle and Venkatesh, 2015; Mackenzie et al., 2011). To identify aggregate dimensions, I examined how distinctive the constructs were from each other and whether eliminating any of them would restrain the domain of the construct in a significant way (Mackenzie et al., 2011). For example, tracking capability in procurement, production, distribution, and transportation share the common focus that an organization is able to access output data from all the systems or devices distributed across organizational boundaries. Moreover, eliminating one of these constructs will restrict the domain of tracking capability in a significant way because they represent four distinct processes of a supply chain where tracking capability is reflected. Therefore, the four constructs were aggregated into tracking capability. Ultimately, BDA in SCM can be understood through the four aggregate dimensions–tracking capability, analytical capability, predictive capability, and decision support capability.

3.2.3 Reporting Phase

Based on the concepts, themes, and aggregate dimensions, I built a data structure (Figure 1) to show how I progressed from raw qualitative data to concepts, themes, and dimensions through a rigorous inductive approach (Gioia et al., 2013). In addition, the 22 constructs related to tracking, analytical, predictive, and decision support capabilities were defined by drawing on the selected articles for content analysis and were presented in Table 2.

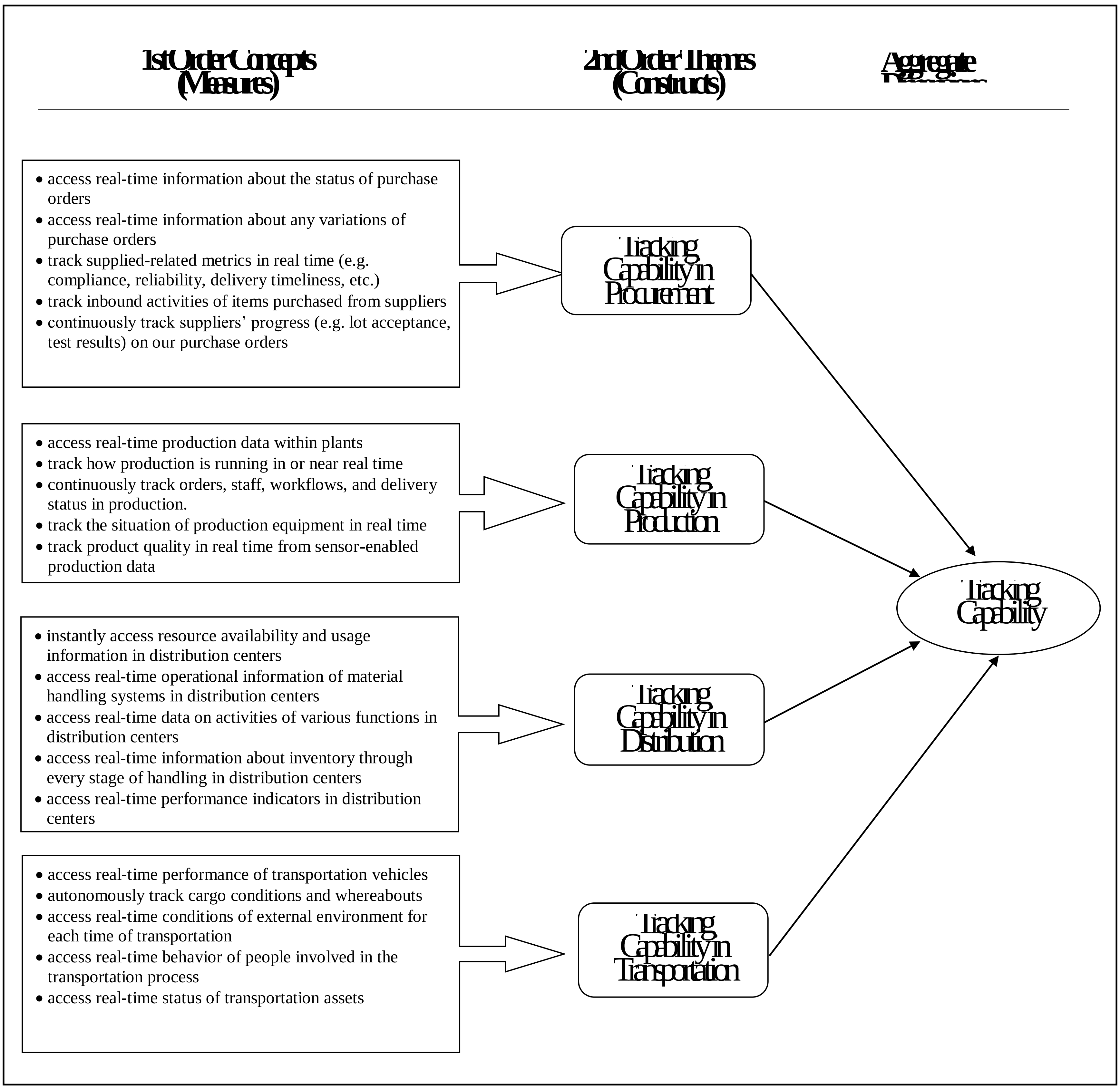

Figure 1. Data Structure

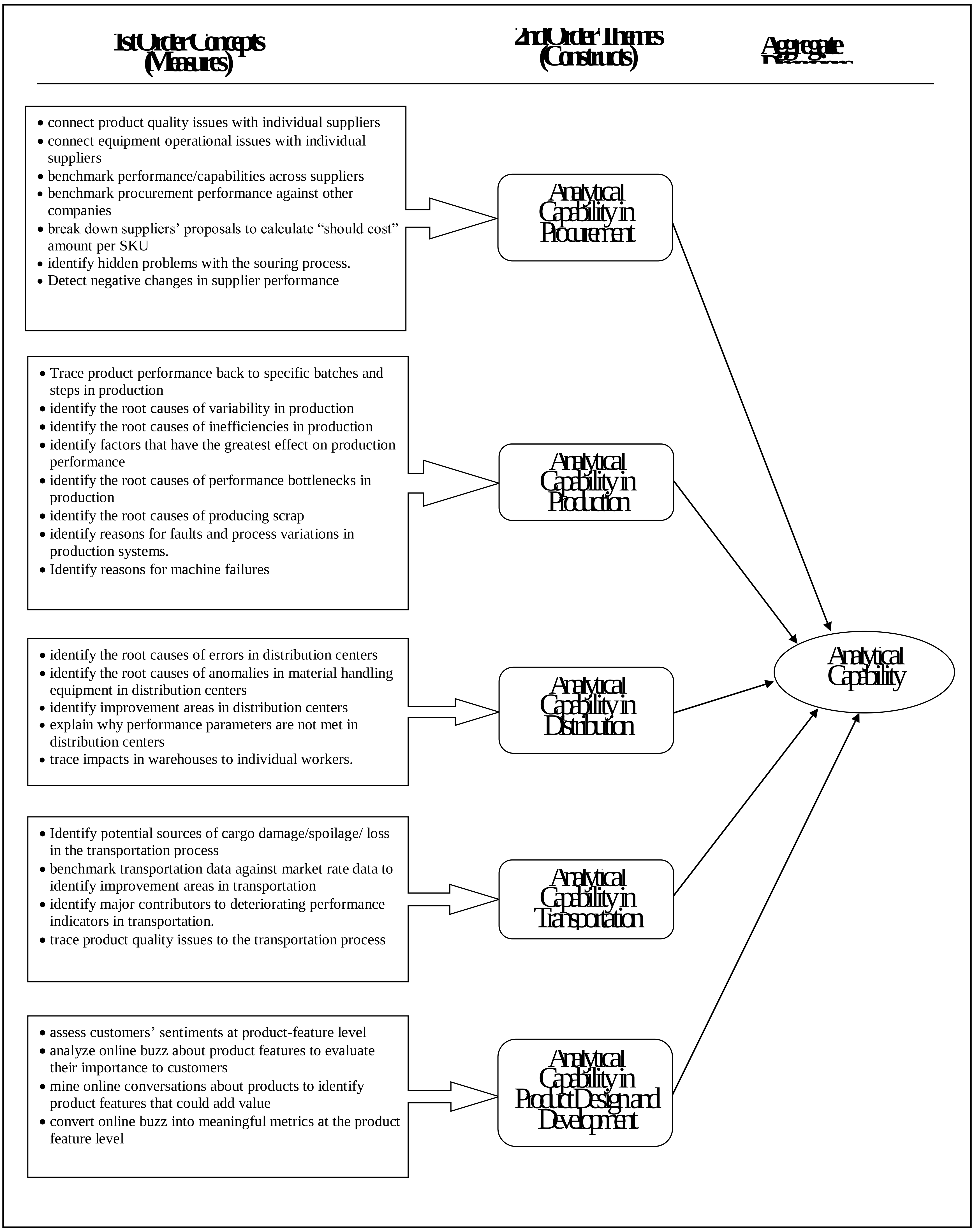

Figure 1. Data Structure (Continued)

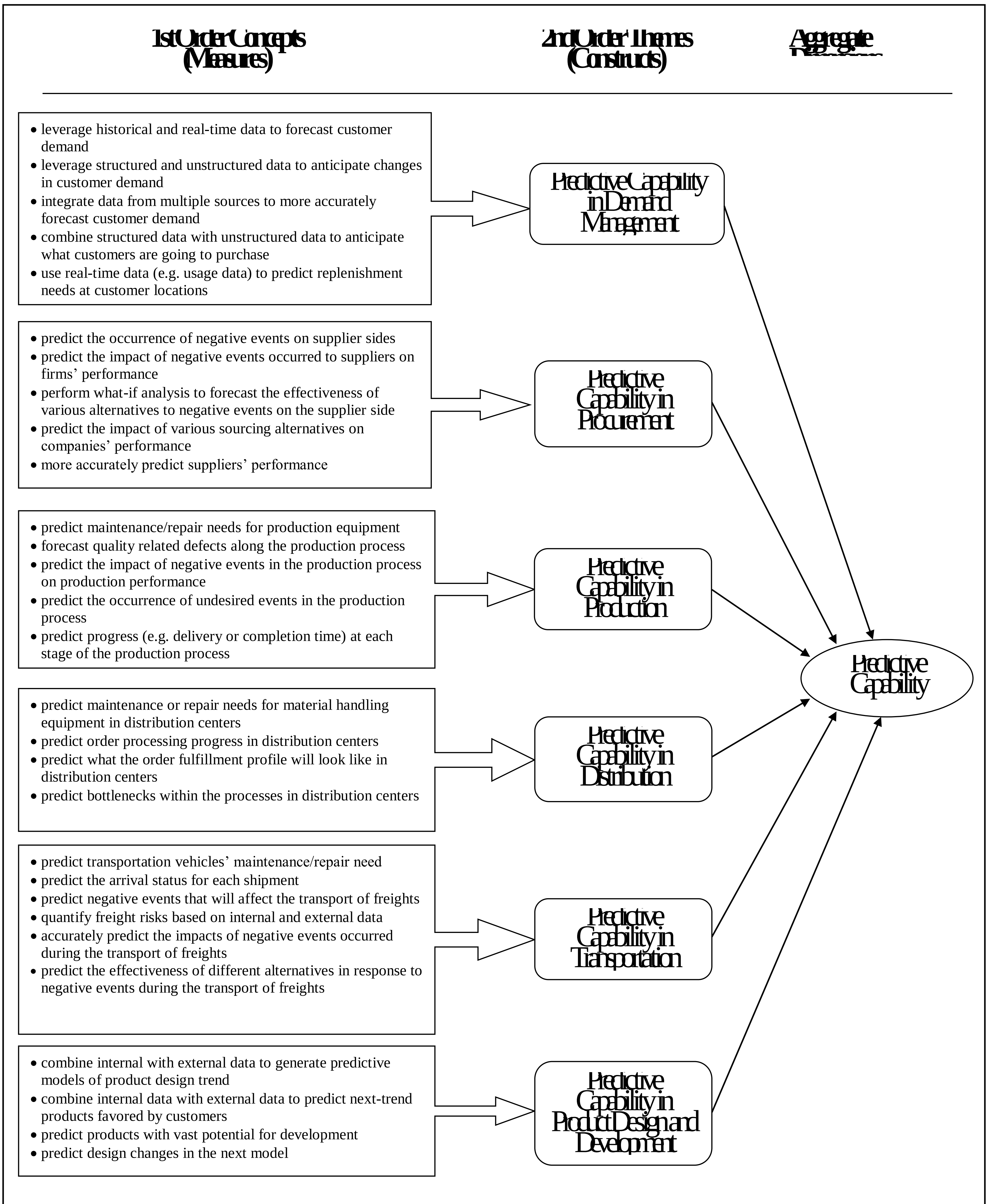

Figure 1. Data Structure (Continued)

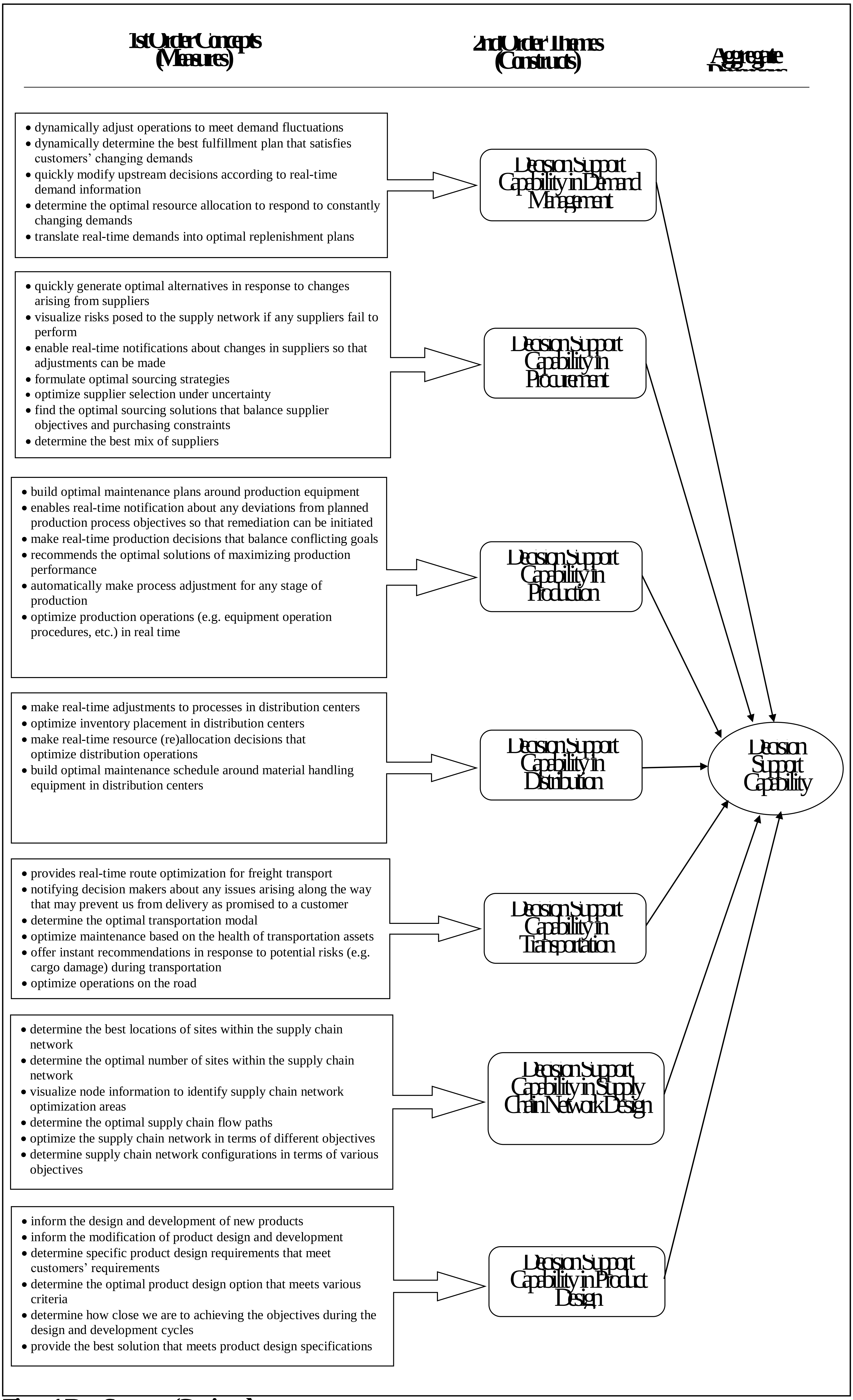

Figure 1. Data Structure (Continued)

Table 2. Constructs of Big Data Analytics in Supply Chain Management

Construct	Definition
Tracking Capability in Procurement	The ability of an organization to access current and past situations regarding purchased orders and suppliers in the supply base.
Tracking Capability in Production	The ability of an organization to access current and past situations of working orders, operators, machinery, and work flows along the production line.
Tracking Capability in Distribution	The ability of an organization to access current and past situations of inventories, workers, equipment, and work flows in distribution centers.
Tracking Capability in Transportation	The ability of an organization to access current and past situations regarding cargos, transportation assets, staff, and surroundings during transportation from supply points to demand points.
Analytical Capability in Procurement	The ability of an organization to examine historical and real-time supply chain data to discover the root causes of undesirable situations in the supply base.
Analytical Capability in Production	The ability of an organization to examine historical and real-time supply chain data to discover the root causes of undesirable situations along the production line.
Analytical Capability in Distribution	The ability of an organization to examine historical and real-time supply chain data to discover the root causes of undesirable situations in distribution centers.
Analytical Capability in Transportation	The ability of an organization to examine historical and real-time supply chain data to discover the root causes of undesirable situations during transportation.
Analytical Capability in Product Design and Development	The ability of an organization to examine historical and real-time supply chain data to discover the contributing factors to successful products.
Predictive Capability in Demand Management	The ability of an organization to discover patterns from historical and real-time supply chain data to forecast customers' demands.
Predictive Capability in Procurement	The ability of an organization to discover patterns from historical and real-time supply chain data to forecast what is likely to happen to its supply base as well as the corresponding impacts.
Predictive Capability in Production	The ability of an organization to discover patterns from historical and real-time supply chain data to forecast what is likely to happen to the production line and the corresponding impacts.
Predictive Capability in Distribution	The ability of an organization to discover patterns from historical and real-time supply chain data to forecast what is likely to happen to order processing in distribution centers.
Predictive Capability in Product Design and	The ability of an organization to discover patterns from historical and real-time supply chain data to forecast customer

Development	preferences for new/next-generation products.
Predictive Capability in Transportation	The ability of an organization to discover patterns from historical and real-time supply chain data to forecast what is likely to happen during transportation.
Decision Support Capability in Demand Management	The ability of an organization to generate the optimal demand execution plan in order to balance customer demand forecast.
Decision Support Capability in Procurement	The ability of an organization to generate the optimal course of actions to guide supplier selection and supply risk mitigation.
Decision Support Capability in Production	The ability of an organization to generate the optimal course of actions to guide adjustment and optimization decisions during production.
Decision Support Capability in Distribution	The ability of an organization to generate the optimal course of actions to guide adjustment and optimization decisions in distribution centers.
Decision Support Capability in Transportation	The ability of an organization to generate the optimal course of actions to guide adjustment and optimization decisions during transportation.
Decision Support Capability in Supply Chain Network Design	The ability of an organization to generate the optimal course of actions to guide supply chain network optimization decisions.
Decision Support Capability in Product Design and Development	The ability of an organization to generate the optimal course of actions to guide product design and development decisions.

4. Results and Discussion

4.1 The Profile of BDA Capability in Supply Chain Management

This section presents discussion around BDA capabilities identified through content analysis. A conceptual framework (Figure 2) is developed to simplify and delineate the capabilities identified. The vertical axis shows four levels of analytics which mirror the way that BDA capabilities (i.e. tracking, analytical, predictive, and decision support capabilities) are realized. The horizontal axis denotes the operational functions of supply chain where BDA is applied. The framework delivers a detailed view of BDA capability in SCM and it can be argued that organizations need to possess all the capabilities in Figure 2 in order to gain value from raw data. Organizations can use this framework to see where they are in building analytics capability for supply chains.

4.1.1 Tracking Capability

Tracking capability is the ability of an organization to access real-time output data from all the systems and devices distributed across organizational boundaries in order to understand the current and past situation of its supply chain. This capability is realized through integrating descriptive analytics with big supply chain data. As organizations increasingly adopt emerging technologies such as sensors, RFID chips, and GPS, huge amounts of data are generated from these devices in real time. Applying descriptive analytics on sensor, RFID, and GPS data provides real-time information regarding the location and conditions of a company's purchased materials, inventories, machinery and so forth in the supply chain (Souza, 2014). Accessing such real-time information allows organizations to monitor processes and identify anomalies and opportunities for their supply chains (Hahn & Packowski, 2015; Tiwari, Wee, and Daryanto, 2018; Wang et al., 2016).

In the procurement process, BDA allows organizations to access real-time information (e.g. live location data streams) on the status of global shipments from suppliers (Bentz, 2014; Leal, 2015), any variations during transit (Bentz, 2014), and the progress of supplied items at suppliers' sites (Fathi, 2011). In the production process, BDA enables organizations to access live data about the status of processes, and the states of production equipment, parts, and tools, product quality, material utilization, production environment, operators and so forth (Garrett, 2016a, Lade, Ghosh, and Srinivasan, 2017, Li et al., 2015, Steppeler, 2013).

Level of Analytics	Demand Management	Procurement	Production	Distribution	Transportation	Product Design and Development	Supply Chain Network Design
Prescriptive Analytics	Decision Support Capability in Demand Management	Decision Support Capability in Procurement	Decision Support Capability in Production	Decision Support Capability in Distribution	Decision Support Capability in Transportation	Decision Support Capability in Product Design and Development	Decision Support Capability in SC Network Design
Predictive Analytics	Predictive Capability in Demand Management	Predictive Capability in Procurement	Predictive Capability in Production	Predictive Capability in Distribution	Predictive Capability in Transportation	Predictive Capability in Product Design and Development	
Diagnostic Analytics		Analytical Capability in Procurement	Analytical Capability in Production	Analytical Capability in Distribution	Analytical Capability in Transportation	Analytical Capability in Product Design and Development	
Descriptive Analytics		Tracking Capability in Procurement	Tracking Capability in Production	Tracking Capability in Distribution	Tracking Capability in Transportation		

Supply Chain Functions

Figure 2. Conceptual Framework for Big Data Capability in Supply Chain Management

With BDA in the distribution process, organizations are able to access real-time information on inventory quantity and availability (Waller and Fawcett, 2013) used on each job (Karlskind, 2014). BDA also provides real-time data on the conditions of components of material handling systems (McCrea, 2014; Partridge, 2014; Trebilcock, 2013) and workers' operational activities within or around a warehouse/distribution center (Douglas, 2014, Michel, 2016, Patridge, 2014, Trebilcock, 2013). Applying BDA during transportation, organizations can track real-time status of transportation assets such as vehicles, trailers, and containers (Abbott, 2017; Chui, Loffler, and Roberts, 2010; Mani, Delgado, Hazen, and Patel, 2017; Trent, 2017), drivers' behavior including speed, breaks, and brakes (Mani et al., 2017; Trent, 2017), and cargo's information on whereabouts and conditions (Ittmann, 2015; Li et al., 2015; Sanders, 2016; Woods, 2017).

4.1.2 Analytical Capability

Analytical capability is the ability of an organization to dig into the past and current supply chain data in order to discover the underlying causes of undesirable situations or events and identify improvement areas. This BDA capability is formed through applying diagnostic analytics on big supply chain data. Once the anomalies impacting the supply chain are identified through BDA tracking capability, exploratory data analysis of existing data (or additional data if required) is implemented using such tools as drill-down, visualization, data discovery, and data mining so as to understand the possible reasons for the occurrences (Banerjee, Bandyopadhyay, and Acharya, 2013; Delen and Zolbanin, 2018). Determining the factors contributing to the outcomes enables directional guidance for making improvements and reacting to problems in the supply chain.

With big data analytics applied to the procurement process, organizations are able to process all forms of supplier data to have a comprehensive view of supplier performance (Richey et al., 2016) and to identify opportunities of bargaining and negotiating with suppliers (Garrett, 2016b; Khan, 2013; Richey et al., 2016). Organizations also leverage BDA to analyze real-time operational data of items purchased from suppliers to pinpoint issues and improvement areas for suppliers' products and services (Chidambaram, Evans, and Entheredge, 2015). In the production process, BDA enables organizations to process large volumes of structured, semi-structured, and unstructured production data generated in real time to spot production problems (Aronow, Burkett, Romano, and Nilles, 2016; Babiceanu and Seker, 2016; Wang et al., 2016; Zhong et al., 2015), and identify root causes for the problems (Hahn and Packowski, 2015; Lade et al., 2017). In the distribution process, BDA allows organizations to look into historical and real-time data of operational activities at warehouse or distribution center level to spot anomalies (Michel, 2016; Partridge, 2014; Trebilcock, 2013) and to find out corresponding root causes (Partridge, 2014; Trebilcock, 2013). In the transportation process, BDA enables organizations to process historical and real-time operation data to evaluate delivery performance (e.g. schedule reliability, item damage) (Douglas, 2014), identify inappropriate driving behavior (Field, 2014; Mani et al., 2017; Szakonyi, 2014), and detect abnormal and inefficient use of transportation assets (Mani et al., 2017; Szakonyi, 2014; Trebilcock, 2014). Applying BDA in the product design and development process, organizations are able to analyze abundant online buzz, measure customers' sentiments at product-feature level, appraise the importance of product features to customers, and identify valuable features (Fedewa, Velarde, and O'Neill, 2016).

4.1.3 Predictive Capability

Predictive Capability refers to the ability of an organization to discover patterns from real-time and historical supply chain data and extrapolate those patterns to project future events, behavior, and outcomes through the supply chain network. This capability is enabled by analyzing big supply chain data using predictive analytics such as data mining, text mining, web mining, and statistical forecasting (Demirkan and Delen, 2013). Building on the understanding of current and past situations through descriptive and diagnostic analytics (Delen and Zolbanin, 2018), predictive analytics looks at all the possible future scenarios within demand planning, procurement, production, and much more. BDA predictive capability helps organizations get out in front of disruptions and events and respond proactively given current and past conditions (Banerjee et al., 2013).

In the demand management process, BDA leverages unstructured data (e.g. engine search data, weather data) to anticipate customers' demand changes at increasingly granular level (Catlin, Scanlan, and Willmott, 2015; Shenkman, Johnson, and Elliot, 2016; Trent, 2017). BDA combines internal data (e.g. purchase history) with external data (e.g. conversations on social media) to more accurately predict customer demands (Brown, Chui, and Manyika, 2011; Hahn and Packowski, 2015; Li et al., 2015; Trent, 2017). BDA looks into historical and real-time data to model future demand trends in real time (Arrow et al., 2016; Sanders, 2016). In procurement, BDA helps organizations integrate internal information with external information to better predict adverse risk events that will disrupt the supply (Mitchell, 2012; Trebilcock, 2015; Trent, 2017). BDA enables procurement team to perform predictive forecasting to understand the impact of supply risk events (Garrett, 2016d; Garrett, 2017; Wang et al., 2016) and the effectiveness of various alternatives in response to the events (Bentz, 2014). BDA is also a powerful tool in forecasting the impact of different sourcing or purchasing alternatives (e.g.

supplier mix, purchase volume) (Boone, Skipper, and Hazen, 2017; Garrett, 2016d). In production, BDA capitalizes on real-time data from sensors deployed on production equipment to precisely predict when equipment will break down or need maintenance (Bughin, Chui, and Manyika, 2015; Garrett, 2016a; Suhas, 2017). BDA helps predict defects during production by detecting emerging negative trends in product quality (Lade et al., 2017; Li et al., 2015). BDA also uses real-time operational data to predict undesired events during production (Babiceanu and Seker, 2016; Dhawan, Singh, and Tuteja, 2014; Kache and Seuring, 2017) and forecast the effects of any variations or disruptions on production performance (Briest et al., 2014; Chae and Olson, 2013; Dhawan et al., 2014). In the distribution process, BDA is able to predict problems of material handling equipment before equipment breaks down based on real-time operational data from sensors attached to the equipment (McCrea, 2014; Trebilcock, 2013). It also helps predict inventory future usage (Hadavi, 2016), the fulfillment profile at the warehouse or distribution center level (Michel, 2016; Partridge, 2014), and when and where an issue may occur to the operations in a warehouse or distribution center (Trebilcock, 2013). In transportation, BDA extracts trends from real-time operational data of transportation vehicles to precisely predict mechanical issues and maintenance needs (Trent, 2017). BDA integrates data from various sources (e.g. radar systems, sensors, social media) to more precisely predict the arrival time of shipments (Borkowsky and Walther, 2014; Kache and Seuring, 2017, Palmquist and Leal, 2016), the negative events (e.g. traffic congestion) that will affect shipping (Field, 2014; Trent, 2017; Wright, 2016), as well as the downstream impacts of such events (Field, 2014). In product design and development, BDA combines internal corporate data with other relevant data from external sources to generate predictive models of design trends (Borkowsky and Walther, 2014; Giannakis and Louis, 2016) and forecast customer preferences for future products (Tan et al., 2015).

4.1.4 Decision Support Capability

Decision support capability is the ability of an organization to determine the optimal course of action for improving supply chain performance, given a complex set of objectives, requirements, and constraints (Demirkan and Delen, 2013; Lustig, Dietrich, Johnson, and Dziekan, 2010; Wang et al., 2016). This capability is achieved through extracting insights from big supply chain data using prescriptive analytics such as simulation and optimization (Banerjee et al., 2013). BDA decision support capability goes beyond predictive, analytical, and tracking capabilities by recommending a set of possible courses of action. Decision makers will better understand the implications of each suggested action by running different scenarios, eventually selecting the action that best impacts the entire company (Viswanathan, 2018). Companies possessing BDA decision support capability are able to successfully optimize supply chain operations, making sure that they are delivering the right product to the right customer at the right time.

In demand management, BDA helps organizations dynamically determine the best way to manage upstream operations based on real-time customers' demands (Dirlea, Ng, Gutierrez, and See, 2016; Garrett, 2016c; Sanders, 2016; Shenkman et al., 2016). For example, BDA recommends the optimal level of inventory across all the SKUs at every level of the supply chain, in responding to constantly changing customer demands (Dirlea et al., 2016; Havadi, 2016; Pearson, 2011; Wang et al., 2016). BDA more precisely models supply and manufacturing capacity (Garrett, 2016d), distribution capacity (Ittmann, 2015), and shipping capacity (Levans, 2011; Michel, 2016) to support forecasted demands. BDA also provides dynamic pricing recommendations to adjust real-time customer demands (Chae and Olson, 2013; Catlin et al., 2015; Garrett, 2016c; Godwin, 2016). In procurement, BDA optimizes decision making related to sourcing, supplier selection and supply risk management. Specifically, BDA leverages massive amounts of data from various sources (e.g. supply markets, suppliers, regulation institutions, procurement function) to determine the optimal sourcing solution (Sanders, 2016; Wang et al., 2016) that for example maximizes the total value between suppliers and procurement (Khan, 2013). Suppliers are selected under uncertainty in most cases. BDA helps organizations find the best supplier or combination of suppliers under uncertainty based on large amounts of supplier data and other influence factors (Li, Tao, Cheng, and Zhao, 2015). In managing risks from the supply side, BDA generates the best alternative or the best combination of alternatives that minimize the impact of supply risks (Bentz, 2014; Mitchell, 2012; Trebilcock, 2015).

In production, BDA generates insights from all available data along production lines to automatically make adjustment decisions that optimize production (Babiceanu and Seker, 2016; Brown et al., 2011) or send notifications to decision makers so that appropriate adjustments can be made (Ward and Gopal, 2014) for any stage of the manufacturing operations. For instance, BDA provides real-time operations centers with the analysis results of the status of production equipment and mechanical systems in order to create optimal maintenance plans to minimize equipment downtime (Brown et al., 2011; Li et al., 2015) and increase the overall performance of the production environment (Kache and Seuring, 2017). In distribution, BDA supports decisions on resource re-allocation through generating insights into material handling and warehousing operations. For example, BDA provides insights into real-time performance and usage of warehouse workforce through which organizations can make labor reallocation decisions on the fly that boost productivity (Michel, 2016; Partridge, 2014). In transportation, BDA supports decisions about shipping modal selection, driver behavior management, fuel efficiency, cargo safety, transportation assets maintenance, and routing. For example, BDA suggests real-time route optimization, taking into account numerous factors such as weather, traffic, location, etc. (Boone et al., 2017; Ittmann, 2015; Waller and Fawcett, 2013; Wang et al., 2016). BDA generates insights into the health of transportation assets which help managers make informed decisions on maintenance (Abbott, 2017; Chui et al., 2010). In product design and development,

BDA provides insights into product functionality and usage which inform the modification of new product models (Bughin et al., 2015). BDA delivers appropriate solutions (e.g. the most economical design) that meet design specification in the conceptual design phase (Li et al., 2015). BDA helps organizations determine specific design requirements, for example, what is required of the parts and the reliability of the materials used in the parts (Wang et al., 2016). In supply chain network design, BDA scrutinizes massive amounts of data across supply chain nodes to drive supply chain network configuration decisions that solve for various objectives (Finley, Blaeser, and Djavairian, 2015; Levans, 2015; Wang et al., 2016). Examples of BDA decision support capability in supply chain network design include determining the right number and location of distribution centers (Bentz, 2014), how products should flow through the supply chain (Levan, 2015; Watson, 2012), and the optimal shipping quantities between locations for minimizing the sum of various cost components (Wang et al., 2016).

5. Implications

5.1 Theoretical Implications

This study makes significant contributions to both theory and practice. First, the findings of this study contribute to the ongoing discussion about BDA in supply chain context by providing insights into different dimensions of BDA capability through a rigorous inductive approach. Extant literature on BDA in SCM has predominantly focused on defining the domain (Waller and Fawcett, 2013; Richey et al.,2016), investigating BDA applications in SCM (Hahn and Packowski, 2015; Wang et al., 2016), building a roadmap to implement BDA (Sanders, 2016), proposing BDA methodologies applied to supply chain issues (Tan et al., 2015; Zhong et al., 2017), assessing the business value of BDA in SCM (Brinch, 2018; Chen et al., 2015), and identifying barriers and opportunities (Kache and Seuring, 2017; Schoenherr and Speier-Pero, 2015). This study builds on the contributions of extant literature and develops the profile of BDA capability in SCM and the embedded construct measures, which expands and enriches the current view on BDA in SCM and moves the understanding of this concept towards a more detailed and granular level.

Additionally, the identification of BDA capability constructs and measures opens a new path for empirical research regarding the business value of BDA in supply chain context. Numerous research agendas call for research on the impacts of BDA on organizations (Loebbecke and Picot, 2015; Sharmar et al., 2014). Although a few studies have empirically assessed the effects of BDA on organizational performance from a BDA capability perspective (e.g. Côrte-Real, Oliveira, and Ruivo, 2017; Fosso Wamba, Gunasekaran, Akter, Ren, and Dubey, 2017), none of the BDA capability models are tailored for supply chain context. Moreover, a few studies have examined analytics capabilities in supply chain context (Chae et al., 2014a; Trkman et al., 2010), however, the supply chain analytics capability measures in those studies are not customized for big data environment in which more advanced analytics technology is needed to deal with large volume, various, and real-time or near real-time data (Liu and Yi, 2017). This study is among the first to understand BDA capability in supply chain context. The conceptualization of BDA capability along with the construct measures will ground new empirical research concerning how to understand the value of BDA in SCM.

5.2 Practical Implications

This study brings some valuable insights to practice. First of all, the comprehension about big data in supply chain context differs tremendously among practitioners (Richey et al., 2016). SCM practitioners may employ the BDA capability framework developed by this study to agree on a unified understanding of big data in SCM from a capability perspective. Moreover, a major concern regarding BDA is that employees on the frontlines are reluctant with using BDA despite considerable investments in this cutting-edge technology (Barton and Court, 2012). The reason for the concern is that front-line employees do not understand or trust BDA. Nevertheless, the accidental return on BDA investments won't materialize unless employees at all levels understand BDA and include it into decision making (Shah, Horne, and Capellá, 2012). Therefore, in order to gain value from BDA, it is necessary for organizations to help employees at all levels have a clear and coherent understanding about BDA which can be achieved from training (Zhu, Gupta, Paradice, and Cegielski, 2018). Since BDA training courses are still in their infancy (Waller and Fawcett, 2013), the framework of BDA capability in SCM may serve as the first step for BDA training by providing all the employees with a preliminary and comprehensive understanding on the capabilities of BDA across organizational boundaries.

The first step towards building an organizational capability is to self-assess the organization's own strengths and weaknesses (Bharadwaj, 2000). SCM practitioners may adopt the BDA capability constructs and measures to appraise their organizations' BDA capability across supply chain processes and identify the capabilities they lack. For example, BDA tracking capability will enable organizations to evaluate the extent to which they can access real-time and historical information along their supply chains. If organizations find themselves lack of tracking capability in, for example, the distribution process, they will formulate strategies to build tracking capability in this process. Some guidance of building BDA tracking capability in distribution includes (Arunachalam et al., 2017; Sanders, 2016): adopt supply chain technologies such as RFID tags and readers, sensors, warehouse management systems to generate data regarding the distribution process; put in the right infrastructure to support the generation, pre-processing, and storage of different types, huge volumes, and real-time and near real-time data; share intelligence extracted from the distribution process across supply chain functions and partners; identify strategically aligned metrics and key performance indicators; apply descriptive analytics on data from the distribution process to continuously monitor this process; create visual presentation of analytics insights through data visualization tools.

6. Limitations and Future Research

This study has several limitations. First, the search of relevant papers was based on a limited number of databases (i.e. WoS and ABI). Moreover, only papers written in English were included in content analysis. Since big data is a global phenomenon, future research may expand this study by including non-English papers from more databases. It will be interesting to see if we can discover any different BDA capabilities in SCM from non-English papers.

Second, the content-analysis based conceptualization of BDA capability involved authors' subjective judgement and interpretation of selected articles and authors' knowledge in BDA field. However, to ensure the reliability, we invited two supply chain executives with more than 20 years of work experience in Fortune 500 companies to independently review and verify the BDA capability constructs and measurements. They both believed the conceptualization of BDA capability in SCM reflects supply chain practices in the real world.

Lastly, since BDA application in SCM is at an early stage, the proposed BDA capability profile in this study was drawn on the current state of BDA in supply chain context. As more and more organizations start adopting and implementing BDA to solve their supply chain issues, the profile of BDA capability in SCM may continue to evolve. Thus, the proposed profile in this study cannot guarantee that all the aspects of BDA capability in SCM have been considered at this time. Future research may enhance the proposed profile by discovering new BDA capabilities as the use of BDA in SCM advances.

7. Conclusion

There is a growing interest in the application of BDA in SCM from both practitioners and academics. Practitioners have been heavily investing in BDA to exploit the true potential of BDA capability in order to improve supply chain decision-making skills. Academics, however, are still in the early stage of understanding BDA in SCM (Fosso Wamba et al., 2018; Waller and Fawcett, 2013). There is a lack of a comprehensive conceptualization for BDA capability in SCM. Based on content analysis of 19 academic articles and 110 practitioner articles, this work conceptualized BDA capability in SCM as consisting of 22 first-order themes that can be aggregated into four dimensions. A data structure was proposed to display the aggregate dimensions, the first-order themes, and the associated measures. This research broadens the knowledge of BDA in SCM and moves the understanding of this field towards a deeper and more detailed level. Practitioners may incorporate the conceptualization into organizational BDA training and the assessment of organizational BDA capability along supply chain functions.

References

Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(2), 3.

Abbott, G. (2017). Connected fleets leverage analytics for better business. *Material Handling & Logistics,* Retrieved from http://spot.lib.auburn.edu/login?url=https://search.proquest.com/docview/1877731616?accountid=8421 (Accessed on July 10, 2017).

Addo-Tenkorang, R., & Helo, P. T. (2016). Big data applications in operations/supply-chain management: A literature review. *Computers & Industrial Engineering*, *101*, 528-543.

Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology*, *76*(5), 732.

Aronow, S., Burkett, M., Romano, J., & Nilles, K. (2016). The 2016 supply chain TOP 25: Lessons from leaders. *Supply Chain Management Review, 20*(5), 10-21.

Arunachalam, D., Kumar, N., & Kawalek, J. P. (2017). Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. *Transportation Research Part E: Logistics and Transportation Review*.

Babiceanu, R. F., & Seker, R. (2016). Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook. *Computers in Industry*, *81*, 128-137.

Barton, D., & Court, D. (2012). Making advanced analytics work for you. *Harvard business review*, *90*(10), 78-83.

Bentz, B. (2014). "Supply Chain Control Towers Help Organizations Respond to NEW PRESSURES." *Supply Chain Management Review* 18(4), 34-9.

Boone, C. A., Skipper, J. B., & Hazen, B. T. (2017). A framework for investigating the role of big data in service parts management. *Journal of cleaner production*, *153*, 687-691.

Borkowsky, M., & Walther, B. (2014). The next big thing. *Air Transport World, 51*(6), 34.

Briest, P., Dilda, V., & Somers, K. (2014). Taming manufacturing complexity with advanced analytics. *Mckinsey Quarterly,* (4), 24-25.

Brinch, M. (2018). Understanding the value of big data in supply chain management and its business processes: Towards a conceptual framework. *International Journal of Operations & Production Management*.

Brown, B., Chui, M., & Manyika, J. (2011). Are you ready for the era of 'big data'? (cover story). *Mckinsey Quarterly*, (4), 24-35.

Bughin, J., Chui, M., & Manyika, J. (2015). An executive's guide to the Internet of Things. *Mckinsey Quarterly*, (4), 92-101.

Catlin, T., Scanlan, J., & Willmott, P. (2015). Raising your digital quotient. *Mckinsey Quarterly*, (3), 30-43.

Chae, B., & Olson, D. L. (2013). Business analytics for supply chain: A dynamic-capabilities framework. *International Journal of Information Technology & Decision Making*, *12*(01), 9-26.

Chae, B., Olson, D., & Sheu, C. (2014 a). The impact of supply chain analytics on operational performance: A resource-based view. *International Journal of Production Research*, *52*(16), 4695-4710.

Chae, B. K., Yang, C., Olson, D., & Sheu, C. (2014 b). The impact of advanced analytics and data accuracy on operational performance: A contingent resource-based theory (RBT) perspective. *Decision Support Systems*, *59*, 119-126.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, *36*(4).

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. *Journal of Management Information Systems*, *32*(4), 4-39.

Chidambaram, V., Evans, H., & Etheredge, K. (2015). Big data: Is the energy industry starting to see real applications? *Supply Chain Management Review*, *19*(1), 62-64.

Chui, M., Loffler, M., & Roberts, R. (2010). The Internet of Things. *Mckinsey Quarterly*, (2), 70-79.

Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, *70*, 379-390.

De Oliveira, M. P. V., McCormack, K., & Trkman, P. (2012). Business analytics in supply chains–The contingent effect of business process maturity. *Expert Systems with Applications*, *39*(5), 5488-5498.

Demirkan, H., & Delen, D. (2013). Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems, 55*(1), 412-421.

Dhawan, R., Singh, K., & Tuteja, A. (2014). When big data goes lean. *McKinsey Quarterly,* (2), 97-102.

Dirlea, V., Ng, H., Gutiérrez, M., & See, B. (2016). The power of modern supply chain modeling. *Supply Chain Management Review, 20*(5), 52-54.

Douglas, M. (2014). Leveraging Big Data. *Inbound Logistics,* Retrieved from http://www.inboundlogistics.com/cms/article/leveraging-big-data/ (Accessed on July 10, 2017).

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing*, *62*(1), 107-115.

Fathi, N. (2011). A Holistic View of Quality. *Supply & Demand Chain Executive 12*(3), 14-16.

Fedewa, D., Velarde, G. L., & O'Neill, B. (2016). Using buzz analytics to gain a product and marketing edge. *Mckinsey Quarterly,* (2), 14-15.

Field, A. M. (2014). Putting data to work. (cover story). *Journal of Commerce (1542-3867),* 4A-12A.

Finley, F., Blaeser, J., & Djavairian, A. (2015). Harnessing big data: Building & maintaining capabilities that deliver results. *Supply Chain Management Review, 19*(5), 24-28,30-31.

Fosso Wamba, S., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356-365.

Garrett, R. (2016a). Climbing the IoT Mountain. *Supply & Demand Chain Executive,* Retrieved from https://www.sdcexec.com/software-technology/article/12219117/climbing-the-iot-mountain (Accessed on July 17, 2017).

Garrett, R. (2016b). Set your sights on strategic sourcing. *Supply & Demand Chain Executive,* Retrieved from https://www.sdcexec.com/sourcing-procurement/article/12247791/set-your-sights-on-strategic-sourcing (Accessed on July 17, 2017).

Garrett, R. (2016c). Taming big data in the IoT. *Supply & Demand Chain Executive,* Retrieved from https://www.sdcexec.com/software-technology/article/12271216/taming-big-data-in-the-iot (Accessed on July 17, 2017).

Garrett, R. (2016d). Go big or go home. *Supply & Demand Chain Executive,* Retrieved from https://www.sdcexec.com/sourcing-procurement/article/12174233/go-big-or-go-home (Accessed on July 17, 2017).

Garrett, R. (2017). The benefits of digital transformation. *Supply & Demand Chain Executive,* Retrieved from https://www.sdcexec.com/sourcing-procurement/article/12328491/the-benefits-of-digital-transformation (Accessed on July 17, 2017).

Giannakis, M., & Louis, M. (2016). A multi-agent based system with big data processing for enhanced supply chain agility. *Journal of Enterprise Information Management*, *29*(5), 706-727.
Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational research methods*, *16*(1), 15-31.

Godwin, K. (2016). How distributors can win with scan-based trading. *Industrial Distribution (Online),* Retrieved from https://www.inddist.com/article/2016/06/how-distributors-can-win-scan-based-trading (Accessed on July 14, 2017).

Goes, P. B. (2014). Big Data and IS Research. *MIS Quarterly*, *38*(3), iii-viii.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049-1064.

Hadavi, C. (2016). Self-Improving supply chains have arrived. *Inbound Logistics,* Retrieved from http://www.inboundlogistics.com/cms/article/self-improving-supply-chains-have-arrived/ (Accessed on July 11, 2017).

Hahn, G. J., & Packowski, J. (2015). A perspective on applications of in-memory analytics in supply chain management. *Decision Support Systems*, *76*, 45-52.

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, *154*, 72-80.

Hazen, B. T., Skipper, J. B., Ezell, J. D., & Boone, C. A. (2016). Big Data and predictive analytics for supply chain sustainability: A theory-driven research agenda. *Computers & Industrial Engineering*, *101*, 592-598.

Hoehle, H., & Venkatesh, V. (2015). Mobile Application Usability: Conceptualization and Instrument Development. *Mis Quarterly*, *39*(2).

Holsapple, C., Lee-Post, A., & Pakath, R. (2014). A unified foundation for business analytics. *Decision Support Systems*, *6Ex4*, 130-141.

Hurwitz, J., Nugent, A., Hapler, F., Kaufman, M., 2013. Big Data for Dummies. John Wiley & Sons, Hoboken, New Jersey.

Ittmann, H. W. (2015). The impact of big data and business analytics on supply chain management. *Journal of Transport and Supply Chain Management*, *9*(1), 1-9.

Kache, F., & Seuring, S. (2017). Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. *International Journal of Operations & Production Management*, *37*(1), 10-36.

Karimi, J., Somers, T. M., & Bhattacherjee, A. (2007). The role of information systems resources in ERP capability building and business process outcomes. *Journal of Management Information Systems*, *24*(2), 221-260.

Karlskind, M. (2014). How 5 companies use mobile apps to manage big data. *Material Handling & Logistics*, Retrieved from http://spot.lib.auburn.edu/login?url=https://search.proquest.com/docview/1637434936?accountid=8421 (Accessed on July 14, 2017).

Kartal, H., Oztekin, A., Gunasekaran, A., & Cebi, F. (2016). An integrated decision analytic framework of machine learning with multi-criteria decision making for multi-attribute inventory classification. *Computers & Industrial Engineering*, *101*, 599-613.

Khan, K. (2013). The transformative power of advanced analytics. *Supply Chain Management Review*, *17*(3), 48-49.

Kim, S., Kim, H., & Park, Y. (2017). Early detection of vessel delays using combined historical and real-time information. *Journal of the Operational Research Society*, *68*(2), 182-191.

Krippendorff, K. (2012). *Content analysis: An introduction to its methodology*. Sage Publications.

Lade, P., Ghosh, R., & Srinivasan, S. (2017). Manufacturing Analytics and Industrial Internet of Things. *IEEE Intelligent Systems*, *32*(3), 74-79.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT sloan management review*, *52*(2), 21.

Levans, M. (2011). 2011 technology roundtable: Putting data into context. *Logistics Management (2002)*, *50*(5), 42-44,46.

Li, J., Tao, F., Cheng, Y., & Zhao, L. (2015). Big data in product lifecycle management. *The International Journal of Advanced Manufacturing Technology*, *81*(1-4), 667-684.

Liu, P., & Yi, S. P. (2017). A study on supply chain investment decision-making and coordination in the Big Data environment. *Annals of Operations Research*, 1-19.

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems*, *24*(3), 149-157.

Lustig, I., Dietrich, B., Johnson, C., & Dziekan, C. (2010). The analytics journey. *Analytics Magazine*, *3*(6), 11-13.

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS quarterly*, *35*(2), 293-334.

Mani, V., Delgado, C., Hazen, B. T., & Patel, P. (2017). Mitigating supply chain risk via sustainability using big data analytics: evidence from the manufacturing supply chain. *Sustainability*, *9*(4), 608.

Matthias, O., Fouweather, I., Gregory, I., & Vernon, A. (2017). Making sense of Big Data–can it transform operations management? *International Journal of Operations & Production Management*, *37*(1), 37-55.

McCrea, B. (2014). Overhead Handling: Raise Worker Safety. *Modern Materials Handling 69*(12), 34-7.

Michel, R. (2016). Turning data capture into decisions. *Modern Materials Handling, 71*(7), 48-50,52,54-55.

Mishra, D., Gunasekaran, A., Papadopoulos, T., & Childe, S. J. (2016). Big Data and supply chain management: a review and bibliometric analysis. *Annals of Operations Research*, 1-24.

Mitchell, P. (2012). SUPPLY ANALYTICS: An overlooked opportunity. *Supply Chain Management Review*, *16*(4), 26-33.

Nguyen, T., Zhou, L., Spiegler, V., Ieromonachou, P., & Lin, Y. (2017). Big data analytics in supply chain management: A state-of-the-art literature review. *Computers & Operations Research*.

Ou, T. Y., Cheng, C. Y., Chen, P. J., & Perng, C. (2016). Dynamic cost forecasting model based on extreme learning machine-A case study in steel plant. *Computers & Industrial Engineering*, *101*, 544-553.

Palmquist, D., Leal, T. (2016). The IoT supply chain: 3 things you should know. *Inbound Logistics*, Retrieved from http://www.inboundlogistics.com/cms/article/iot-in-the-supply-chain/ (Accessed on July 11, 2017).

Partridge, A. R. (2014). Materials Handling Meets Big Data. *Inbound Logistics*, Retrieved from http://www.inboundlogistics.com/cms/article/materials-handling-meets-big-data/ (Accessed on July 10, 2017).

Pavlou, P. A., & El Sawy, O. A. (2010). The “third hand”: IT-enabled competitive advantage in turbulence through improvisational capabilities. *Information systems research*, *21*(3), 443-471.

Pearson, M. (2011). Predictive analytics: Looking forward to better supply chain decisions. *Logistics Management (2002), 50*(9), 22-23.

Perreault, W. D., & Leigh, L. E. (1989). Reliability of nominal data based on qualitative judgments. *Journal of marketing research*, *26*(2), 135.

Rai, A., Pavlou, P. A., Im, G., & Du, S. (2012). Interfirm IT capability profiles and communications for cocreating relational value: evidence from the logistics industry. *MIS quarterly*, *36*(1), 233-262.

Richey Jr, R. G., Morgan, T. R., Lindsey-Hall, K., & Adams, F. G. (2016). A global exploration of big data in the supply chain. *International Journal of Physical Distribution & Logistics Management*, *46*(8), 710-739.

Ryan, G.W., & Bernard, H. R. (2000). Data management and analysis methods. N. Denzin, Y. Lincoln, eds. *Handbook of Qualitative Research,* 2nd ed. Thousand Oaks, CA, Sage, 769-802.

Sanders, N. R. (2016). How to use big data to drive your supply chain. *California Management Review*, *58*(3), 26-48.

Schoenherr, T., & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: Current state and future potential. *Journal of Business Logistics*, *36*(1), 120-132.

Shah, S., Horne, A., & Capellá, J. (2012). Good data won't guarantee good decisions. *Harvard Business Review*, *90*(4).

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems*, *23*(4), 433-441.

Shenkman, D., Johnson, C., & Elliot, J. (2016). How they did it: Walgreens' talent strategy. *Supply Chain Management Review, 20*(4), 35-39.

Souza, G. C. (2014). Supply chain analytics. *Business Horizons*, *57*(5), 595-605.

Steppeler, B. (2013). Make: Take One Tablet and Get Instant Relief. *Material Handling & Logistics*.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research techniques*. Sage publications.

Suhas, S. (2017). Four ways the Internet of Things will transform the supply chain. *Material Handling & Logistics,* Retrieved from http://www.mhlnews.com/technology-automation/four-ways-internet-things-will-transform-supply-chain (Accessed on July 11, 2017).

Szakonyi, M. (2014). Big data to the rescue. *Journal of Commerce (1542-3867), 15*(10), 64-65.

Tan, K. H., Zhan, Y., Ji, G., Ye, F., & Chang, C. (2015). Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. *International Journal of Production Economics*, *165*, 223-233.

Tiwari, S., Wee, H. M., & Daryanto, Y. (2018). Big data analytics in supply chain management between 2010 and 2016: Insights to industries. *Computers & Industrial Engineering*, *115,* 319-330.

Trebilcock, B. (2013). The big picture on BIG DATA. *Logistics Management, 52*(10), 53-57.

Trebilcock, B. (2015). Cisco's warehouse of the future. *Modern Materials Handling, 70*, 24-28.

Trent, R. J. (2017). 5 supply chain trends happening now. *Supply Chain Management Review*, *21*(2), 18-24.

Trkman, P., McCormack, K., De Oliveira, M. P. V., & Ladeira, M. B. (2010). The impact of business analytics on supply chain performance. *Decision Support Systems*, *49*(3), 318-327.

Urquhart, C., Lehmann, H., & Myers, M. D. (2010). Putting the 'theory' back into grounded theory: guidelines for grounded theory studies in information systems. *Information systems journal*, *20*(4), 357-381.

Viswanathan, N. (2018). Using Prescriptive Analytics to Decrease Supply Chain Disruptions. *Supply Chain Management Review.* Retrieved from http://www.scmr.com/article/are_you_using_prescriptive_analytics_to_decrease_supply_chain_disruptions_y (Accessed on July 29, 2018).

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77-84.

Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, *176*, 98-110.

Wang, Y., & Hajli, N. (2017). Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, *70*, 287-299.

Wang, Y., *Kung*, L., Wang, W. Y. C., & Cegielski, C. G. (2017). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management*.

Ward, T., & Gopal, V. (2014). How they did it: Moving IBM's smarter supply chain to the cloud. *Supply Chain Management Review, 18*(2), 26-31.

Watson, M. (2012). Discovering the value of analytics. *Inbound Logistics*, Retrieved from http://www.inboundlogistics.com/cms/article/discovering-the-value-of-analytics/ (Accessed on July 11, 2017).

Weill, P., & Vitale, M. (2002). What IT infrastructure capabilities are needed to implement e-business models. *MIS quarterly Executive*, *1*(1), 17-34.

Wixom, B. H., Yen, B., & Relich, M. (2013). Maximizing Value from Business Analytics. *MIS Quarterly Executive*, *12*(2), 111-123.

Woods, R. (2017). How smart ULDs are becoming essential Big-Data tools. *Air Cargo World, 107*(3), 22-24.

Wright, J. (2016). Turning data into insight. *Inbound Logistics*, Retrieved from http://www.inboundlogistics.com/cms/article/joanne-wright-turning-data-into-insight/ (Accessed on July 11, 2017).

Zhong, R. Y., Huang, G. Q., Lan, S., Dai, Q. Y., Chen, X., & Zhang, T. (2015). A big data approach for logistics trajectory discovery from RFID-enabled production data. *International Journal of Production Economics*, *165*, 260-272.

Zhong, R. Y., Lan, S., Xu, C., Dai, Q., & Huang, G. Q. (2016 a). Visualization of RFID-enabled shop floor logistics Big Data in Cloud Manufacturing. *The International Journal of Advanced Manufacturing Technology*, *84*(1-4), 5-16.

Zhong, R. Y., Newman, S. T., Huang, G. Q., & Lan, S. (2016 b). Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives. *Computers & Industrial Engineering*, *101*, 572-591.

Zhong, R. Y., Xu, C., Chen, C., & Huang, G. Q. (2017). Big Data Analytics for Physical Internet-based intelligent manufacturing shop floors. *International Journal of Production Research*, *55*(9), 2610-2621.

Zhu, S., Song, J., Hazen, B. T., Lee, K., & Cegielski, C. (2018). How supply chain analytics enables operational supply chain transparency: An organizational information processing theory perspective. *International Journal of Physical Distribution & Logistics Management*, *48*(1), 47-68.

Zhu, S., Gupta, A., Paradice, D., & Cegielski, C. (2018). Understanding the Impact of Immersion and Authenticity on Satisfaction Behavior in Learning Analytics Tasks. *Information Systems Frontiers*, 1-24.

Appendices for Essay I

Appendix 1
Advanced data modeling recently helped one global chemical maker to cut through all these problems in its flagship plant. Company experts in sales, production, and optimization assembled raw-material and product price curves, market-size forecasts, historical equipment-performance data, and more than 600 decision variables into a mathematical model describing the plant's production yields under various operating conditions. The resulting model offers managers a precise understanding of the effects that variations anywhere along the value chain can have on the production network as a whole. The company can now, for example, easily fine-tune the mix of raw materials and finished products, as well as the routing of manufacturing flows, in real time -- while constantly identifying opportunities for improvement (exhibit). All told, these changes increased the plant's EBIT returns by more than 50 percent, and the company is now applying this model across its full factory network so that production capacity can shift in modular fashion.
Concept: predict the effects of variations in operating conditions along the value chain on production yields.
Theme: Predictive Capability in Production.

Essay II

Big Data Analytics Capability in Supply Chain Management: Construct Development and Measurement Validation

1. Introduction

The widespread deployment of digital technologies (e.g. sensors, barcodes, Internet of Things, etc.) at the periphery of enterprise supply chain networks has dramatically increased the volume, variety, and velocity of data generated from supply chain operations. This poses a challenge for organizations because it complicates the identification and extraction of useful insights for managing the supply chain (Kache and Seuring, 2017). Big data analytics (BDA) emerges in this respect to provide organizations with better means for obtaining value from massive amounts of data (Chen, Chiang, and Storey, 2012). The application of BDA in the supply chain domain has been reported in a number of studies (see reviews from Wang et al., 2016 and Addo-Tenkorang and Helo, 2016).

A review of literature on BDA in SCM revealed a lack of research on the mechanism of BDA value creation in SCM. Specifically, studies on the value of BDA have largely focused on the outcomes of BDA by exploring and identifying the potential benefits (Kache and Seuring, 2017; Richey et al., 2016; Schoenherr and Speier-Pero, 2015). Few research digs into BDA itself to understand what it is about BDA that creates value for supply chain and through which process the value creation is realized. Literature regarding business value of Information Technology (IT) has highlighted and demonstrated the importance of IT capability in improving organizational capabilities and competitiveness (Muller et al., 2010; Pavlou and ElSawy, 2010; Rai et al., 2012). Following this logic, a comprehensive investigation of BDA capability in SCM will be an appropriate start to clarifying the underlying mechanism of BDA value creation in the supply chain context. In this study, BDA capability in SCM is defined as the ability to convert high volume, variety, and velocity of data generated along supply chain network into meaningful knowledge, which helps decision makers make decisions and take actions to solve supply chain issues in a timely fashion.

A review of the literature on BDA capability in SCM indicated no context-specific and precise instruments for measuring this construct. First, current measurements of analytics capability in SCM were not customized for the big data environment (e.g. Chae, Olson, and Sheu, 2014a; Chae, Yang, Olson, and Sheu, 2014b) in which more advanced analytics technology is needed to deal with large, various, and real-time or near real-time data (Liu and Yi, 2017). Instead, these measurements were tailored to traditional analytics applications that are only able to deal with explicit (structured) data gathered from internal and limited external sources (Kache and Seuring, 2017). Second, some studies used the “use” of BDA or analytics related resources to measure BDA capability in SCM (e.g. Chae et al., 2014a; Chae et al., 2014b; Chen, Preston, and Swink, 2015; Hahn and Packowski, 2015; Trkman, McCormack, De Oliveira, and Ladeira, 2010; Zhu, Song, Hazen, Lee, and Cegielski, 2018). This practice seems problematic because it will lead to interpretational confounding (Burt, 1976) which occurs when the empirical meaning of an unobserved variable differs from the meaning assigned by the author (Bollen, 2007). The use of BDA or analytics related resources is positively associated with the formation of BDA capability, but both concepts are conceptually distinct and should be treated as such. To address the above issues, this study answers the following research questions: *(1) What are the constructs that constitute big data analytics capability in supply chain management? (2) How to measure those constructs?*

To solve the research questions, this study developed a context-specific and fine-grained instrument for measuring BDA capability in SCM. 110 professional articles and 19 academic articles were analyzed and coded to derive the BDA capability constructs. Mckenzie et al. (2010)'s guideline of construct development and validation was followed to define each BDA capability construct, generate initial measurement items, and test the reliability and validity of the scales.

2. Major Issues of Measuring Big Data Analytics Capability in Supply Chain Management

A search of literature on the assessment of BDA capability in SCM only revealed two studies (i.e. Chen et al. (2015) and Hahn and Packowski (2015)). To properly position the present work, available research on related topic (e.g. business analytics in SCM, supply chain analytics) was also reviewed. The literature review suggested a lack of a context-specific and precise instrument for measuring BDA capability in SCM (See Table 1). Specifically, two major issues were identified.

Table 1. Prior Conceptualizations and Measurement Approaches for BDA Capability in SCM

Elements of BDA Capability	Conceptualization	Measurements	Study Type	Study
Data management resources (DMR), IT-enabled planning resources (IPR), performance management resources (PMR)	Supply chain analytics is a combination of three sets of data and IT-enabled SCM resources that improve operational performance.	DMR includes the use of IT-related resources and analytics resources in data acquisition and management. IPR includes the use of different software tools for supply chain planning areas. PMR includes the degree of investment in the data-orientated performance improvement processes.	Empirical	Chae et al. (2014a)
Advanced analytics	The use of different levels of data analysis methods in manufacturing	The use of three data analysis methods (experience, statistical methods, and	Empirical	Chae et al. (2014b)

	decision-making areas.	mathematical optimization) in manufacturing decision-making areas.		
Big data analytics use	The extent of using BDA across a breath of supply chain processes is associated with an organization's BDA capability.	The extent of BDA use across 10 supply chain domains.	Empirical	Chen et al. (2015)
Data management capability, analytical supply chain process capability, supply chain performance management capability	Supply chain analytics capability is an IT-enabled analytical dynamic capability for managing data, supply chain process, and supply chain performance.	Lack of appropriate measurements for each construct	Conceptual	Chae and Olson (2013)
Analytics capability in plan, analytics capability in source, analytics capability in make, and analytics capability in deliver	Derived from analytics practices in each area of the Supply Chain Operations Reference model.	The extent of use of business analytics practices in the plan, source, make, and deliver area.	Empirical	Trkman et al. (2010)
Monitor-and-navigate, sense-and-respond, predict-and-act, plan-and-optimize	Four use cases of in-memory analytics applications in supply chain management	Lack of appropriate measurements for each dimension.	Conceptual	Hahn and Packowski (2015)
Analytics capability in plan, analytics capability in source, analytics capability in make, and analytics capability in deliver	Derived from analytics practices in each area of the Supply Chain Operations Reference model.	The extent of use of business analytics practices in the plan, source, make, and deliver process.	Emprical	Zhu et al. (2010)

First, the measurements of analytics capability constructs were old style and did not reflect the novelty of big data analytics (e.g. Chae et al, 2014a; Chae et al., 2014b) that enables the extraction of insights from both implicit and explicit data captured from both internal and external sources (Kache and Seuring 2017). Instead, these measurements focused on traditional analytics applications that are only able to deal with explicit (structured) data gathered from internal and limited external sources. For example, Chae et al. (2014a) conceptualized supply chain analytics as a combination of three resources: data management, IT-enabled supply chain planning, and supply chain performance management. The use of three levels of analytics resources (i.e. experience, statistical methods, and mathematical optimization) was applied to assess data management capability. The use of six levels of IT resources (i.e. no formal system, manual, desktop software, customer software, commercial software, and modified commercial software) in the supply chain planning process was employed to evaluate IT-enabled supply chain planning capability. The extent of these IT and analytics resources invested in managing supply chain executions were used to measure the capability of supply chain performance management. Although the findings demonstrated the positive impacts of supply chain analytics on operational performance, a major limitation of their study was that the instrument was not customized for the big data environment in which more advanced analytics technology is needed to deal with large, various, and real-time or near real-time data (Liu and Yi, 2017). Simply adopting such instrument to measure BDA capability in SCM is problematic because you cannot motivate a phenomenon (i.e. big data) without actually delivering on it.

Second, some studies used the “use” of BDA or analytics related resources to measure analytics capability in SCM (Chae et al., 2014a; Chae et al., 2014b; Chen et al., 2015; Hahn and Packowski, 2015; Trkman et al., 2010). For example, Chen et al. (2015) assessed an organization's information processing capability in SCM through the usage of BDA across ten supply chain primary activities (e.g. sourcing, producing, distribution, etc.). Chae et al. (2014b) used the level of data analysis methods (i.e. experience, statistical methods, and mathematical optimization) used in manufacturing decision areas to assess the level of analytical capability. Trkman et al. (2010) measured supply chain analytics capability through the extent of use of analytics practices in each SC decision area (i.e. plan, source, make, and deliver). IT literature has differentiated between IT resources and IT capabilities and has further empirically demonstrated that IT resources are positively related to IT capabilities (Karimi et al., 2007; Ravichandran and Lertwongsatien, 2005; Wang et al., 2012). In other words, a firm's IT capabilities emerge through the process of developing, implementation, and use of IT resources (Wang et al., 2012). Several big data studies also assert that big data analytics capabilities are generated through the configuration of big data analytics resources (e.g. big data architecture) (Wang and Hajli, 2016; Wang et al., 2017). Therefore, Therefore, if we adopt the extent of analytics-related resources usage to measure BDA capability, we will run into interpretational confounding (Burt, 1976) because the empirical meaning of BDA capability in SCM is not aligned with the meaning assigned by the authors (i.e. the use of analytics-related resources) (Bollen, 2007). The use of BDA or analytics related resources is positively associated with the formation of BDA capability, but both concepts are conceptually distinct and should be treated as such.

3. Conceptualization and Instrument Development

A construct is an abstract representation of a phenomenon of interest to researchers. BDA capability in SCM is the construct of interest in this study. The methodology of this study follows the construct development and validation procedures outlined by Mackenzie et al.

(2011). Mackenzie, Podsakoff, and Podsakoff's methodology has been adopted previously for construct development in the field of information technology, analytics, and supply chain management (see, e.g. Chowdhury and Quaddus, 2017; Gupta and George, 2016; Roberts and Grover, 2012).

3.1 Conceptualization of Constructs

The first step of scale development and validation is to define the conceptual domain of the construct because the lack of a clear and precise conceptualization of the focal construct can undermine the construct validity (Mackenzie, 2003; Mackenzie et al., 2011).

Academic journal databases (Web of Science and ABI/INFORM complete) as well as major supply chain trading periodicals (e.g. Supply Chain Management Review, Mckinsey Quarterly, Logistics Management) were explored to capture articles at the nexus of BDA and SCM published between 2007 and 2017. The following keywords were used both separately and in combination (using "AND"/ "OR") for the first-round data selection: big data, analytics, supply chain, logistics, operations management, demand management, supply management, procurement, sourcing, production, manufacturing, distribution, inventory, warehouse, and transportation. This search generated 1,578 articles from academic journals and 1,800 articles from trading periodicals. Next, I read the titles and abstracts of each paper and kept the ones with a focus on big data analytics application in the supply chain domain. The second-round selection lead to 408 and 110 articles from trading periodicals and academic journals respectively. In the third-round selection, I carefully read the full text of each article and kept the ones that primarily focus on describing how an organization's implementation of BDA enables knowledge derivation and thereby better decision-making to address supply chain problems. The final dataset contained 19 academic articles and 110 practitioner papers.

Content analysis technique is an empirically grounded method for extracting themes and topics from text (Krippendorff, 2012). It can help with the identification of key aspects or attributes of a construct's domain required in the conceptualization step (Mackenzie et al., 2011). To ensure a better understanding of BDA capability in SCM, a three-phase content analysis (i.e. preparation, organizing, and reporting) proposed by Elo and Kyngäs (2008) was conducted for each selected article. Moreover, this study adopted inductive content analysis because there is not enough former knowledge on the application of BDA within SCM (Fosso Wamba, Angappa, Papadopoulos, and Ngai, 2018; Schoenherr and Speier-Pero, 2015; Waller and Fawcett, 2013).

The goal of the preparation phase is to understand the coding process, in terms of the selection of unit of analysis, the level of analysis, and the purpose of evaluation (Elo and Kyngäs, 2008). "Themes" were selected as the unit of analysis, which primarily expresses an idea that can be sentences, paragraphs, or a portion of a page (Krippendorff, 2012). The level of analysis is related to organizations that engage in BDA implementation in SCM. The purpose of evaluation is to identify specific aspects of BDA capability in SCM.

The second phase is to organize qualitative data. This is the phase through which all the aspects of BDA capability are identified and conceptualized (Elo and Kyngäs, 2008). I initially read each article several times and highlighted statements related to how analytics can transform big data into useful insights to solve supply chain problems. A total of 428 statements were obtained and copied and pasted into the spreadsheet, which served as the basis for further analysis. The selected statements were then send to the second coder, a senior professor whose research focuses on the intersection of supply chain and big data. In order to process the highlighted statement, Strauss and Corbin's (1998) open, axial, and selective coding procedures were followed to identify conceptually similar aspects. Open coding refers to "the analytical process through which concepts are identified and their properties and dimensions are discovered in data" (Strauss and Corbin, 1998, p.101). Axial coding is "the process of relating categories to their subcategories, termed 'axial' because coding occurs around the axis of a category, linking categories at the level of properties and dimensions" (Strauss and Corbin, 1998, p.123). Selective coding is "the process of integrating and refining the theory" (Strauss and Corbin, 1998, p.143). Following this approach, both coders (i.e. the senior professor and I) independently coded each statement and organize the codes in a Microsoft Excel spreadsheet.

In open coding, all the 428 statements were analyzed using Strauss and Corbin's (1998) line-by-line analysis. Statements found to express similar themes were grouped together and provided with more abstract concepts (i.e. categories). Both coders classified statements expressing similar themes into constructs (i.e. categories) that describe different aspects of BDA capability in supply chain domain. In axial coding, the coders identified subcategories associated with each category to form more precise and complete explanations about BDA capability. The subcategories were abstracted from the statements to describe the content of each category and move beyond description to a higher level of abstraction (Urquhart et al., 2010). As the example displayed in Appendix A, a passage from Briest, Dilda, and Somers (2014) described that big data analytics allows chemical manufacturers to model the effect of various operating conditions on production yields through advanced mathematical modeling. This passage was labeled as "Big Data Analytics Capability in Supply Chain Management–Predictive Capability in Production" in the open coding. Then, the subcategory "predict the effects of variations in operating conditions along the value chain on production yields" was created to describe the predictive capability in the production process during axial coding. In the selective coding, the coders focused on finalizing all the categories and subcategories through comparing, integrating, and refining similar codes emerged during open and axial coding.

Once the coders completed the coding process, they met and compared the coding outcomes. The coding reliability was calculated using the reliability index developed by Perreault and Leigh (1989). The interrater reliability was 0.88, which was higher than the convention of 0.7 (Ryan and Bernard, 2000). Most disagreements occurred between the two coders were on where to put the concepts that describe how BDA solve issues in inventory management, material handling, warehousing, and transportation. A third researcher with more than 15 years of research experience in IT and supply chain management was invited to facilitate the discussion of the disagreement in order to reach an agreement. The three researchers re-analyzed each disputed statement and eventually reached a consensus: combine concepts describing BDA in inventory management, material handling, and warehousing to formulate BDA capability in distribution constructs (i.e. tracking capability in distribution, predictive capability in distribution, analytical capability in distribution, and decision support capability in distribution); relate the concept describing BDA application in transportation to BDA capability in transportation construct (i.e. tracking capability in transportation, predictive capability in transportation, analytical capability in transportation, and decision support capability in transportation). In the end, the coding process lead to 22 initial constructs that represent different aspects of big data analytics capability in supply chain management. Following the guideline of Mackenzie et al. (2011), each construct was conceptualized by identifying the entity to which it applies and the type of property it represents. Table 2 lists all the constructs and their definitions.

Table 2. Constructs, Construct Entities, and Construct Definitions

Construct Name	**Entity to which the Construct Applies and General Property (GP) the**	**Construct Definition**

	Construct Represents	
Tracking Capability in Procurement	E = Organization; GP= the ability to access current and past situations regarding purchased orders and suppliers in the supply base.	The ability of an organization to access current and past situations regarding purchased orders and suppliers in the supply base.
Tracking Capability in Production	E = Organization; GP = the ability to access current and past situations of working orders, operators, machinery, and work flows along the production line.	The ability of an organization to access current and past situations of working orders, operators, machinery, and work flows along the production line.
Tracking Capability in Distribution	E = Organization; GP = the ability to access current and past situations of inventories, workers, equipment, and work flows in distribution centers.	The ability of an organization to access current and past situations of inventories, workers, equipment, and work flows in distribution centers.
Tracking Capability in Transportation	E = Organization; GP = the ability to access current and past situations regarding cargos, transportation assets, staff, and surroundings during transportation from supply points to demand points.	The ability of an organization to access current and past situations regarding cargos, transportation assets, staff, and surroundings during transportation from supply points to demand points.
Analytical Capability in Procurement	E = Organization; GP = the ability to examine historical and real-time supply chain data to discover the root causes of undesirable situations in the supply base.	The ability of an organization to examine historical and real-time supply chain data to discover the root causes of undesirable situations in the supply base.
Analytical Capability in Production	E = Organization; GP = ability to examine historical and real-time supply chain data to discover the root causes of undesirable situations along the production line.	The ability of an organization to examine historical and real-time supply chain data to discover the root causes of undesirable situations along the production line.
Analytical Capability in Distribution	E = Organization; GP = the ability to examine historical and real-time supply chain data to discover the root causes of undesirable situations in distribution centers.	The ability of an organization to examine historical and real-time supply chain data to discover the root causes of undesirable situations in distribution centers.

Analytical Capability in Transportation	E = Organization; GP = the ability to examine historical and real-time supply chain data to discover the root causes of undesirable situations during transportation.	The ability of an organization to examine historical and real-time supply chain data to discover the root causes of undesirable situations during transportation.
Analytical Capability in Product Design and Development	E = Organization; GP = the ability to examine historical and real-time supply chain data to discover the contributing factors to successful products.	The ability of an organization to examine historical and real-time supply chain data to discover the contributing factors to successful products.
Predictive Capability in Demand Management	E = Organization; GP= ability to discover hidden patterns, correlations, or trends from big data to forecast customer demand.	An organization’s ability to discover hidden patterns, correlations, or trends from immense volume, variety, and velocity of data to forecast customers’ demands.
Predictive Capability in Procurement	E = Organization; GP = the ability to discover hidden patterns, correlations, or trends from big data to forecast what is likely to happen to its supply base as well as the corresponding impacts.	The ability of an organization to discover hidden patterns, correlations, or trends from immense volume, variety, and velocity of data to forecast what is likely to happen to its supply base as well as the corresponding impacts.
Predictive Capability in Production	E = Organization; GP = the ability to discover hidden patterns, correlations, or trends from big data to forecast what is likely to happen to the production line and the corresponding impacts.	The ability of an organization to discover hidden patterns, correlations, or trends from immense volume, variety, and velocity of data to forecast what is likely to happen to the production line and the corresponding impacts.
Predictive Capability in Distribution	E = Organization; GP = the ability to discover hidden patterns, correlations, or trends from big data to forecast what is likely to happen to order processing in distribution centers.	The ability of an organization to discover hidden patterns, correlations, or trends from immense volume, variety, and velocity of data to forecast what is likely to happen to order processing in distribution centers.
Predictive Capability in Product Design and Development	E = Organization; GP = the ability to discover patterns, trends, or correlations from big data to forecast customer preferences for new/next-	The ability of an organization to discover patterns, trends, or correlations from immense volume, variety, or velocity of data to forecast customer preferences for new/next-generation products.

	generation products.	
Predictive Capability in Transportation	E = Organization; GP = the ability to discover patterns, trends, or correlations from big data to forecast what is likely to happen during transportation.	The ability of an organization to discover patterns, trends, or correlations from immense volume, variety, and velocity of data to understand what is likely to happen during the transportation.
Decision Support Capability in Demand Management	E = Organization; GP = ability to generate an optimal demand execution plan.	The ability of an organization to generate an optimal demand execution plan that balances the forecasted customer demands.
Decision Support Capability in Procurement	E = Organization, GP= the ability to generate the optimal course of actions to guide supplier selection and supply risk mitigation.	The ability of an organization to generate the optimal course of actions to guide supplier selection and supply risk mitigation.
Decision Support Capability in Production	E = Organization, GP = the ability to generate the optimal course of actions to guide adjustment and optimization decisions during production.	The ability of an organization to generate the optimal course of actions to guide adjustment and optimization decisions during production.
Decision Support Capability in Distribution	E = Organization, GP = the ability to generate the optimal course of actions to guide adjustment and optimization decisions in distribution centers.	The ability of an organization to generate the optimal course of actions to guide adjustment and optimization decisions in distribution centers.
Decision Support Capability in Transportation	E = Organization, GP = the ability to generate the optimal course of actions to guide adjustment and optimization decisions during transportation.	The ability of an organization to generate the optimal course of actions to guide adjustment and optimization decisions during transportation.
Decision Support Capability in Supply Chain Network Design	E = Organization, GP = the ability to generate the optimal course of actions to guide supply chain network optimization decisions.	The ability of an organization to generate the optimal course of actions to guide supply chain network optimization decisions.
Decision Support Capability in Product Design and Development	E= Organization, GP = the ability to generate the optimal course of actions to guide product design and development decisions.	The ability of an organization to generate the optimal course of actions to guide product design and development decisions.

After defining all the constructs, the next phase of construct conceptualization is the identification of higher-order constructs. Constructs sharing a common theme and similar essential characteristics should be theoretically abstracted to a higher level (Hoehle and Venkatesh, 2015; Mackenzie et al., 2011). To identify potential high-order constructs, I carefully examined the constructs in Table 2. I thought through how distinctive the constructs were from each other and whether eliminating any of them would restrain the domain of the construct in a significant way (Mackenzie et al., 2011). For example, tracking capability in procurement, tracking capability in production, tracking capability in distribution, and tracking capability in transportation share the common theme that an organization is able to track output data from all the systems or devices to understand the historical and current situation of each supply chain process. Moreover, eliminating one of these constructs will restrict the domain of tracking capability in a significant way because they represent four distinct facets or processes of supply chain where tracking capability is reflected. Therefore, the four constructs were abstracted to a higher level and identified as tracking capability. Finally, four second-order constructs (i.e. tracking capability, analytical capability, predictive capability, and decision support capability) were identified to represent the aggregation of the 22 first-order constructs. Table 3 displays the four second-order constructs and their definitions.

Table 3. Second-Order Constructs, Construct Entities, and Construct Definitions

Construct Name	**Entity (Entity) to which the Construct Applies and General Property (GP) the Construct Represents**	**Construct Definition**
Tracking Capability	E = Organization; GP = access past and real-time output data from all the systems and devices distributed across the supply chain.	The ability of an organization to access historical and real-time output data from all the systems and devices distributed across organizational boundaries in order to understand the current and past

		situation of its supply chain.
Analytical Capability	E = Organization; GP = the ability to dig into the past and current supply chain data in order to discover the underlying causes of undesirable situations or events and identify improvement areas.	The ability of an organization to dig into the past and current supply chain data in order to discover the underlying causes of undesirable situations or events and identify improvement areas.
Predictive Capability	E = Organization; GP = the ability to project future events, behavior, and outcomes through the supply chain network.	The ability of an organization to discover patterns from real-time and historical supply chain data and extrapolate those patterns to project future events, behavior, and outcomes through the supply chain network.
Decision Support Capability	E = Organization; GP = the ability to determine the optimal course of action for improving supply chain performance, given a complex set of objectives, requirements, and constraints.	The ability of an organization to determine the optimal course of action for improving supply chain performance, given a complex set of objectives, requirements, and constraints.

Second-order Construct: Tracking Capability

Tracking capability is the ability of an organization to access real-time output data from all the systems and devices distributed across organizational boundaries in order to understand the current and past situation of its supply chain. This capability is realized through integrating descriptive analytics with big supply chain data. As organizations increasingly adopt emerging technologies such as sensors, RFID chips, and GPS, huge amounts of data are generated from these devices in real time. Applying descriptive analytics on sensor, RFID, and GPS data provides real-time information regarding the location and conditions of a company’s purchased materials, inventories, machinery and so forth in the supply chain (Souza, 2014). Accessing such real-time information allows organizations to monitor processes and identify anomalies and opportunities for their supply chains (Hahn & Packowski, 2015; Tiwari, Wee, and Daryanto, 2018; Wang et al., 2016).

Second-order Construct: Analytical Capability

Analytical capability is the ability of an organization to dig into the past and current supply chain data in order to discover the underlying causes of undesirable situations or events and identify improvement areas. This BDA capability is formed through applying diagnostic analytics on big supply chain data. Once the anomalies impacting the supply chain are identified through BDA tracking capability, exploratory data analysis of existing data (or additional data if required) is implemented using such tools as drill-down, visualization, data discovery, and data mining so as to understand the possible reasons for the occurrences (Banerjee, Bandyopadhyay, and Acharya, 2013; Delen and Zolbanin, 2018). Determining the factors contributing to the outcomes enables directional guidance for making improvements and reacting to problems in the supply chain.

Second-order Construct: Predictive Capability

Predictive Capability refers to the ability of an organization to discover patterns from real-time and historical supply chain data and extrapolate those patterns to project future events, behavior, and outcomes through the supply chain network. This capability is enabled by analyzing big supply chain data using predictive analytics such as data mining, text mining, web mining, and statistical forecasting (Demirkan and Delen, 2013). Building on the understanding of current and past situations through descriptive and diagnostic analytics (Delen and Zolbanin, 2018), predictive analytics looks at all the possible future scenarios within demand planning, procurement, production, and much more. BDA predictive capability helps organizations get out in front of disruptions and events and respond proactively given current and past conditions (Banerjee et al., 2013).

Second-order Construct: Decision Support Capability

Decision support capability is the ability of an organization to determine the optimal course of action for improving supply chain performance, given a complex set of objectives, requirements, and constraints (Demirkan and Delen, 2013; Lustig, Dietrich, Johnson, and Dziekan, 2010; Wang et al., 2016). This capability is achieved through extracting insights from big supply chain data using prescriptive analytics such as simulation and optimization (Banerjee et al., 2013). BDA decision support capability goes beyond predictive, analytical, and tracking capabilities by recommending a set of possible courses of action. Decision makers will better understand the implications of each suggested action by running different scenarios, eventually selecting the action that best impacts the entire company (Viswanathan, 2018). Companies possessing BDA decision support capability are able to successfully optimize supply chain operations, making sure that they are delivering the right product to the right customer at the right time.

After identifying higher-order constructs, the next question to consider is the nature of the relationship between the sub-dimensions and the higher-order construct, (Mackenzie et al., 2011). Drawing on the guidelines of Mackenzie et al. (2011), the sub-dimensions were thought of as formative indicators of the four second-order constructs. For example, tracking capability in procurement, tracking capability in production, tracking capability in distribution, and tracking capability in transportation were viewed as formative indicators of tracking capability. This is because it seems reasonable that an increase in the level of tracking capability in procurement is associated with an increase in the level of tracking capability, without necessarily being associated with any changes in tracking capability in production, in distribution, or in transportation.

3.2 Item Development

Once the constructs have been well defined, the next step is to generate items that "fully capture all of the essential aspects of the domain of the focal construct" (Mackenzie et al., 2011, p. 304). The subcategories derived from the axial coding process were leveraged to create the initial items because they contained key words describing the domains of constructs identified in the previous section. Eventually, 132 items were initially produced to capture the most essential aspects of the constructs summarized in Table 2.

Face validity check was performed to examine the wording and simplicity of the items. Face validity check is useful and necessary when items are newly developed and have never been validated by individuals (DeVellis, 2011; Hoehle and Venkatesh, 2015). A face validity check focuses on the items themselves and does not ask respondents to rate the items. A prerequisite for participating in the face validity check was that the participant has backgrounds in both supply chain management and analytics, which would ensure that he/she understands the context of the items. One professor whose research focuses on big data and supply chain area, one lecturer with more than 20 years of work experience in supply chain function, and one supply chain manager from a Fortune 500 firm voluntarily participated in the face validity check. The participants were provided with a Microsoft Word document that included all the 132 items. They were asked to examine all the items and to mark/comment on any items that they thought the wording was vague or confusing. In total, 34 items were identified as confusing or worded unclearly. Through a discussion with the two participants, I reworded one item and deleted 33 items from the initial item pool. This led to a total of 99 items that would be used in the next step: content validity check.

3.3 Assess the Content Validity of the Items

Content validity concerns “the degree to which items in an instrument reflect the content universe to which the instrument will be generalized” (Straub et al., 2004, p. 424). MacKenzie et al. (2011) recommended using a variance analysis approach suggested by Hinkin and Tracey (1999). This approach includes the use of a matrix in which the new items are placed in rows and construct definitions are placed in columns. Raters are asked to indicate the extent to which each item captures construct domains using a 5-point Likert scale (i.e. 1= not at all, 5=complete). Then, a one-way ANOVA is performed to assess whether an item’s mean rating on one construct is significantly different from its rating on other constructs.

Although this approach provides precise assessments (Yao et al., 2007), one disadvantage is that rating all the item-construct combinations can overburden raters when the number of items is too large (Hoehle and Venkatesh, 2015). Hoehle and Venkatesh (2015) used variance analysis approach on 82 items and 20 constructs and indicated that the rating would overburden raters even though the item pool was split into several matrixes. My item pool contains 99 items and 22 constructs, which is larger than that in Hoehle and Venkatesh’s work. Given the length of my survey and Hoehle and Venkatesh’s suggestion, I used the method suggested by Anderson and Gerbing (1991) to assess the content adequacy of new items. This method assumes each item represents a single construct. Raters are asked to select only one construct for each item rather than rate each item-construct combination, which is especially suited for large survey instruments (Yao et al., 2007). According to Anderson and Gerbing’s method, the content validity survey was developed to ask respondents to assign only one construct to each item.

MacKenzie et al. (2011) and Anderson and Gerbing (1991) recommended that the target respondents for content validity check should be representative of the population of interest. Thus, I recruited participants whose job function lies in supply chain or logistics area from MBA alumni panel at Auburn University College of Business. Ten professionals agreed to participate in my study. The content validity survey link was send to each participant and 10 responses were collected. The average complete time is 24 minutes and 59 seconds. There is one response with complete time of 3 minutes and 54 seconds. I believe the responder of this response did not pay sufficient attention to the questions because the complete time of the rest of responses ranges from 16 minutes and 34 minutes to 43 minutes. Thus, I excluded that response and 9 responses were used to assess the content validity of items.

Two indices were calculated from the survey data based on Anderson and Gerbing's work. First, I computed the proportion of substantive agreement (P_{SA}), which refers to the proportion of respondents who assign an item to the intended construct by using the following formula:

$$Psa = \frac{nc}{N}$$

Where *nc* is the number of respondents assigning an item to its intended construct, and N is the number of respondents in total. The value of P_{SA} ranges from 0 to 1, with a larger value indicating higher extent that an item reflects its intended construct (Anderson and Gerbing, 1991).

Second, I computed the substantive validity coefficient (C_{SV}). C_{SV} refers to the extent to which respondents assign an item to its intended construct more than to any other construct. C_{SV} is calculated by the following formula:

$$Csv = \frac{\mathrm{nc-no}}{\mathrm{N}}$$

Where *nc* is the number of respondents assigning an item to its intended construct, *no* is the highest number of assignments to any other construct, and N is the number of respondents in total. The value of C_{SV} ranges from -1 to 1. A positive value indicates that an item is assigned to the intended construct more than it is assigned to any other construct. A negative value indicates the opposite (Anderson and Gerbing, 1991).

The results of content validity assessment are displayed in Appendix B. Consistent with Hoehle and Venkatesh (2015), I applied a threshold of 0.6 as a cut-off value for P_{SA} and C_{SV}. The 0.6 cut-off value reflects that at least 60 percent of respondents assign the items to the intended construct definitions.

Overall, the content validity indices were high, suggesting that most respondents assigned majority of items their intended construct definitions. Out of 99 items, 13 items did not meet the 0.6 cut-off value. Appendix B displays that the C_{SV} values of TrD1, TrD3, TrT1, ACPC1, ACPC2, ACD2, PCPC5, PCD2, PCD4, PCD5, DSCDM4, DSCT6, DSCPDD4 were lower than the threshold 0.6. The P_{SA} values of TrD1, PCPC5, PCD5, DSCDM4, and DSCT6 were lower than the threshold 0.6. Thus, I carefully inspected the wording of these items and compared them with construct definitions. In some cases, I reworded the item (i.e. TrD3, TrT1, ACPC1, ACPC2, ACD2, PCD2, PCD4, DSCPDD4) to better align with the posited construct definitions. In some cases, the item (i.e. TrD1, PCPC5, PCD5, DSCDM4, DSCT6) was deleted because the wording was so generic that it can cover different aspects of supply chain process. Appendix B shows the item pool after content validity check including 94 items for measuring big data analytics capability in supply chain management.

3.4 Formally Specify the Measurement Model

Once a set of content valid items are generated, the next step is to formally specify the measurement model. Following the guidelines of assessing hierarchical construct models developed by Wetzels, Odekerken-Schröder, and Oppen (2009), I formally specified the relationships between indicators, sub-dimensions, and higher-order constructs. First, the first-order latent constructs were constructed by relating them to their indicators as Mode A (reflective). Then, I constructed the second-order latent constructs by connecting them with their first-order latent constructs as formative dimensions and repeating the indicators of their first-order constructs using Mode B (formative). Therefore, tracking capability consisted of the indicators of tracking capability in procurement, production, distribution, and transportation. Analytical capability was made up of the indicators of analytics capability in procurement, production, distribution, transportation, and new product design and development. Predictive capability was linked to the indicators of predictive capability in demand management, procurement, production, distribution, transportation, and new product design and development. Decision support capability was connected to the indicators of demand management, procurement, production, distribution, transportation, new product design and development, and supply chain network design and development. Likewise, the third-order latent construct (BDA capability in SCM) was constructed by relating it with its second-order latent constructs as formative dimensions and repeating the indicators of its second-order latent constructs using Mode B (formative).

3.5 Collect Data to Conduct Pretest

I created a survey including instructions for participants and the items finalized in the step of content validity check. All the items were measured using a 7-point Likert scale, with 1 = Strongly Disagree and 7 = Strongly Agree. The survey participants were recruited through LinkedIn. The target respondents are supply chain/logistics managers or higher level whose company produces or manufactures products. In total, 147 responses were received. All responses were scrutinized for the survey completion duration and the quality of answers.

Responses that were completed in less than 7 minutes and/or included straight-line answers were excluded from the sample. A cut-off threshold of 7 minutes as applied because a response ratio of 15 questions per minute would infer that the respondent did not pay enough attention to the survey questions (Hoehle and Venkatesh, 2015). Consequently, out of 147 responses, 134 responses were usable for data analysis. Table 4 displays the summary of respondent demographics. The sample included firms from a variety of industry (e.g. food and beverage, electronics, health care, etc.) and the positions of respondents included manager, senior manager, director, senior director, vice president, senior vice president, and executives of supply chain/logistics.

Table 4. Respondent Demographics

		Study I	
Demographic	**Category**	**N=134**	**%**
Position	Manager of supply chain/logistics	61	45.5
	Senior manager of supply chain/logistics	25	18.7
	Director of supply chain/logistics	31	23.1
	Vice President of supply chain/logistics	10	7.5
	Senior Vice President of supply chain/logistics	4	3
	Supply chain/Logistics Executives	3	2.2
Number of Employees	Less than 100	17	12.7
	100 - 250	15	11.2
	250 - 500	4	3
	500 - 1,000	14	10.4
	1,000 - 5,000	15	11.2
	5,000 - 10,000	14	10.4
	More than 10,000	55	41.1
Annual Revenue	Less than $1Million	3	2.2
	$1 Million - $5 Million	7	5.2
	$5 Million - $10 Million	3	2.2
	$10 Million - $20 Million	10	7.5
	$20 Million - $50 Million	16	11.9
	$50 Million - $100 Million	16	11.9
	$100 Million - $500 Million	11	8.2
	More than $500 Million	68	50.9
Industry	Aerospace	13	9.7
	Agriculture	2	1.5
	Automotive	10	7.5

	Chemical	9	6.7
	Computer	3	2.2
	Construction	2	1.5
	Defense	5	3.7
	Electronics	13	9.7
	Food & Beverage	16	11.9
	Healthcare	9	6.7
	Industrial Equipment	6	4.5
	Oil & Gas	5	3.7
	Pulp & Paper	3	2.2
	Steel	9	6.7
	Telecommunication	5	3.7
	Textile	4	3
	Others	20	15.1

3.6 Scale Purification and Refinement

I used SPSS and SmartPLS 3.0 to conduct data analysis. Given that I had a large number of indicators (i.e. 94 indicators), I first performed an exploratory factor analysis (EFA) in SPSS using principal component analysis and varimax rotation. Twenty-two factors emerged (eigen value > 1) from the analysis and all items loaded on the hypothesized constructs.

Following the guideline of Mackenzie et al (2011), I used SmartPLS 3.0 to (1) assess the validity of the set of indicators at construct level, (2) assess the reliability of the set of indicators at construct level, and (3) assess individual indicator validity and reliability.

3.6.1 Assess the validity of the set of indicators at construct level

The assessment methods for reflective and formative constructs are different. For the first-order reflective constructs, convergent validity can be evaluated by calculating average variance extracted (AVE) (Fornell and Larcker, 1981; Mackenzie et al., 2011). The AVE of all constructs are greater than 0.50 (Table 6), indicating that the latent construct on average accounts for a majority of variance in its indicators. The square roof of AVE of each latent construct is greater than its correlation with any other constructs (Table 6), supporting the discriminant validity of first-order reflective constructs (Fornell and Larcker, 1981). For the second-order and third-order formative constructs, I evaluated the validity of the set of sub-dimensions serving as formative indicators of the second-order and third-order construct using Edward's (2001) adequacy coefficient (R_a^2). R_a^2 is calculated by summing the squared correlation between each higher construct and its sub-dimensions and then dividing by the number of sub-dimensions. All R_a^2 values exceed 0.50 (Table 5), suggesting that a majority variance in the subdimensions (i.e. the first-order and second-order sub-dimensions) is shared with the second-order and third-order constructs.

Table 5. Second- and Third-order Constructs Validation

Construct	Sub-dimension	Weight	Significance	R_a^2	VIF
Tracking Capability (TC)	TC in Procurement	0.286	p<0.001	0.72	1.62
	TC in Production	0.313	p<0.001		2.51
	TC in Distribution	0.334	p<0.001		3.19
	TC in Transportation	0.3	p<0.001		1.92
Analytical Capability (AC)	AC in Procurement	0.364	p<0.001	0.72	1.6
	AC in Production	0.225	p<0.001		3.32
	AC in Distribution	0.206	p<0.01		3.68
	AC in Transportation	0.222	p<0.001		2.88
	AC in Product Design and Development	0.219	p<0.001		2.45
Predictive Capability (PC)	PC in Demand Management	0.192	p<0.01	0.76	2.56
	PC in Procurement	0.182	p<0.001		2.14
	PC in Production	0.204	p<0.01		3.21
	PC in Distribution	0.213	p<0.001		5.2
	PC in Transportation	0.199	p<0.01		3.01
	PC in Product Design and Development	0.206	p<0.01		3.37
Decision Support Capability (DSC)	DSC in Demand Management	0.135	p<0.05	0.71	2.51
	DSC in Procurement	0.161	p<0.05		2.6
	DSC in Production	0.164	p<0.05		2.85
	DSC in Distribution	0.193	p<0.01		3.98
	DSC in Transportation	0.167	p<0.05		3.65
	DSC in Product Design and Development	0.2	p<0.01		2.89
	DSC in SC Network Design and Development	0.203	p<0.01		2.66
Big Data Analytics Capability in	Tracking Capability	0.254	p<0.001	0.89	3.57
	Analytical Capability	0.271	p<0.001		8.3

Supply Chain Management	Predictive Capability	0.267	p<0.001		7.13
	Decision Support Capability	0.267	p<0.001		6.95

Table 6. Reliability and Inter-correlations of First-order Constructs

	Construct	CR	Cronbach's α	AVE	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	**ACD**	0.9	0.9	0.7	**0.84**																					
2	**ACPC**	0.89	0.89	0.58	0.55	**0.76**																				
3	**ACPDD**	0.97	0.97	0.89	0.56	0.49	**0.94**																			
4	**ACPD**	0.85	0.85	0.66	0.77	0.51	0.69	**0.81**																		
5	**ACT**	0.82	0.82	0.61	0.74	0.56	0.65	0.59	**0.78**																	
6	**DSCDM**	0.91	0.91	0.73	0.64	0.62	0.53	0.66	0.62	**0.85**																
7	**DSCD**	0.91	0.91	0.71	0.81	0.52	0.64	0.66	0.79	0.69	**0.84**															
8	**DSCPDD**	0.94	0.94	0.76	0.61	0.57	0.78	0.72	0.63	0.63	0.66	**0.87**														
9	**DSCPC**	0.91	0.91	0.67	0.52	0.78	0.62	0.62	0.53	0.62	0.58	0.71	**0.82**													
10	**DSCPD**	0.92	0.92	0.7	0.61	0.5	0.58	0.8	0.58	0.61	0.62	0.72	0.68	**0.84**												
11	**DSCSCN**	0.93	0.93	0.74	0.73	0.56	0.58	0.54	0.7	0.65	0.76	0.58	0.54	0.54	**0.86**											
12	**DSCT**	0.92	0.92	0.69	0.64	0.58	0.64	0.69	0.8	0.57	0.77	0.68	0.67	0.72	0.67	**0.83**										
13	**PCDM**	0.9	0.89	0.68	0.65	0.58	0.66	0.71	0.6	0.8	0.72	0.61	0.6	0.54	0.63	0.58	**0.83**									
14	**PCD**	0.91	0.9	0.71	0.83	0.67	0.62	0.78	0.72	0.76	0.79	0.7	0.7	0.71	0.68	0.77	0.74	**0.84**								
15	**PCPDD**	0.94	0.94	0.79	0.63	0.61	0.76	0.8	0.58	0.6	0.62	0.81	0.67	0.69	0.49	0.69	0.69	0.81	**0.89**							
16	**PCPC**	0.89	0.89	0.68	0.49	0.68	0.63	0.63	0.54	0.63	0.54	0.64	0.8	0.52	0.51	0.6	0.62	0.62	0.62	**0.82**						
17	**PCPD**	0.94	0.94	0.79	0.64	0.55	0.66	0.86	0.6	0.65	0.62	0.72	0.64	0.82	0.54	0.63	0.69	0.8	0.75	0.62	**0.89**					
18	**PCT**	0.94	0.94	0.79	0.57	0.67	0.64	0.63	0.76	0.55	0.69	0.63	0.68	0.69	0.56	0.88	0.59	0.78	0.7	0.66	0.67	**0.89**				
19	**TCD**	0.9	0.9	0.7	0.74	0.44	0.67	0.81	0.61	0.64	0.7	0.62	0.53	0.59	0.61	0.61	0.67	0.72	0.65	0.52	0.63	0.54	**0.84**			
20	**TCPC**	0.87	0.87	0.63	0.5	0.55	0.46	0.44	0.52	0.54	0.52	0.48	0.57	0.39	0.4	0.41	0.53	0.53	0.46	0.56	0.47	0.54	0.55	**0.79**		
21	**TCPD**	0.9	0.9	0.7	0.65	0.41	0.51	0.8	0.42	0.62	0.54	0.54	0.5	0.72	0.43	0.46	0.59	0.6	0.59	0.45	0.66	0.48	0.75	0.57	**0.84**	
22	**TCT**	0.89	0.89	0.63	0.59	0.42	0.62	0.66	0.76	0.51	0.62	0.52	0.42	0.54	0.54	0.75	0.53	0.58	0.53	0.48	0.49	0.66	0.68	0.48	0.5	**0.79**

Note: (1) ACD–Analytical Capability in Distribution, ACPC– Analytical Capability in Procurement, ACPDD– Analytical Capability in Product Design & Development, ACPD– Analytical Capability in Production, ACT– Analytical Capability in Transportation, DSCDM–Decision Support Capability in Demand Management, DSCD– Decision Support Capability in Distribution, DSCPDD– Decision Support Capability in Product Design & Development, DSCPC– Decision Support Capability in Procurement, DSCPD– Decision Support Capability in Production, DSCSCN– Decision Support Capability in Supply Chain Design & Development, DSCT– Decision Support Capability in Transportation, PCDM–Predictive Capability in Demand Management, PCD– Predictive Capability in Distribution, PCPDD– Predictive Capability in Product Design & Development, PCPC– Predictive Capability in Procurement, PCPD– Predictive Capability in Production, PCT– Predictive Capability in Transportation, TCD–Tracking Capability in Distribution, TCPC– Tracking Capability in Procurement, TCPD– Tracking Capability in Production, TCT– Tracking Capability in Transportation (2) The bold values on the diagonal are the square root of AVEs.

3.6.2 Assess the reliability of the set of indicators at construct level

For the first-order reflective constructs, Cronbach's α and Fornell and Larker's (1981) construct reliability were used to evaluate the internal consistency reliability of measures (Mackenzie et al., 2011). The values of Cronbach's α and construct reliability are above 0.7 for all first-order constructs (Table 6), implying internal consistency among the measures. Since the notion of internal consistency cannot be applied to the set of sub-dimensions serving as formative indicators of second- and third-order constructs, Cronbach's α and construct reliability were not reported for higher-order formative constructs.

3.6.3 Assess individual indicator validity and reliability

I inspected the item loadings for the first-order reflective constructs. Table 7 shows the results. All items loaded significantly on the hypothesized latent construct, with item loadings ranging from 0.641 to 0.963 and only TrPC1 loading lower than 0.7, thus indicating individual indicator validity (Fornell and Larcker, 1981). Because each item loads on only one latent construct in my study, the measures for assessing individual indicator validity and reliability will be equal (Mackenzie et al., 2011). Therefore, individual indicator reliability is also supported. For higher-order formative constructs, I evaluated the weights of sub-dimensions (i.e. first-order and second-order constructs) on respective higher-order constructs (i.e. second-order and third-order constructs). The results are shown in Table 5. All weights are statistically significant, which suggests that each lower-order construct contributes significantly to its higher-order construct. Table 6 shows that all values of Fornell and Larcker's (1981) construct reliability are greater than 0.7, thus supporting the reliability of each individual sub-dimension. Mackenzie et al. (2011) emphasized the importance of examining the multicollinearity among formative indicators. Variance Inflation Factor (VIF) below 10 in general implies low multicollinearity.

The VIF values of all second-order and third-order formative constructs are less than 10 (Table 5), suggesting that multicollinearity is not an issue in this study.

The results of data analysis demonstrated high validity and reliability of first-order reflective constructs and second-order and third-order formative constructs. I did not feel any items needed to be eliminated from the instrument.

Table 7. Item Loadings

Construct Name	Loadings	Construct Name	Loadings	Construct Name	Loadings
Tracking Capability in Procurement (TrPC 1-4)	0.641	Predictive Capability in Demand Management (PCDM 1-4)	0.725	Decision Support Capability in Production (DSCPD 1-5)	0.76
	0.756		0.883		0.786
	0.82		0.838		0.884
	0.919		0.846		0.857
Tracking Capability in Production (TrPD 1-4)	0.792	Predictive Capability in Procurement (PCPC 1-4)	0.672		0.889
	0.816		0.868	Decision Support Capability in Transportation (DSCT 1-5)	0.775
	0.828		0.928		0.789
	0.903		0.801		0.805
Tracking Capability in Distribution (TrD 1-4)	0.834	Predictive Capability in Production (PCPD 1-4)	0.862		0.929
	0.873		0.894		0.834
	0.857		0.877	Decision Support Capability in Product Design and Development (DSCPDD 1-5)	0.874
	0.771		0.912		0.924
Tracking Capability in Transportation (TrT 1-5)	0.82	Predictive Capability in Distribution (PCD 1-4)	0.813		0.831
	0.75		0.88		0.852
	0.818		0.863		0.871
	0.788		0.801	Decision Support Capability in Supply Chain Network Design and Development (DSCSCN 1-5)	0.865
	0.783	Predictive Capability in Transportation (PCT 1-4)	0.937		0.944
Analytical Capability in Procurement (ACPC 1-6)	0.779		0.884		0.847
	0.79		0.877		0.808
	0.759		0.86		0.831
	0.83	Predictive Capability in Product Design and Development (PCPDD 1-4)	0.897		
	0.711		0.871		
	0.701		0.885		
Analytical Capability in Production (ACPD 1-3)	0.774		0.902		
	0.851	Decision Support Capability in Demand Management (DSCDM 1-4)	0.867		
	0.813		0.898		
Analytical Capability in Distribution (ACD 1-4)	0.867		0.806		
	0.793		0.835		
	0.846	Decision Support Capability in	0.85		
	0.831		0.811		

Analytical Capability in Transportation (ACT 1-3)	0.765	Procurement (DSCPC 1-5)	0.875	
	0.752		0.774	
	0.819		0.763	
Analytical Capability in Product Design and Development (ACPDD 1-4)	0.899	Decision Support Capability in Distribution (DSCD 1-4)	0.861	
	0.949		0.802	
	0.955		0.882	
	0.963		0.818	

4. Discussion

This research developed and validated a BDA capability in SCM conceptualization and survey instrument following the construct measurement and validation procedures proposed by MacKenzie et al. (2011). Content analysis of 129 academic and practical articles was used to develop the conceptualization and instrument. I conceptualized BDA capability in SCM as a hierarchical construct which includes four second-order constructs that are formed through twenty-two first-order constructs. The scale development process embodied item generation, face validity assessment, and content validity assessment. Once the measurement model was specified, I performed the first-round data collection to examine scale properties and refine the items. 134 usable responses were collected from supply chain/logistics managers or above level. The scales were tested and found to be reliable and valid.

4.1 Theoretical Implications

This research contributes to the existing body of big data literature in the following ways. First, it advances the expansion of big data literature by presenting a context specific and precise theoretical framework of BDA capability in SCM and providing empirical evidence to support the proposed framework. Specifically, the literature review suggested that existing measures of analytics capability in supply chain domain are not tailored to big data context. To solve this problem, I developed and validated analytics capability constructs unique to big data context. For instance, tracking capability refers to the ability of an organization to access historical and real-time output data from all the systems and devices distributed across organizational boundaries so as to understand the past and current situation of its supply chains. This capability indicates the volume, variety, and velocity characteristics of big data and emphasizes that advanced analytics can process big data to understand what happened and what is happening to the supply chain. In addition, previous research drew on scales developed for IT capability for measuring BDA capability (e.g. Akter, Fosso Wamba, Gunasekaran, Dubey, and Childe (2016) and Fosso Wamba, Gunasekaran, Akter, Ren, Dubey, and Childe (2017)). Although it is reasonable to use measurements for IT capability as a starting point, such measurements may not capture the important aspects relevant to big data analytics. This study asserts that BDA capability is different from IT capability. IT capabilities include compatibility, connectivity, and modularity of different systems, the management of IT resources in accordance with business needs and priorities, and IT staff's professional ability to undertake assigned tasks (Kim, Shin, and Kwon, 2012). In contrast, the four second-order BDA capabilities developed in this study–tracking, analytical, predictive, and decision support capability–are regarding the abilities to extract insights from big data so as to be aware of current and historical situations, analyze the reasons, predict future situations, and make informed decisions. The literature review also showed that some studies used the extent of use of BDA resources in SCM to measure BDA capability in SCM, which will lead to interpretational confounding. To address this issue, this study offered more precise conceptualization and measurements of BDA capability in SCM. All the BDA capabilities developed in this study are the capabilities formed through the use of BDA resources, ensuring the match between the nominal and empirical meaning of BDA capability in SCM.

Given that this study has provided more context-specific, holistic, and accurate conceptualization and instrument for BDA capability in SCM, researchers can use the BDA capability constructs and measures to further the study of the emerging stream of big data in supply chain domain. For instance, there is considerable interest in investigating the value creation mechanism of BDA in SCM (Chen et al., 2015; Fosso Wamba et al., 2015; Waller and Fawcett, 2013). I believe that the BDA capability constructs and instrument developed in this study can be an ideal candidate for exploring the value creation process in more depth. Specifically, future studies could use the constructs and instrument, in combination with other theories such as resource-based theory (Barney, 2001), contingency theory (Drazin and Van De Ven, 1985), and resource-picking and capability-building view (Makadok, 2001; Karimi, Somers, and Bhattacherjee, 2007), to study how BDA resources help build BDA capabilities which create value for supply chains and how context factors moderate the value creation process.

4.2 Practical Implications

This study has important implications for practitioners because critical BDA capability elements have been identified in the supply chain context. First, this study measured BDA capability in SCM at a granular level. Practitioners could use the conceptualization and instrument to inform the implementation of big data analytics to drive supply chains. Sanders (2016) presents a prescriptive framework showing that companies should use BDA along all supply chain functions in a functionally linked manner rather than optimize single functions or single decision areas in isolation to ensure the coordination of activities. Therefore, all supply chain functions should build BDA capabilities so that insights from big data offered by one function could be linked to and utilized by the rest to drive the entire supply chain. Companies may adopt the instrument to determine the BDA capabilities in abundance and the capabilities they lack in each process. For example, scales pertaining to predictive capability in procurement will allow organizations to assess the extent to which they possess the ability to project future events, behavior, and outcomes on the supplier side. Lack of predictive capability (e.g. not able to predict the impacts of negative events happening to suppliers) may prevent downstream operations from responding to the events in a timely manner. If a company identifies itself lack of predictive capability in the procurement function, it could formulate plans to establish predictive capability in this process in order that insights regarding suppliers could be generated and passed to downstream operations for timely responses.

In addition, the instrument can help organizations evaluate the maturity of implementing BDA along SC functions and whether organizations are ready to move to the next maturity level. Sanders (2016) proposed a four-stage maturity map for best accomplishing BDA implementation. The first stage is ensuring the data gathered are clean and in good quality so that they can be used for further analysis. If a company aims to build tracking capability in the production process, it could utilize the first-order tracking capability in production construct to determine what data to be collected to obtain that capability and whether such data are clean, structured, and well-organized. The second maturity stage is making the right data available when and where needed. Organizations may refer to the instrument to evaluate whether the data required for building BDA capability are available in a usable form in respective supply chain functions. For example, scales concerning tracking capability in transportation represent the ability of an organization to access the past and current situations regarding cargos, transportation assets, staff, and surroundings during the transport from supply points to demand points. Companies concentrating on big data efforts in the transportation process could use the scales to check whether data about cargos, drivers, vehicles, external environment and so on are available in a usable format in transportation departments. The third maturity stage is applying basic analytics to the data. Scales with respect to tracking and analytical capability were derived from statements about the use of descriptive and diagnostic analytics (i.e. basic analytics). If an organization focuses on establishing BDA capability in the distribution process for instance, it could apply the scales to examine whether tracking and analytical capabilities have been built for this process through using basic analytics and whether it is ready for the organization to move to the next maturity level. The fourth maturity stage is applying advanced analytics. Scales with regard to predictive and decision support capability were originated from statements about the use of advanced analytics such as predictive and prescriptive analytics. Organizations could use the scales to verify whether they have gained predictive and decision support capabilities by using advanced analytic and whether they have reached the highest maturity level of BDA implementation across the supply chain.

5. Limitations and Future Research

This study has several limitations. First, since the incorporation of BDA in the field of supply chain management is a very recent phenomenon, the conceptualization and instrument developed in this study may not cover every aspect of BDA capability at this time. Currently, a number of organizations are still in the process of evaluating and implementing BDA along supply chains and BDA tools with new features are emerging. Hence, new BDA capabilities may arise and the proposed BDA capability in SCM in this study may continue to evolve. Researchers in the future could complement and enhance this study by identifying other new BDA capabilities in the supply chain sector as BDA keeps advancing steadily and BDA implementation across supply chains becomes mature.

Second, nomological validity of the instrument was not tested in this study. However, this is something I plan to do in the future. Nomological validity will be assessed by including two constructs (cost efficiency and customer satisfaction) to examine the relationship between BDA capability in SCM and organizational performance. A survey including both refined measures of BDA capability in SCM and measure for organizational performance will be sent to supply chain/logistics managers or above level. The responses will be used for reexamining the scales and evaluating nomological validity.

Third, this study collected data only from companies in the United States. As big data is a global phenomenon and organizations in other countries are also implementing BDA to drive supply chain, futures research could expand this study by collecting a broader sample of companies from different countries. It will be interesting to see whether the country-level differences influence the scale properties of BDA capability in SCM constructs as well as the relationship between BDA capability in SCM and organizational performance.

6. Conclusion

Despite big data analytics being a prevalent research topic in various disciplines, the literature review revealed that there is a lack of theoretical clarity on holistically evaluating BDA capability in supply chain domain. This study bridges this gap by providing a context-specific conceptualization of BDA capability in SCM at granular level and developing a reliable and valid instrument for measuring this capability. This work makes important contributions to both big data and supply chain research because it grounds future empirical research at the nexus of BDA and SCM and aids theory development in the field of BDA value creation for supply chains. Practitioners may adopt the conceptualization and instrument to inform their BDA implementation along supply chain functions.

References

Addo-Tenkorang, R., & Helo, P. T. (2016). Big data applications in operations/supply-chain management: A literature review. *Computers & Industrial Engineering*, *101*, 528-543.

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics, 182*, 113-131.

Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology*, *76*(5), 732.

Barney, J. B. (2001). Resource-based theories of competitive advantage: A ten-year retrospective on the resource-based view. *Journal of management*, *27*(6), 643-650.

Bollen, K. A. (2007). Interpretational confounding is due to misspecification, not to type of indicator: Comment on Howell, Breivik, and Wilcox (2007). Psychological Methods, 12(2), 219-228.

Briest, P., Dilda, V., & Somers, K. (2014). Taming manufacturing complexity with advanced analytics. *Mckinsey Quarterly,* (4), 24-25.

Burt, R. S. (1976). Interpretational confounding of unobserved variables in structural equation models. *Sociological methods & research*, *5*(1), 3-52.

Chae, B., & Olson, D. L. (2013). Business analytics for supply chain: A dynamic-capabilities framework. *International Journal of Information Technology & Decision Making*, *12*(01), 9-26.

Chae, B., Olson, D., & Sheu, C. (2014 a). The impact of supply chain analytics on operational performance: A resource-based view. *International Journal of Production Research*, *52*(16), 4695-4710.

Chae, B. K., Yang, C., Olson, D., & Sheu, C. (2014 b). The impact of advanced analytics and data accuracy on operational performance: A contingent resource-based theory (RBT) perspective. *Decision Support Systems*, *59*, 119-126.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, *36*(4).

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. *Journal of Management Information Systems*, *32*(4), 4-39.

Chowdhury, M. M. H., & Quaddus, M. (2017). Supply chain resilience: Conceptualization and scale development using dynamic capability theory. *International Journal of Production Economics*, *188*, 185-204.

DeVellis, R. 2011. *Scale Development: Theory and Applications* (3rd ed.), Thousand Oaks, CA: Sage Publications.

Drazin, R., & Van de Ven, A. H. (1985). Alternative forms of fit in contingency theory. *Administrative science quarterly*, 514-539.

Edwards, J. R. (2001). Multidimensional constructs in organizational behavior research: An integrative analytical framework. *Organizational research methods*, *4*(2), 144-192.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing*, *62*(1), 107-115.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 39-50.

Fosso Wamba, S., Angappa, G., Papadopoulos, T., & Ngai, E. (2018). Big data analytics in logistics and supply chain management. *International Journal of Logistics Management, The*, (just-accepted), 00-00.

Fosso Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356-365.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049-1064.

Hahn, G. J., & Packowski, J. (2015). A perspective on applications of in-memory analytics in supply chain management. *Decision Support Systems*, *76*, 45-52.

Hinkin, T. R., & Tracey, J. B. (1999). An analysis of variance approach to content validation. *Organizational Research Methods*, *2*(2), 175-186.

Hoehle, H., & Venkatesh, V. (2015). Mobile Application Usability: Conceptualization and Instrument Development. *Mis Quarterly*, *39*(2).

Kache, F., & Seuring, S. (2017). Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. *International Journal of Operations & Production Management*, *37*(1), 10-36.

Karimi, J., Somers, T. M., & Bhattacherjee, A. (2007). The role of information systems resources in ERP capability building and business process outcomes. *Journal of Management Information Systems*, *24*(2), 221-260.

Kim, G., Shin, B., & Kwon, O. (2012). Investigating the value of sociomaterialism in conceptualizing IT capability of a firm. *Journal of Management Information Systems*, *29*(3), 327-362.

Krippendorff, K. (2012). *Content analysis: An introduction to its methodology*. Sage Publications.

Liu, P., & Yi, S. P. (2017). A study on supply chain investment decision-making and coordination in the Big Data environment. *Annals of Operations Research*, 1-19.

MacKenzie, S. B. (2003). The dangers of poor construct conceptualization. *Journal of the Academy of Marketing Science, 31*(3), 323-326.

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS quarterly*, *35*(2), 293-334.

Makadok, R. (2001). Toward a synthesis of the resource-based and dynamic-capability views of rent creation. *Strategic management journal*, *22*(5), 387-401.

Pavlou, P. A., & El Sawy, O. A. (2010). The "third hand": IT-enabled competitive advantage in turbulence through improvisational capabilities. *Information systems research*, *21*(3), 443-471.

Perreault, W. D., & Leigh, L. E. (1989). Reliability of nominal data based on qualitative judgments. *Journal of marketing research*, *26*(2), 135.

Rai, A., Pavlou, P. A., Im, G., & Du, S. (2012). Interfirm IT capability profiles and communications for cocreating relational value: evidence from the logistics industry. *MIS quarterly*, *36*(1), 233-262.

Ravichandran, T., & Lertwongsatien, C. (2005). Effect of information systems resources and capabilities on firm performance: A resource-based perspective. *Journal of management information systems*, *21*(4), 237-276.

Richey Jr, R. G., Morgan, T. R., Lindsey-Hall, K., & Adams, F. G. (2016). A global exploration of big data in the supply chain. *International Journal of Physical Distribution & Logistics Management*, *46*(8), 710-739.

Roberts, N., & Grover, V. (2012). Leveraging information technology infrastructure to facilitate a firm's customer agility and competitive activity: An empirical investigation. *Journal of Management Information Systems*, *28*(4), 231-270.

Ryan, G.W., & Bernard, H. R. (2000). Data management and analysis methods. N. Denzin, Y.

Lincoln, eds. *Handbook of Qualitative Research*, 2nd ed. Thousand Oaks, CA, Sage, 769-802.

Sanders, N. R. (2016). How to use big data to drive your supply chain. *California Management Review*, *58*(3), 26-48.

Schoenherr, T., & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: Current state and future potential. *Journal of Business Logistics*, *36*(1), 120-132.

Souza, G. C. (2014). Supply chain analytics. *Business Horizons*, *57*(5), 595-605.

Straub, D., Boudreau, M. C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *The Communications of the Association for Information Systems*, *13*(1), 63.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research techniques*. Sage publications.

Trkman, P., McCormack, K., De Oliveira, M. P. V., & Ladeira, M. B. (2010). The impact of business analytics on supply chain performance. *Decision Support Systems*, *49*(3), 318-327.

Urquhart, C., Lehmann, H., & Myers, M. D. (2010). Putting the ‘theory’ back into grounded theory: guidelines for grounded theory studies in information systems. *Information systems journal*, *20*(4), 357-381.

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77-84.

Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, *176*, 98-110.

Wang, N., Liang, H., Zhong, W., Xue, Y., & Xiao, J. (2012). Resource structuring or capability building? An empirical study of the business value of information technology. *Journal of Management Information Systems*, *29*(2), 325-367.

Wang, Y., & Hajli, N. (2017). Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, *70*, 287-299.

Wang, Y., Kung, L., Wang, W. Y. C., & Cegielski, C. G. (2018). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management*, *55*(1), 64-79.

Wetzels, M., Odekerken-Schröder, G., & Van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS quarterly*, 177-195.

Yao, G., Wu, C. H., & Yang, C. T. (2008). Examining the content validity of the WHOQOL-BREF from respondents' perspective by quantitative methods. *Social Indicators Research*, *85*(3), 483-498.

Zhu, S., Song, J., Hazen, B. T., Lee, K., & Cegielski, C. (2018). How supply chain analytics enables operational supply chain transparency: An organizational information processing theory perspective. *International Journal of Physical Distribution & Logistics Management*, *48*(1), 47-68.

Appendices for Essay II

Appendix 1
Advanced data modeling recently helped one global chemical maker to cut through all these problems in its flagship plant. Company experts in sales, production, and optimization assembled raw-material and product price curves, market-size forecasts, historical equipment-performance data, and more than 600 decision variables into a mathematical model describing the plant's production yields under various operating conditions. The resulting model offers managers a precise understanding of the effects that variations anywhere along the value chain can have on the production network as a whole. The company can now, for example, easily fine-tune the mix of raw materials and finished products, as well as the routing of manufacturing flows, in real time -- while constantly identifying opportunities for improvement (exhibit). All told, these changes increased the plant's EBITreturns by more than 50 percent, and the company is now applying this model across its full factory network so that production capacity can shift in modular fashion.
Category: Supply chain analytics capability–Predictive Capability in Production.
Subcategory: predict the effects of variations in operating conditions along the value chain on production yields.

Appendix 2
Initial Items Based on the Content Validity Check

Initial Items for Tracking Capability in Procurement (TrPC)
TrPC1 Big data analytics enables us to access real-time information about the status of our purchase orders.
TrPC2 Big data analytics enables us to access real-time information about any variations of purchase orders.
TrPC3 Big data analytics enables us to track inbound activities of items purchased from suppliers.
TrPC4 Big data analytics enables us to continuously track suppliers' progress (e.g. lot acceptance, test results, etc.) on our purchase orders.
Initial Items for Tracking Capability in Production (TrPD)
TrPD1 Big data analytics enables us to access real-time production data of production orders.
TrPD2 Big data analytics enables us to access real-time status of the production process.
TrPD3 Big data analytics enables us to track real-time performance indicators of the production process.
TrPD4 Big data analytics enables us to track real-time states of component parts (e.g. operators, machines, etc.) of the production process.
Initial Items for Tracking Capability in Distribution (TrD)
TrD1 Big data analytics enables us to instantly access resource availability (e.g. labor availability) in distribution centers.
TrD2 Big data analytics enables us to access real-time operational information about material handling systems (e.g. conveyors, lift trucks, etc.) in distribution centers.
TrD3 Big data analytics enables us to access real-time data on activities of various functions in distribution centers.
TrD4 Big data analytics enables us to access real-time information about inventory through every stage of handling in distribution centers.
Initial Items for Traceability in Transportation (TrT)
TrT1 Big data analytics enables us to access real-time information on vehicles' condition (e.g. location, engine health, etc.) during transportation.
TrT2 Big data analytics enables us to access real-time information on cargo status (e.g. location, transit duration, etc.) from origin to destination.
TrT3 Big data analytics enables us to real-time track external environment (e.g. weather, traffic congestion, etc.) for each time of transportation.
TrT4 Big data analytics enables us to track employees' real-time driving behavior during transportation.
TrT5 Big data analytics enables us to access real-time status of transportation assets (e.g. trailers, refrigerated units, etc.).
Initial Items for Analytical Capability in Procurement (ACPC)
ACPC1 Big data analytics enables us to benchmark performance across suppliers to pinpoint improvement areas for suppliers' products and services
ACPC2 Big data analytics enables us to attribute issues of purchased products and services to specific suppliers.

ACPC3	Big data analytics enables us to benchmark procurement performance against other companies to identify improvement areas for the procurement processes.
ACPC4	Big data analytics enables us to calculate "should-cost" amount per SKU to identify opportunities of negotiation with suppliers.
ACPC5	Big data analytics enables us to analyze spend data across suppliers to uncover hidden saving opportunities.
ACPC6	Big data analytics enables us to connect customer complaints about product performance with specific suppliers.
Initial Items for Analytical Capability in Production (ACPD)	
ACPD1	Big data analytics enables us to identify root causes of variability in the production process.
ACPD2	Big data analytics enables us to identify root causes of inefficiency in the production process.
ACPD3	Big data analytics enables us to identify root causes of producing scrap (defective products).
Initial Items for Analytical Capability in Distribution (ACD)	
ACD1	Big data analytics enables us to explain why performance parameters (e.g. packing performance) are not met in distribution centers.
ACD2	Big data analytics enables us to analyze historical and real-time operational data from material handling equipment to identify the causes of anomalies in the equipment in distribution centers.
ACD3	Big data analytics enables us to identify the root causes of performance differences among employees within distribution centers.
ACD4	Big data analytics enables us to identify the root causes of errors (e.g. picking errors) in distribution centers.
Initial Items for Analytical Capability in Transportation (ACT)	
ACT1	Big data analytics enables us to identify the potential sources of cargo spoilage or loss in the transportation process.
ACT2	Big data analytics enables us to identify reasons for deteriorating performance indicators in the transportation process.
ACT3	Big data analytics enables us to connect product performance issues to the transportation process.
Initial Items for Analytical Capability in Product Design and Development (ACPDD)	
ACPDD1	Big data analytics enables us to mine data from online conversations about products (e.g. product reviews) to assess customers' sentiments at product-feature level.
ACPDD2	Big data analytics enables us to analyze online buzz about product features to evaluate their importance to customers.
ACPDD3	Big data analytics enables us to mine online conversations about products to identify product features that could add value to us.
ACPDD4	Big data analytics enables us to convert online buzz into meaningful metrics at the product feature level.
Predictive Capability in Demand Management (PCDM)	
PCDM1	Big data analytics enables us to leverage structured and unstructured data (e.g. historical transactions, search data, etc.) to anticipate changes in customer demand.
PCDM2	Big data analytics enables us to integrate data from multiple sources (i.e. internal and external) to more accurately forecast customer demand.

PCDM3	Big data analytics enables us to combine structured data with unstructured data to anticipate what customers are going to purchase.
PCDM4	Big data analytics enables us to use real-time data (e.g. usage data) to predict replenishment needs at customer locations.
Initial Items for Predictive Capability in Procurement (PCPC)	
PCPC1	Big data analytics enables us to predict the occurrence of negative events (e.g. delay in shipments, natural disasters, etc.) on supplier sides.
PCPC2	Big data analytics enables us to predict the impact of negative events occurred to suppliers on our firm's performance (e.g. revenue, service level, etc.).
PCPC3	Big data analytics enables us to perform "what-if" analysis to forecast the effectiveness of various alternatives to negative events on the supplier side.
PCPC4	Big data analytics enables us to predict the impact of various sourcing alternatives (e.g. different supplier mixes, etc.) on our company's performance.
Initial Items for Predictive Capability in Production (PCPD)	
PCPD1	Big data analytics enables us to discover patterns from real-time machine data to predict maintenance/repair needs for production equipment.
PCPD2	Big data analytics enables us to discover correlations from real-time and historical data along production line to forecast quality related defects.
PCPD3	Big data analytics enables us to predict the impact of negative events in the production process (e.g. machine failure, etc.) on production performance.
PCPD4	Big data analytics enables us to discover patterns from real-time and historical data to predict the occurrence of undesired events in the production process (e.g. process failure, etc.).
Initial Items for Predictive Capability in Distribution (PCD)	
PCD1	Big data analytics enables us to predict maintenance or repair needs for material handling equipment in distribution centers.
PCD2	Big data analytics enables us to combine real-time with historical data to predict order processing progress in distribution centers (e.g. the completion time of zones).
PCD3	Big data analytics enables us to model real-time and historical data to predict what the order fulfillment profile will look like in distribution centers (e.g. labor and inventory needs, etc.).
PCD4	Big data analytics enables us to discover patterns from real-time data to predict bottlenecks within the processes (e.g. picking, packing) in distribution centers.
Initial Items for Predictive Capability in Product Design and Development (PCPDD)	
PCPDD1	Big data analytics enables us to combine internal with external data (e.g. historical transactions, web browsing, etc.) to generate predictive models of product design trends.
PCPDD2	Big data analytics enables us to combine internal data with external data to predict next-trend products favored by customers.
PCPDD3	Big data analytics enables us to cross-reference internal data (e.g. customer transactions, etc.) with external data (e.g. social media opinions, etc.) to predict products with vast potential for development.
PCPDD4	Big data analytics enables us to discover patterns from data of launched products (e.g. usage data) to predict design changes in the next model.
Initial Items for Predictive Capability in Transportation (PCT)	
PCT1	Big data analytics enables us to leverage real-time and historical data to predict

	negative events (e.g. traffic congestion, etc.) that will affect the transport of freights.
PCT2	Big data analytics enables us to quantify freight risks based on internal and external data (e.g. shipping data, public data, etc.).
PCT3	Big data analytics enables us to accurately predict the impacts of negative events (e.g. hub congestion, etc.) occurred during the transport of freights.
PCT4	Big data analytics enables us to perform what-if analysis to predict the effectiveness of different alternatives in response to negative events during the transport of freights.
Initial Items for Decision Support Capability in Demand Management (DSCDM)	
DSCDM1	Big data analytics enables us to dynamically adjust operations to meet demand fluctuations.
DSCDM2	Big data analytics enables us to dynamically determine the best fulfillment plan that satisfies customers' changing demands.
DSCDM3	Big data analytics enables us to quickly modify upstream decisions (e.g. production, pricing decisions, etc.) according to real-time demand information.
DSCDM4	Big data analytics enables us to translate real-time demands into optimal replenishment plans.
Initial Items for Decision Support Capability in Procurement (DSCPC)	
DSCPC1	Big data analytics enables us to quickly generate optimal alternatives in response to changes (e.g. disruptive events, etc.) arising from suppliers.
DSCPC2	Big data analytics enables us to visualize risks posed to the supply network if any suppliers fail to perform.
DSCPC3	Big data analytics enables real-time notifications about changes in suppliers (e.g. suppliers' shipment delay) so that adjustments can be made.
DSCPC4	Big data analytics enables us to optimize supplier selection under uncertainty.
DSCPC5	Big data analytics enables us to find the optimal sourcing solutions that balance supplier objectives and purchasing constraints.
Initial Items for Decision Support Capability in Production (DSCPD)	
DSCPD1	Big data analytics enables us to build optimal maintenance plans around production equipment.
DSCPD2	Big data analytics enables real-time notification about any deviations from planned production process objectives so that remediation can be initiated to production operations.
DSCPD3	Big data analytics recommends us optimal solutions of maximizing production performance.
DSCPD4	Big data analytics enables us to automatically make process adjustment for any stage of production.
DSCPD5	Big data analytics informs us of how to optimize production operations (e.g. equipment operation procedures, etc.) in real time.
Initial Items for Decision Support Capability in Distribution (DSCD)	
DSCD1	Big data analytics enables us to make real-time adjustments to processes (e.g. storing, picking, etc.) in distribution centers.
DSCD2	Big data analytics helps us optimize inventory placement in distribution centers.
DSCD3	Big data analytics enables us to make real-time resource (re)allocation decisions that optimize distribution operations.
DSCD4	Big data analytics enables us to build optimal maintenance schedule around material

	handling equipment in distribution centers.
Initial Items for Decision Support Capability in Transportation (DSCT)	
DSCT1	Big data analytics provides real-time route optimization for freight transport according to numerous factors (e.g. weather, safety, etc.).
DSCT2	Big data analytics provides means for notifying decision makers about any issues (e.g. cargos derivate from predetermined route, etc.) arising along the way that may prevent us from delivery as promised to a customer.
DSCT3	Big data analytics allows us to optimize maintenance based on the health of transportation assets (e.g. trucks, containers, trailers, etc.).
DSCT4	Big data analytics gives drivers instant recommendations on safe driving practices during transportation.
DSCT5	Big data analytics allows us to optimize fuel efficiency during transportation.
Initial Items for Decision Support Capability in Supply Chain Network Design (DSCSCN)	
DSCSCN1	Big data analytics provides us means for determining the optimal number and locations of facilities (e.g. manufacturing plants, warehouses, etc.) within the supply chain network.
DSCSCN2	Big data analytics enables us to visualize node information (e.g. space requirements, product flow paths, etc.) to identify supply chain network optimization areas.
DSCSCN3	Big data analytics supports decisions determining the optimal supply chain flow paths.
DSCSCN4	Big data analytics enables us to optimize the supply chain network in terms of different objectives (e.g. cost, service levels, etc.).
DSCSCN5	Big data analytics helps us determine supply chain network configurations in terms of various objectives (e.g. cost, service levels, etc.).
Decision Support Capability in Product Design and Development (DSCPDD)	
DSCPDD1	Big data analytics provides insights to inform the design and development of new products.
DSCPDD2	Big data analytics provides insights to inform the modification of product design and development.
DSCPDD3	Big data analytics helps us determine specific product design requirements (e.g. quality, reliability) that meet customers’ requirements.
DSCPDD4	Big data analytics helps us determine the optimal product design option that meets various criteria (e.g. quality, reliability, etc.).
DSCPDD5	Big data analytics helps us determine how close we are to achieving the objectives during the design and development cycles.

Appendix 3
Content Validity: Proportion of Substantive Agreement and Substantive Validity Coefficient

Construct Name	Label	P_{SA}	C_{SV}	Construct Name	Label	P_{SA}	C_{SV}
Traceability in Procurement	TrPC1	0.89	0.78	Predictive Capability in Production	PCPD1	1	1
	TrPC2	1	1		PCPD2	1	1
	TrPC3	0.89	0.78		PCPD3	1	1
	TrPC4	0.89	0.78		PCPD4	1	1
Traceability in Production	TrPD1	1	1	Predictive Capability in Distribution	PCD1	0.89	0.78
	TrPD2	1	1		PCD2	0.78	0.56
	TrPD3	1	1		PCD3	0.78	0.67
	TrPD4	1	1		PCD4	0.78	0.56
Traceability in Distribution	TrD1	0.33	0		PCD5	0.56	0.11
	TrD2	0.89	0.78	Predictive Capability in Transportation	PCT1	1	1
	TrD3	0.67	0.56		PCT2	1	1
	TrD4	1	1		PCT3	1	1
	TrD5	1	1		PCT4	1	1
Traceability in Transportation	TrT1	0.67	0.33	Predictive Capability in Product Design and Development	PCDD1	1	1
	TrT2	0.89	0.78		PCDD2	0.89	0.78
	TrT3	0.89	0.78		PCDD3	1	1
	TrT4	0.89	0.78		PCDD4	1	1
	TrT5	1	1	Decision Support Capability in Demand Management	DSCDM1	1	1
Analytical Capability in Procurement	ACPC1	0.67	0.56		DSCDM2	1	1
	ACPC2	0.67	0.44		DSCDM3	1	1
	ACPC3	0.89	0.78		DSCDM4	0.56	0.33
	ACPC4	1	1		DSCDM5	0.89	0.78
	ACPC5	1	1	Decision Support Capability in Procurement	DSCPC1	0.89	0.78
	ACPC6	1	1		DSCPC2	1	1
Analytical Capability in Production	ACPD1	1	1		DSCPC3	0.89	0.78
	ACPD2	0.89	0.78		DSCPC4	1	1
	ACPD3	0.89	0.78		DSCPC5	1	1
Analytical Capability in Distribution	ACD1	0.78	0.67	Decision Support Capability in Production	DSCPD1	1	1
	ACD2	0.67	0.44		DSCPD2	0.78	0.67
	ACD3	1	1		DSCPD3	1	1
	ACD4	0.78	0.67		DSCPD4	1	1
Analytical Capability in Transportation	ACT1	0.89	0.78		DSCPD5	1	1
	ACT2	1	1	Decision Support Capability in Distribution	DSCD1	0.89	0.78
	ACT3	0.89	0.78		DSCD2	1	1
Analytical Capability in Product Design and Development	ACPDD1	1	1		DSCD3	0.89	0.78
	ACPDD2	1	1		DSCD4	1	1
	ACPDD3	1	1	Decision Support Capability in Transportation	DSCT1	1	1
	ACPDD4	0.89	0.78		DSCT2	1	1
Predictive	PCDM1	1	1		DSCT3	1	1

Capability in Demand Management	PCDM2	0.89	0.78		DSCT4	0.89	0.78
	PCDM3	0.78	0.56		DSCT5	1	1
	PCDM4	0.89	0.78		DSCT6	0.44	0.22
Predictive Capability in Procurement	PCPC1	0.89	0.78	Decision Support Capability in Product Design and Development	DSCPDD1	1	1
	PCPC2	1	1		DSCPDD2	1	1
	PCPC3	0.89	0.78		DSCPDD3	1	1
	PCPC4	1	1		DSCPDD4	0.67	0.56
	PCPC5	0.56	0.33		DSCPDD5	1	1
				Decision Support Capability in Supply Chain Network Design	DSCSCN1	1	1
					DSCSCN2	1	1
					DSCSCN3	1	1
					DSCSCN4	1	1
					DSCSCN5	1	1

Essay III

Impact of Big Data Analytics Initiative on Operational Efficiency and Business Growth

1. Introduction

Big data analytics (BDA) refers to the use of advanced analytical techniques against data characterized by 5 Vs (i.e. volume, variety, velocity, veracity, and value) to gain actionable insights for delivering sustained value and building competitive advantages (Fosso Wamba, Akter, Edwards, Chopin, and Gnanzou, 2015; Fosso Wamba, Gunasekaran, Akter, Ren, Dubey, and Childe, 2017). Organizations are increasingly implementing big data analytics (BDA) to improve performance. Both anecdotal surveys and empirical research show that BDA is primarily used in the areas of new product development or engineering, operations, marketing and sales, customer relationship management, and fraud and compliance (Datameer, 2016; International Institute for Analytics, 2016; NewVantage, 2017) to reduce expenses, find new innovation avenues, increase revenues, launch new product or service offerings, accelerate the speed of current efforts, and transform business for the future (Barn, 2017; Chen, Preston, and Swink, 2015; Ghasemaghaei, Hassanein, and Turel, 2017b; Popovič, Hackney, Tassabehji, and Castelli, 2018; Wang, Kung, Wang, and Cegielski, 2018). However, evidence regarding BDA benefits has been predominantly survey-based or case-based. Respondents usually provide perceived benefits of their organizations' BDA efforts and only successful BDA implementation cases are selected to be reported. There is a lack of more objective evidence on the impact of organizations' BDA initiative on organizations' actual performance.

The value created by BDA depends on the strategic objectives of using BDA (Ghoshal, Larson, Subramanyam, and Shaw, 2014). The strategic goals for firms' use of BDA relate to two aspects: improving established business models through incremental enhancement and innovating business models (Günther, Mehrizi, Huysman, and Feldberg, 2017; Woerner and Wixom, 2015). For the first aspect, firms invest in BDA to enhance the efficiency of existing operations through reducing cost, eliminating waste, or appropriating staffing. For the second aspect, firms invest in BDA to grow and innovate businesses through discovering new value avenues, targeting new customers, or developing new products and services. The goals of BDA initiative form two measures of BDA value: operational efficiency and business growth. Accordingly, the first question this study aims to answer is: *What is the influence of organizations' BDA initiative on organizations' performance in terms of operational efficiency and business growth?*

Current research seems to implicitly assume the value of BDA is universal. In other words, BDA is positively associated with all measures of performance. Nevertheless, many firms have heavily invested in BDA but still have not realized the outcomes they desired (Brown and Gottlieb, 2016; Henke, Bughin, and Chui, 2016), which indicates that the success of BDA projects is contingent on certain circumstances. My study selects industry environment as an important circumstance that may influence BDA value generation. Such selection is motivated by the Information Systems (IS)-business strategy alignment literature, which emphasizes that decision makers should align IT objectives with business strategies (Sabherwal and Chan, 2001), and by the structure-conduct-performance paradigm, which suggests that a company's behavior and performance is influenced by the structure of industry within in which the company competes (Bain, 1968). Prior studies (e.g. Chen et al., 2015) have examined the role of industry environment in moderating the impact of perceived BDA usage on perceived organizational performance. Few studies have explored how industry environment moderates the objective performance outcomes associated with reported BDA initiative. Thus, the second research question this study aims to address is: *What is the effect of industry environment on moderating the relationships of BDA initiative with operational efficiency and business growth?*

I draw on the dynamic capability perspective and conceptualize organizational BDA initiative as a dynamic information processing capability which will bring competitive advantages to organizations. Following the strategic literature (e.g. Dess and Beard, 1984; Keats and Hitt, 1988), I conceptualize industry environment in terms of environmental dynamism, munificence, and complexity. This study solves the research questions using multiple-industry panel data collected from Nexus Uni and COMPUSTAT databases. Two dynamic panel data models are constructed to test the effect of BDA initiative on performance and how industry environment moderates the BDA-performance relationships. System generalized method of moments (GMM) is selected to estimate the models. The analysis indicates that: organizational BDA initiative enhances operational efficiency and facilitates business growth; at high level of environmental complexity, BDA initiative is associated with a greater increase in both operational efficiency and business growth; at lower level of industry dynamism and munificence, BDA initiative has a greater positive impact on operational efficiency; at higher level of industry dynamism, BDA initiative is associated with a greater increase in business growth. These findings provide a theory-based understanding about the economic benefits of BDA as well as offer guidance regarding what practitioners can expect from BDA initiative and how firms can realize value from BDA given the characteristics of industries they are operating in.

The remainder of this paper is organized as follows. Section 2 reviews the literation on BDA initiative and rationalizes BDA initiative as organizational dynamic capability. Section 3 develops the research model and hypotheses. Section 4 describes the data collection, variable operationalization, analysis method, and empirical results. Section 5 discusses the theoretical and practical implications of this research. Section 6 and 7 include the research limitation, future research directions, and conclusion.

2. Theoretical Background

2.1 Big Data Analytics Initiative

BDA is the use of advanced analytic techniques against data characterized by 5 Vs (i.e. volume, variety, velocity, veracity, and value) to gain actionable insights for delivering sustained value, measuring performance, and building competitive advantages (Fosso Wamba et al., 2015; Fosso Wamba et al., 2017). With a view to embrace the infinite pool of data to enhance decision making, firms increasingly adopt BDA for various business purposes such as decreasing expenses, adding revenue, and launching new products and services (Bean, 2017), which in this research refers to BDA initiative. With references to previous studies (e.g. Datameer, 2016; International Institute for Analytics, 2016; NewVantage, 2017), I categorize firms' BDA initiative into five classes: (1) new product development/engineering, (2) operations, (3) customer service and customer relationship management, (4) sales and marketing, and (5) fraud and compliance.

Although researchers have paid increasing attention to the big data phenomenon, the empirical evidence on the benefits of BDA has been primarily survey-based or case-based evidence (Table 1). Specifically, a number of studies examine the business value of BDA based on data collected through surveys, where respondents rate their perceived use of BDA and perceived benefits from BDA efforts. For example, drawing on survey data from 161 U.S. firms, Chen, Preston, and Swink (2015) investigated the key drivers of BDA use and the impacts of BDA use on two key components of value creation (i.e. asset productivity and business growth).

Gupta and George (2016) developed and validated an instrument for measuring BDA capability and tested the relationship between BDA capability and business performance through surveying BDA managers and executives.

In addition, much case-level evidence has begun to emerge from BDA literature about how BDA generates value for specific firms in different industries. For instance, through 36 case descriptions of BDA implementation in healthcare organizations, Wang, Kung, Wang, and Cegielski (2018) revealed the key components, links, and path-to-value chains that explain the path of how BDA can be leveraged to deliver business value in healthcare setting. Based on in-depth analysis of case studies conducted with four companies, Gunasekaran, Yusuf, Adeleye, and Papadopoulos (2017) found that the integrative deployment of big data business analytics leads to agile manufacturing enablers which in turn improve competitiveness and business performance. Existing research on BDA value creation for organizations focuses either on adopters' perceptions about BDA usage and benefits or on exploratory studies using qualitative case study approach. Few studies, except Teo, Nishant, Koh (2016) and Tambe (2014), concentrate on utilizing quantitative research methods to explore the impact of BDA initiative on firms' actual performance in the long term. My study is one of the initial attempts to quantitatively investigate the actual impact of BDA initiative at the firm level.

Table 1. Current Empirical Research on BDA Value Creation for Organizations

Study	Evidence Type	Data Analysis	Summary
Aker et al. (2016)	Survey-based	Partial Least Squares	Akter et al. leverage resource-based view and sociomaterialism perspective to conceptualize BDA capability as a high-order construct that consists of three dimensions. In addition, they examine the direct impact of BDA capability on firm performance as well as the moderating effect of analytics capability-business strategy alignment on the BDA capability-performance relationship.

Chae et al. (2014a)	Survey-based	Partial Least Squares	Drawing on resource-based view, Chae et al. conceptualize supply chain analytics as the integration of three resources: data management resources, IT-enabled supply chain planning resources, and performance management resources. They explore the relationships among these resources, supply chain planning satisfaction, and operational performance.
Chae et al. (2014b)	Survey-based	Partial Least Squares	Chae et al. examine the impact of two business analytics resources: accurate data and advanced analytics on manufacturers' operational performance. They also investigate the role of complementary resources by testing the moderating and mediating effect of SCM initiatives on the relationships between primary resources and performance.
Chen et al. (2015)	Survey-based	Partial Least Squares	Motivated by the fact that not many firms capitalize on BDA, Chen et al. develop and test a research model showing the key drivers of BDA usage and the influence of BDA usage on two key components of value creation: asset productivity and business growth.
Corte-Real et al. (2017)	Survey-based	Partial Least Squares	Corte-Real et al. examine the paths of how BDA leverage knowledge assets to create business value for European firms.
Ghasemaghaei et al. (2017a)	Survey-based	Partial Least Squares	Ghasemaghaei et al. conceptualize, operationalize, and validate the concept of Data Analytics Conceptency, and empirically examine the impact of Data Analytics Competency on decision making performance (i.e. decision efficiency and decision quality). The findings demonstrate significant and positive relationship between data analytics competency and decision -making performance.
Ghasemaghaei et al. (2017b)	Survey-based	Partial Least Squares and ANOVA	Ghasemaghaei et al. draw on dynamic capability theory to examine the impact of data analytics use on organizational agility and leverage the fit perspective to study when and how the effect of using data analytics accrues.

Gunasekaran et al. (2017b)	Case-based	Qualitative Research Method	Based on in-depth multiple case studied conducted with four organizations, Gunasekaran et al. develop and validate a theoretical framework for understanding the role of big data business analytics in agile manufacturing practices, given a particular level of market turbulence.
Gupta and George	Survey-based	Partial Least Squares	Gupta and George identify resources for building a firm's BDA capability, develop an instrument to measure BDA capability of the firm, and empirically test the relationship between BDA capability and business performance. The results validate the instrument and provide evidence that BDA capability leads to superior business performance.
Ji-fan Ren et al. (2017)	Survey-based	Partial Least Squares-Structural Equation Modeling	Ren et al. examine the impact of big data analytics quality dynamics on both business value and firm performance.
Krishnamoorthi and Mathew (2018)	Case-based	Qualitative Research Method	Krishnamoorthi and Mathew theorize on how BA contribute to business value of firms. Drawing on multiple case studies, they identify key concepts constituting Analytics Technology Assets, BA Capability, as well as the mechanism through which business analytics contributes to business performance. Based on case study results, they propose a research model that reveals key predictors and moderators and their relationships with business performance.
Seddon et al. (2017)	Case-based	Qualitative Research Method	Seddon et al. present and preliminarily assess a model that shows how business analytics contributes to business value. The model consists of two parts: a process and a variance model. The results suggest the business analytics success model is likely to be a useful basis for future research.
Tambe (2014)	Secondary data from LinkedIn and COMPUSTAT	Complementary Test, Regression analysis	Tambe investigates how labor market factors shape early returns on big data investment in terms of Hadoop-based systems.

	database		
Teo et al. (2016)	Secondary data on business analytics announcements and market reactions		Teo et al. investigate market reaction to announcements of business analytics news or events. Specifically, they examine whether type of BA vendor, type of BA, the extent of implementation, firm-specific characteristics, and stock characteristics influence shareholders' reaction to BA announcements.
Trkman et al. (2010)	Survey-based	Partial Least Squares	Trkman et al. examine the impact of business analytics in key supply chain processes on supply chain performance. They also study the moderating role of information systems support and business process orientation on analytics-performance relationships.
Fosso Wamba et al. (2015)	Case-based	Qualitative Research Method	Through a comprehensive literature review of "big data" articles, Fosso Wamba et al. provide a general taxonomy that helps understand the business value of big data and issues of value creation from big data. The in-depth analysis of a longitudinal case study reveals insights into how to implement big data projects to achieve business value in emergency service environments.
Fosso Wamba et al. (2017)	Survey-based	Partial Least Squares-Structural Equation Modeling	Drawing on the resource-based view, literature on BDA and IT capability, Fosso Wamba et al. propose a conceptual model for measuring big data analytics capability and examine the direct effect and indirect effect of BDA capability on firm performance.
Wang and Hajli (2017)	Case-based	Qualitative Research Method	Drawing on the resource-based view, IT capability building view, and multidimensional benefit framework, Wang and Hajli propose a big data analytics-enabled business value model to explain how big data analytics capabilities are developed and the potential benefits obtained from these capabilities in the healthcare context. The proposed model is validated through the analysis of 109 case descriptions of big data analytics implementation in healthcare organizations.

Wang et al. (2018)	Case-based	Qualitative Research Method	Based on Practice-based View, Wang et al. propose a BDA-enabled transformation model that links BDA capabilities with IT-enabled transformation practices and then with benefit dimensions and business values. The model is tested in healthcare setting by analyzing big data implementation cases. The results reveal essential elements (BDA capabilities, practices, benefits, and values), links, and path-to-value chains specific to the healthcare industry.
Xie et al. (2016)	Case-based	Qualitative Research Method	Xie et al. propose a theoretical framework of how big data transforms from resources into cooperative assets to promote value co-creation between firms and customers. This study provides a theoretical perspective on how big data interconnects customers and firms in promoting value co-creation.

2.2 Big Data Analytics Initiative as Enterprise Dynamic Capabilities

Researchers have developed the dynamic capability perspective to describe "firms' ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments" (Teece, Pisano, and Shuen, 1997). Dynamic capabilities can be disaggregated into three capacities: the ability to (1) sense and shape opportunities and threats, (2) seize opportunities, and (3) maintain competitiveness through reconfiguring enterprises' tangible and intangible assets (Teece, 2009). Dynamic capabilities can also be described as firms' identifiable and specific processes that use resources to match and even make market changes (Eisenhardt and Martin, 2000). Such processes are collections of firms' routines that focus on integrating, reconfiguring, gaining, and releasing resources. Dynamic capabilities are therefore firms' routines through which "firms achieve new resource configurations as markets emerge, collide, split, evolve, and die" (Eisenhardt and Martin, 2000). The new resource configurations are the sources of competitive advantages for firms within in dynamic markets.

We conceptualize organizational BDA initiative as dynamic capabilities for two reasons. First, BDA initiative helps establish, particularly for firms operating in dynamic markets, knowledge derivation and decision-making routines that convert data into information that leads to knowledge that leads to decisions that lead to value-creating actions (Abbasi, Sarker, and Chiang, 2016; Seddon, Constantinidis, Tamm, and Dod, 2017). The knowledge derivation and decision-making routines are dynamic capabilities identified in the literature (Eisenhardt and Martin, 2000). Moreover, such routines create business value by capitalizing on existing organizational resources or changing (adding or dropping) organizational resources (Seddon et al., 2017). The value of BDA initiative thus lies in the resource configurations that they create.

Second, the three key capacities of dynamic capability are reflected widely in both academic and industry's claims about BDA. The capacity to sense and shape opportunities and threats focuses on gathering and filtering market, technology, and competitive information from both inside and outside of the organization, making sense of it, and extracting implications for actions (Teece, 2009). Prior literature has emphasized BDA's ability to combine data from different sources, which allows enterprises to have more-panoramic and more-granular views of their business environment (Barton and Court, 2012). For example, companies are increasingly analyzing real-time data related to suppliers to gain insight into suppliers' technical capabilities, financial health, weather and political risks, and so forth so that proper actions are taken in response to potential disruptions (Davenport, 2014).

Additionally, the capacity to seize opportunities relates to making sound decisions under uncertainty and executing well on the decisions (Teece, 2009). A number of studies on big data are about its role in "replacing/supporting human decision making with automated algorithms" (see a review from Fosso Wamba et al. (2015)). Through improved decision making within organizations, organizations are able to handle uncertainties and variabilities (Wang, Gunasekaran, Ngai, and Papadopoulos, 2016) and act faster and more wisely (Woerner and Wixom, 2015). Last, reconfiguring enterprises' tangible and intangible assets involves business model redesign, asset re-alignment, and the revamping of routine to adapt to the dynamic environment (Teece, 2009). Previous literature has indicated BDA's potential to transform the way companies do business (Barton and Court, 2012; McAfee and Brynjolfsson, 2012). One way of using BDA is through accessing new data and insight, and consequently acting to refine and optimize established processes in terms of efficiency and effectiveness (Loebbecke and Picot, 2015; Woerner and Wixom, 2015). Companies are also using BDA to innovate business models through monetizing insights by selling, bartering, and wrapping, or by creating or moving to new industries (Woerner and Wixom, 2015). Therefore, viewing BDA initiative as dynamic capabilities makes theoretical sense and such view helps understand the impact of BDA initiative on value creation.

3. Hypothesis Development

3.1 BDA Initiative and Operational Efficiency

In accordance with the dynamic capability perspective (Eisenhardt and Martin, 2000), I view BDA initiative as a way that organizations adopt to enhance organizational information processing capability (Chen et al., 2015) that enables them to collect, integrate, and convert fast-moving data of large volumes and variety into more comprehensive and accurate insights and to synthesize insights to decision makers within different functional departments. The insights from BDA support a wide range of aspects in relation to firms' day-to-day operations. For example, through a comparative of case study of three manufacturing firms, Popovič, Hackney, Tassabehji, and Castelli (2018) find that insights from BDA enable firms to better predict potentially unfavorable events, defective products, and demand changes, improve schedule and cost variance, maximize equipment uptime, real-time monitor business processes, reduce production resource waste, and optimize capacity utilization. The utilization of BDA also enables optimal operational shift planning by appropriate staffing for efficient output (Addo-Tenkorang and Helo, 2016; Jeske, Grüner, and Weiß, 2013). Moreover, BDA enabled fact-based decision making empowers employees to reconfigure operations/processes in order to timely respond to unplanned events or changes in the environment (Popovič et al., 2018). As a result, firms realize operational benefits in the forms of cost reductions, better operations planning, elimination of waste and process downtime, better organization of labor forces, and better resource consumptions. Therefore, I hypothesize that:

H1. BDA initiative is positively associated with operational efficiency.

3.2 BDA Initiative and Business Growth

In dynamic markets, business growth is a function of the capabilities of creating a series of temporary competitive advantages (Eisenhardt and Martin, 2000). By accessing, integrating, and analyzing large volumes of data from various sources, BDA produces innovative insights in different areas such as customer demands and preferences, product or service trends, markets, inventory management, etc. These insights support decision-making in established business model improvement (Günther et al., 2017; Loebbecke and Picot, 2015; Pigni, Piccoli, and Watson, 2016; Woerner and Wixom, 2015) to create a series of temporary competitive advantages. For instance, through analyzing retailers' granular data together with economic, weather, demographic, and geographic data, suppliers are able to better understand the markets and retailers' businesses and to figure out how to better sell products. Suppliers therefore are able to increase sales by leveraging BDA to better develop marketing campaigns and design promotions (Najjar and Kettinger, 2013).

The BDA-enabled insights also lead organizations to reconfigure resources to innovate and transform business models (Günther et al., 2017; Pigni et al., 2016; Woerner and Wixom, 2015) to create a set of temporary competitive advantages. Specifically, the utilization of BDA allows firms to "develop whole new value propositions, target different customers, or interact with customers in different ways" (Günther et al., 2017, pp. 197). Netflix is an example of an organization that utilized big data and analytics to overturn its historical operating model and moved from a disc rental model to an on-demand streaming model (Lycett, 2013). Analyzing data from different sources (e.g. catalogue data, search terms, film reviews, and social data), Netflix has enhanced the interaction between providers and subscribers by offering dynamic and personalized recommendations. Moreover, analysis of years' worth of user behavior data even informs content Netflix may produce to drive up the subscriber base, which moves Netflix from streaming content to producing content. Accordingly, I hypothesize that:

H2. BDA initiative is positively associated with business growth.

3.3 Industry Environments

3.3.1 Environmental Complexity

Environmental complexity is conceptualized as "the heterogeneity of and range of an organization's activities" (Child 1972, p. 3). Firms operating in complex environment tend to have many different inputs (e.g. suppliers, components, materials) and outputs (e.g. customers and products) (Dess and Beard, 1984). The multiple inputs and outputs as well as the interplay among them increase the chance of errors and make it difficult to identify, diagnose, and respond to problems in operations (Azadegan, Patel, and Parida, 2013; Azadegan, Patel, Zangoueinezhad, and Linderman, 2013), which raises the complexity of operations in organizations (Dess and Beard, 1984). Thus, organizations facing complex environments have greater information requirements than those facing simple environments (Dess and Beard, 1984).

The use of BDA addresses the information processing requirements by offering insights that organizations may use to sense, diagnose, and respond to problems in operations. For example, organizations in complex environments have a number of different suppliers, which increases the likelihood of experiencing supply disruption. Failure by a supplier to deliver products or services on time and in the right quantity and quality would impose a negative impact on organizations. BDA gives organizations the ability to identify events and trends by monitoring publicly available news or social media data related to suppliers or the sourcing markets (Wang et al., 2016). Organizations are able to sense the changes on suppliers, evaluate risks, and respond to changes with contingency plans, thus mitigating the cost resulting from supply disruptions and enhancing the efficiency of inbound logistics. According, I hypothesize that:

H3. Environmental complexity moderates the relationship between BDA initiative and operational efficiency, such that with increasing frequency of BDA initiative in complex environments operational efficiency is improved.

Concentration-dispersion dimension also underlines the environmental complexity (Dess and Beard). A concentrated industry may have a less complex environment as just a few firms dominate the industry (Xue, Ray, and Sambamurthy, 2012). On the contrary, a disperse industry may have a more complex environment because organizations need to deal with a greater number of competitors. In a disperse industry (i.e. more complex environment), it is difficult for firms to know about each competitor and to predict their competitors' likely behavior. Firms may have to compete in unpredictable ways. In such environments, firms need to respond very quickly to competitors' competitive actions by focusing on identifying and pursuing new product/service and market opportunities. Therefore, organizational information requirements increase so that firms can stay ahead of the competition in complex environments.

The use of BDA resolves the requirements by providing insights into a wide range of product/service and market conditions, trends, and events, so that firms could grow through radical product/service and process innovations and market development. For instance, the consumer fashion industry has a complex environment with numerous incumbents and new entrants. Companies implementing BDA in this industry are able to collect and analyze various sources of data such as existing customers' registered information and shopping history, videos and photos of available products, and social media data. Through BDA, companies will create a holistic understanding of potential new product development ideas and trends, thus gaining competitive advantages by improve innovation capabilities (Tan, Zhan, Ji, Ye, and Chang, 2015). Accordingly, I hypothesize that:

H4. Environmental complexity moderates the relationship between BDA initiative and business growth, such that BDA initiative is associated with greater growth in businesses in complex environments than in simple environments.

3.3.2 Environmental Dynamism and Munificence

Environmental dynamism refers to the unpredictability of environmental change (Dess and Beard, 1984). In a stable environment, firms face with lasting technology, steady supply, and stable and predictable customer demands. They can produce and sell homogeneous products or services in large volumes to attain economies of scale. Environments with lower level of dynamism enable firms to compete on incremental product/service and process improvement

(Tushman and Anderson, 1986; Xue et al., 2012) that reduces cost and drives up operational efficiency.

Environmental munificence is conceptualized as the ability of an environment to support sustained growth of an organization (Aldrich 1979). Low level of munificence indicates the scarcity of resources for organizational growth and stability (Dess and Beard, 1984). Although firms in such environments may face less competition from new entrants because the environments themselves are already short of resources, they still have to compete with existing players. Due to the lack of resources in environments with low munificence, deployment of any resources away from the core product/service market may not have any positive effects on performance (Goll and Rasheed, 2004). Thus, firms operating in less munificent environments focus greater attention on the core product/service markets and compete by reducing cost, minimizing waste, and improving the efficiency of current operations.

Thus, in less dynamic and munificent environments, firms' BDA initiative support operational planning and control to enhance operational efficiency. For example, chemical industry has a stable environment with steady changes in supply and demand for inputs and outputs (Davis, Eisenhardt, and Bingham, 2009; Lau, Man, and Chow, 2004). In addition, this industry is less munificent compared to the munificent high-tech industries such as computers and Internet industries. Chemical firms focus BDA on extracting insights from vast troves of production and process data to improve operations. The utilization of BDA reveals a number of previously unseen insights regarding the factors influencing overall yield. By resetting the identified parameters accordingly, chemical companies are able to reduce waste and variability and significantly improve product quality and yield without incurring additional capital expenditures (Auschitzky, Hammer, and Rajagopaul, 2014). Accordingly, I hypothesize that:

H5. Environmental dynamism moderates the relationship between BDA initiative and operational efficiency, such that in less dynamic environments, BDA initiative is associated with a greater increase in operational efficiency.

H6. Environmental munificence moderates the relationship between BDA initiative and operational efficiency, such that in less munificent environments, BDA initiative is associated with a greater increase in operational efficiency.

In a more dynamic environment, firms face high level of unpredictability in change in production and service technologies and customers' needs and preferences. The uncertain and rapid shift in technologies and markets requires firms to engage in continuous and relentless change for survival (Brown and Eisenhardt, 1997). Firms in such environments compete through radical innovation (Koberg, Detienne, and Heppard, 2003; Tushman and Anderson, 1986) which focuses on developing new products and services and exploring and entering new areas of opportunities.

Munificent environments have abundant resources for organizations to pursue sustained growth. The strategic priority of firms in munificent industries is to expand in terms of scale and/or scope (Dess and Beard, 1984). Moreover, munificent environments tend to attract new entrants (Aldrich 1979) with new assets and capabilities (Xue et al., 2012). As a result, munificent industries compel firms to create new products/services and seek new areas of opportunities for business growth. In more dynamic and munificent environments, firms' goal is to release new products and services and expand into new areas of opportunities ahead of competitors, thereby achieving a set of temporary competitive advantages.

In more dynamic and munificent environments, firms should utilize insights extracted from BDA to continuously redefine products, services, process, and markets to implement radical innovations. E-commerce is a dynamic and munificent industry with high level of unpredictability in customer needs but many opportunities for growth. An e-commerce giant, Amazon, recently has implemented big data, including purchase orders and historical product searches, to predict specifically when a customer will place an order. The knowledge allows Amazon to pre-ship the product to the nearest depot before the customer actually makes a purchase online or realizes that he/she needs it (Bensinger, 2014). Amazon utilizes the insights extracted from BDA to radically transform its distribution strategy and processes, which greatly expands its base of loyal customers and improves revenues (Erevelles, Fukawa, and Swayne, 2016). Accordingly, I hypothesize that:

H7. Environmental dynamism moderates the relationship between BDA initiative and business growth, such that in more dynamic environments, BDA initiative is associated with a greater increase in business growth.

H8. Environmental munificence moderates the relationship between BDA initiative and business growth, such that in more munificent environments, BDA initiative is associated with a greater increase in business growth.

4. Research Method

4.1 Data sources and sample

The data used in this study were collected from two sources: Firms' BDA initiatives were drawn from the Nexis Uni database; data for industry characteristics and firms' performance were obtained from the COMPUSTAT database.

I assembled a set of BDA initiative announcements by carrying out a detailed search in the Nexis Uni database that includes major newspapers and business wire news. The announcements were collected for Fortune 500 firms for the seven-year period from 2010 (when interest in big data began (Marr, 2015)) to 2016. I chose Fortune 500-firm criterion because such firms are extensively reported by the media. The search focused on firms that have a December 31 fiscal year end to facilitate data analysis and to ensure comparable performance periods for all firms (Bharadwaj, Bharadwaj, and Konsynski, 1999). Consistent with prior literature, I took year 2012 as the baseline year and included every Fortune 500 firm that meets the December year-end criterion of that year in the search (Altinkemer, Ozcelik, and Ozdemir, 2011). The company name within 75 words of the words “big data” or “analytics” or “real-time analysis” or “analyze real-time data” or “intelligence solution” or “intelligence solutions” or “data mining” or “machine learning” or “business intelligence” or “Hadoop” or “Map-reduce” or “Internet of Things” or “IoT” were used to identify potential news articles. I then read the title of each article as well as the section that mentioned the above keywords to carefully retain articles that pertained to the focal firm’s BDA projects.

When a press release was qualified, I identified and recorded the starting year of the BDA project. In some cases, a press release discussed the focal firm’s BDA project, but the project starting year was missing. I then looked for additional informative keywords within the information sources, such as the name of the purchased analytics software. A second round search using such additional keywords was performed for the focal firm to help retrieve the starting year. If the focal firm’s BDA project starting year still could not be recovered using additional keywords, the article that contained the firm’s BDA project was deleted given the intent to measure performance influences over the years (Altinkemer et al., 2011). In addition, a news article that included information about BDA projects within multiple firms was counted as multiple BDA initiatives, each relating to one of the firms involved (Altinkemer et al., 2011;

Subramani and Walden, 2001). When an article regarding a BDA initiative published in one newspaper was republished in other outlets, these duplicates were only counted once.

As my dynamic panel data models need a one-year lag between the dependent and independent variables, I collected data related to industrial factors and firm performance covering 2010 to 2017. Specifically, I needed a minimum of one piece of consecutive two-year data for each firm because the dynamic panel data models required one-year lagged performance and industry characteristics measure for analysis. 98 firms were found to have at least one piece of consecutive two-year data for all industry and performance measures. As the industry and performance data covered 8 years, I had a maximum of 7 pairs of consecutive two-year data (i.e. a maximum of 784 consecutive firm-years). However, not all the firms had the full ten-year consecutive data in the industry and performance measures. After excluding 12 consecutive firm-years with missing data in all performance and industry measures, I obtained a total of 772 consecutive firm-years.

These 98 firms represent 31 industries based on 2-digit SIC codes. The top 15 industries of the samples firms are displayed in Table 2. Table 2 shows that the sample firms come from a wide variety range of industries. The top five industry sectors based on 2-digit SIC codes include: (1) Electric, Gas, & Sanitary Service, (2) Chemical & Allied Products, (3) Industrial Machinery & Equipment, (4) Oil & Gas Extraction, and (5) Food & Kindred Products. Table 3 presents the characteristics of sample firms in terms of total assets, sales, number of employees, capital expenditure, operating income, and age.

Table 2. Industries of Sample Firms

2-digit SIC Codes	**Industries**	**Firm Frequency**	**Firm Percentage (%)**
49	Electric, Gas, & Sanitary Service	16	16.3
28	Chemical & Allied Products	15	15.3
35	Industrial Machinery & Equipment	6	6.1

13	Oil & Gas Extraction	5	5.1
20	Food & Kindred Products	4	4.1
55	Automotive Dealers & Service Stations	4	4.1
56	Apparel & Accessory Stores	4	4.1
26	Paper & Allied Products	3	3.1
33	Primary Metal Industries	3	3.1
50	Wholesale Trade – Durable Goods	3	3.1
Other SIC codes	Other Industries	35	35.6

Table 3. Characteristics of Sample Firms

Variable	**Unit**	**Mean**	**Std. deviation**	**Minimum**	**Maximum**
Age	Years	79.92	45.35	6.00	211.00
Assets	Millions of dollars	30522.08	30779.33	1404.60	212949.00
Capital Expenditure	Millions of dollars	1590.92	1950.65	6.50	19099.00
Cost of Goods Sold	Millions of dollars	16528.79	20234.60	551.941	200990.00
Number of Employees	Thousands	66.46	88.75	0.92	539.00
Operating Income	Millions of dollars	2931.16	3522.45	-13353.00	25042.00
Sales	Millions of dollars	24319.13	24940.57	1809.58	230859.00

4.2 Measurement of variables

Big data analytics initiative. I collected BDA initiatives from Nexis Uni database and counted the number of initiatives to measure firms' BDA efforts.

Operational efficiency. I adopted the Stochastic Frontier Estimation (SFE) methodology to measure firms' operational efficiency (Dutta et al., 2005; Lam et al., 2016; Li et al., 2010). A firm's operational efficiency is viewed as a transformational process. SFE models a firm's operational efficiency as the capability of converting the respective inputs (i.e. number of employees, cost of goods sold, and capital expenditure) into outputs (i.e. operating income) relative to peers within the same industry (Li et al., 2010). The SFE approach measures the relative operational efficiency of a firm in its industry, thus accounting for the industry heterogeneity (Dutta et al., 2005). In addition, SFE offers a more comprehensive measure of a firm's operational efficiency compared to traditional measures using a single indicator such as inventory turnover and labor productivity (Lam et al., 2016).

To implement SFE, I built a stochastic production function to model the level of operational output (i.e. operating income) that can be produced from a given level of operational inputs (i.e. number of employees, cost of goods sold, and capital expenditure). The function is showed below:

$$\ln(Operating\ Income)_{ijt} = \beta_0 + \beta_1 \ln(Capital\ Expenditure)_{ijt} + \beta_2 \ln(Cost\ of\ Goods\ Sold)_{ijt} + \beta_3 \ln(Number\ of\ Employees)_{ijt} + v_{ijt} - u_{ijt} \qquad (1)$$

Where v_{ijt} is the stochastic error term and u_{ijt} represents the technical inefficiency of firm *i* in industry *j* at time *t*. u_{ijt} is constrained to be between 0 and 1, with 0 indicating technically efficient (i.e. a firm's output level is on the frontier in its industry and within a certain year). Hence, the operational efficiency of firm *i* in industry *j* at year *t* can be expressed as below (Li et al., 2010):

$$Operational\ Efficiency_{ijt} = e^{-u_{ijt}} \qquad (2)$$

Tobin's Q. I operationalized Tobin's Q based on the definitions provided in Chung and Pruitt (1994), Bharadwaj, Bharadwaj, and Konsynski (1999), and Bardhan, Krishnan, and Lin (2013). Specifically,

Tobin's Q = (Market Value of Common Equity + Liquidating Value of Preferred Stock + Book Value of Debt) / Book Value of Total Assets

Where:

Market Value of Common Equity= PRCC_F × CSHO from Compustat;

Liquidating Value of Preferred Stock= PSTKL or PSTKRV if PSTKL is missing from Compustat;

Book Value of Debt= LCT – ACT + INVT + DLTT from Compustat;

Book Value of Total Assets= TA from Compustat.

Environmental Characteristics. I measured the environmental characteristics in terms of dynamism, complexity, and munificence based on the existing literature (e.g. Dess and Beard, 1984; Keats and Hitt, 1988; Palmer and Wiseman, 1999; Xue, Ray, and Gu, 2011; Xue, Ray, and Sambamurthy, 2012).

Dynamism refers to the volatility of an industry (Dess and Beard, 1984; Keats and Hitt, 1988). The volatility of sales in a dominant industry over a period of five years was used to measure dynamism (Keats and Hitt, 1988; Xue et al., 2011; Xue et al., 2012). The volatility of industry sales was measured using a two-step procedure. First, the natural logarithm of annual sales of all firm in each 2-digit SIC industry was regressed against an index variable of years, over a period of five years. Then, the antilog of the standard error of the regression slope coefficient was used for industry sales volatility over the period.

Munificence refers to "the availability of environmental resources to support growth" (Keats and Hitt, 1988). I used the sales growth in a dominant industry over a period of five years to measure munificence (Keats and Hitt, 1988; Xue et al., 2011; Xue et al., 2012)). Industry sales growth was measured through two steps. First, I regressed the natural logarithm of annual sales of all firms in each 2-digit SIC industry against the index variable of years, over a period of five years. Next, the antilog of the regression slope coefficient was used as the measure for industry sales growth over the period.

Complexity refers to the concentration-dispersion of task-environment elements (Dess and Beard, 1984; Keats and Hitt, 1988). I used Herfindahl index to measure complexity (Keats and Hitt, 1988; Xue et al., 2011; Xue et al., 2012). Herfindahl index is a well-known measure for market concentration. A large value of Herfindahl index indicates highly concentrated industry where a small number of firms dominate the industry and every firms knows its competitors and how to respond to competitors' actions. Hence, a more concentrated industry is less complex. Because a large value of Herfindahl index suggests low complexity, I adopted the log value of the reciprocal of the Herfindahl index as the measure for complexity.

4.3 Control variables

I included control variables to control for firm, industry, and year effects. First, prior studies suggest that firm size, firm age, and firm profitability influence operational efficiency (Kortmann, Gelhard, Zimmermann, and Piller, 2014; Lam et al., 2016) and business growth (Bardhan et al., 2013; Dezsö and Ross, 2010; Julian and Ofori-Dankwa, 2017). Thus, I included firm size, firm age, and firm profitability to control for firm-level effects. Firm size is measured as the natural logarithm of annual sales (Bardhan et al., 2013). Firm age is measured as the natural logarithm of difference between current year and founding year (Dezsö and Ross, 2010; Li et al., 2010). Firm profitability is measured as firm's return on assets (Lam et al., 2016). Second, I created seven binary dummy variables to take into account the year effects. Third, based on the two-digit SIC codes, I created thirty binary dummy variables to control for any unobservable industry effects. Table 4 provides the descriptive statistics of variables used and Table 5 shows the correlations between these variables.

Table 4. Descriptive Statistics of Variables Used

	Mean	Std. Deviation
BDA Initiative	0.40	0.743
Operational Efficiency	0.481	0.109
Lagged Operational Efficiency	0.481	0.109
Complexity	0.104	0.128
Dynamism	1.094	0.074
Munificence	0.998	0.066
Tobin's Q	1.467	0.993
Firm Size	9.824	0.735
Firm Age	4.188	0.720
Firm Profitability	0.057	0.070

Table 5. Correlation Matrix

	1	2	3	4	5	6	7	8	9	10
1. BDA Initiative	1									
2. Operational Efficiency	-.112**	1								
3. Lagged Operational Efficiency	-0.042	.431**	1							
4. Complexity	0.01	.198**	.183**	1						
5. Dynamism	0.041	.251**	.244**	.608**	1					
6. Munificence	-0.051	0.011	0.017	.290**	0.059	1				
7. Tobin's Q	0.032	-.113**	-.101**	.126**	.166**	-0.015	1			
8. Firm Size	.222**	-.096**	-.082*	.113**	.276**	-.121**	0.041	1		
9. Firm Age	.088*	-0.07	-0.062	-.279**	-.097*	-.125**	0.029	.109**	1	
10. Firm Profitability	0.049	-.163**	-.119**	0.068	0.068	0.049	.622**	.150**	-0.007	1
**. Correlation is significant at the 0.01 level (2-tailed).										
*. Correlation is significant at the 0.05 level (2-tailed).										

4.4 Analysis and Results

4.4.1 Dynamic Data Panel Models

I constructed the following two dynamic data panel models (DDP) to test the hypotheses.

$$\begin{aligned} Operational\ Efficiency_{it} &= \alpha_0 + \alpha_1 Operational\ Efficiency_{it-1} + \alpha_2 BDA\ Initiatives_{it-1} \\ &+ \alpha_3 Dynamism_{it-1} + \alpha_4 Munificence_{it-1} + \alpha_5 Complexity_{it-1} \\ &+ \alpha_6 BDA\ Initiatives_{it-1} \times Dynamism_{it-1} \\ &+ \alpha_7 BDA\ Initiatives_{it-1} \times Munificence_{it-1} \\ &+ \alpha_8 BDA\ Initiatives_{it-1} \times Complexity_{it-1} + \alpha_9 Firm\ Size_{it-1} \\ &+ \alpha_{10} Firm\ Age_{it-1} + \alpha_{11} Firm\ Profitability_{it-1} + \sum Year\ Dummies_{it-1} \\ &+ \sum Industry\ Dummies_{it-1} + \varepsilon_{i\,t-1} \qquad (3) \end{aligned}$$

$$
\begin{aligned}
Tobin's\ Q_{it} = {} & \beta_0 + \beta_1 Tobin's\ Q_{it-1} + \beta_2 BDA\ Initiatives_{it-1} + \beta_3 Dynamism_{it-1} \\
& + \beta_4 Munificence_{it-1} + \beta_5 Complexity_{it-1} \\
& + \beta_6 BDA\ Initiatives_{it-1} \times Dynamism_{it-1} \\
& + \beta_7 BDA\ Initiatives_{it-1} \times Munificence_{it-1} \\
& + \beta_8 BDA\ Initiatives_{it-1} \times Complexity_{it-1} + \beta_9 Firm\ Size_{it-1} \\
& + \beta_{10} Firm\ Age_{it-1} + \beta_{11} Firm\ Profitability_{it-1} + \sum Year\ Dummies_{it-1} \\
& + \sum Industry\ Dummies_{it-1} + \varepsilon_{i\,t-1} \qquad (4)
\end{aligned}
$$

In both models, the dependent variables (i.e. operational efficiency and Tobin's Q) were lagged by one year after the independent variables. Furthermore, the two models involved a dynamic panel specification. I included the lagged dependent variables as independent variables because firm performance such as operational efficiency and business growth could be path dependent and persistent over time (i.e. performance in year t-1 may affect performance in year t) (Lam et al., 2016; Mukherji, Sorescu, Prabhu, and Chandy, 2011). Following prior DDP studies (e.g. Lam et al., 2016; Mukherji et al., 2011), I maintained one-year lag of the dependent variables in both models. These models enable me to evaluate the impact of BDA initiative on operational efficiency and business growth. The coefficients of the interaction terms capture the moderating effects of the industry characteristics.

There are several challenges towards the estimation of the above models. First, the dynamic data panel models cannot be estimated using ordinal least squares because the lagged dependent variables (i.e. $Operational\ Efficiency_{it-1}$ and $Tobin's\ Q_{it-1}$) are correlated with the error terms, which gives rise to "dynamic panel bias" (Nickell, 1981). Second, although Kiviet (1995) argues that the least-square-dummy-variables (LSDV) estimator can handle dynamic panel bias, his approach works only for balanced panels. Therefore, the LSDV estimator is not appropriate for my unbalanced panels where some firms have less observations than other firms. Third, potential endogeneity exists in the independent variables of the DDP models. One source of endogeneity problems is the reverse causality. For example, in my model

I proposed that BDA initiative may affect operational efficiency and business growth. However, organizational performance may influence organizational strategies (Bardhan et al., 2013; Lam et al., 2016). That is to say, operational efficiency and business growth may affect organizations' decisions in BDA adoption. Thus, instruments are needed to address the endogeneity concern (Roodman, 2009). Instruments can be external variables available outside the immediate dataset or internal variables based on lags of instrumented variables (Roodman, 2009). Prior studies (e.g. Bardhan et al., 2013; Lam et al., 2016) have indicated the difficulty of finding strictly exogenous external instruments. Hence, I had to rely on obtaining instruments internally. To tackle the aforementioned challenges, I applied the system GMM approach developed by Arellano and Bover (1995) and Blundell and Bond (1998) to estimate the two DDP models.

System GMM approach employs a system of two equations: the original level equation and the transformed equation by first differencing the variables in the original equation (Arellano and Bover, 1995; Blundell and Bond, 1998). I used the Xtabond2 command in the Stata software to perform system GMM estimation. Independent variables including $Operational\ Efficiency_{it-1}, Tobin's\ Q_{it-1}, BDA\ Initiatives_{it-1}$, $BDA\ Initiatives_{it-1} \times Dynamism_{it-1}$, $BDA\ Initiatives_{it-1} \times Munificence_{it-1}$, and $BDA\ Initiatives_{it-1} \times Complexity_{it-1}$ were considered endogenous and were instrumented with lagged values of the variables for both level and transformed equations. Specifically, I specified lag (2 .) and deeper for the transformed equation and lag 1 for the levels equation, which is the standard treatment for endogenous regressors (Roodman, 2009). I performed the two-step GMM estimation because the two-step estimator is asymptotically efficient and robust to heteroscedasticity (Roodman, 2009).

4.4.2 Results

Table 6 presents the results of system GMM estimation. I discuss the impact of BDA initiative on operational efficiency first. As shown in Table 6 Panel B, the coefficient of BDA initiative is positive and significant ($p<0.01$), suggesting the support of H1. The negative and significant coefficient of *BDA Initiative* × *Dynamism* interaction ($p<0.05$) indicates that BDA initiative has a greater influence on operational efficiency in less dynamic environments than in more dynamic environments. Thus, H5 is supported. The negative and significant coefficient of *BDA Initiative* × *Munificence* interaction ($p<0.01$) infers that BDA initiative has a greater impact on operational efficiency in less munificent environments than in more munificent environments. This result provides support for H6. The positive and significant coefficient of *BDA Initiative* × *Complexity* interaction ($p<0.01$) means that BDA initiative has a greater impact on operational efficiency in more complex environments than in less complex environment. This result supports H3.

Panel B of Table 6 also displays the impact of BDA initiative on business growth. The coefficient of BDA initiative is positive and significant ($p<0.01$), which is consistent with H2. The coefficient of *BDA Initiative* × *Complexity* interaction is positive and significant ($p<0.01$). This indicates that BDA initiative has a greater impact on business growth in more complex environments than in less complex environments. Therefore, H4 is supported. The coefficient of *BDA Initiative* × *Dynamism* is positive and significant ($p<0.05$), indicating that BDA initiative has a greater impact on business growth in more dynamic environments than in less dynamic environments. This finding provides support for H7. The coefficient of *BDA Initiative* × *Munificence* is negative and significant ($p<0.01$), suggesting that BDA initiative has a greater influence on business growth in less munificent environments than in more munificent environments. This result is not consistent with H8.

Table 6. Impact of BDA Initiative on Operational Efficiency and Business Growth

Panel A: The Base Model		
	Operational Efficiency	Standardized Tobin's Q
Lagged Operational Efficiency	0.316**	0.601**
BDA Initiative	0.001	-0.024
Complexity	0.526	-2.685
Dynamism	-1.173*	-0.610
Munificence	0.131**	-0.183
Firm Size	-0.050**	0.015
Firm Age	0.015†	-0.042
Firm Profitability	-0.086	0.338†
F Statistic	1188.06	29474.87
Hansen Test	0.358	0.676
AR Test: AR(1)	0.000	0.022
AR Test: AR(2)	0.189	0.203
Year dummies and industry dummies are included in the analysis, but their coefficients are not reported for brevity.		
Panel B: The Full Model		
	Operational Efficiency	Standardized Tobin's Q
Lagged Standardized Tobin's Q	0.308**	0.640**
BDA Initiative	0.366**	1.128**
Complexity	-0.117	-1.868**
Dynamism	-0.952**	0.653
Munificence	0.184**	0.481**
BDA Initiative × Complexity	0.077**	0.406**
BDA Initiative × Dynamism	-0.138*	0.131*
BDA Initiative × Munificence	-0.229**	-1.373**
Firm Size	-0.042**	0.031
Firm Age	0.007	-0.026
Firm Profitability	-0.091*	0.803**
F Statistic	1140000	1340000
Hansen Test	0.999	1.000
AR Test: AR(1)	0.000	0.015
AR Test: AR(2)	0.154	0.123
Year dummies and industry dummies are included in the analysis, but their coefficients are not reported for brevity.		

Note: †$p<0.1$; *$p<0.05$; **$p<0.01$

5. Discussion

My research proposes and provides empirical evidence about the influence of BDA initiative on operational efficiency and business growth as well as how industry environment moderates the BDA-organizational performance relationship. I find that, consistent with the hypotheses, BDA initiative is related to a greater increase in operational efficiency in less dynamic and munificent industry environments. Meanwhile, BDA initiative is related to a greater increase in business growth in more dynamic and complex industry environments.

5.1 Theoretical Implications

Although extant research has provided evidence that BDA enhances a number of performance measures, including operational efficiency and business growth, the evidence is primarily survey-based or case-based. Evidence from surveys tends to be subjective and usually only successful cases are reported. My study departs from this research stream as it uses secondary data in a longitudinal setting to investigate the real impact of the strategic use of BDA on organizational performance. Therefore, this study extends and enriches existing big data literature by providing more objective evidence regarding BDA value creation.

The insignificant main effects in the base model and significant effects in the full model suggest that the value generation of BDA is not universal to organizations. This finding is consistent with the fact that organizations report mixed success in achieving their analytics objectives (Brown and Gottlieb, 2016). Many organizations have invested in BDA at scale but still have not yielded the payoff they expected (Henke et al., 2016). My study provides insights into the conditions under which BDA initiative brings value to organizations. Specifically, it demonstrates that BDA initiative enables a greater increase in operational efficiency in less dynamic and munificent but more complex environment. At high level of complexity and dynamism but low level of munificence, BDA initiative contributes to greater business growth. The results further extend the environment-strategy-performance perspective (Child, 1972; Dyer and Singh, 1998) in that the strategic use of BDA alignment with industry contexts of organizations is important for generating positive returns.

5.2 Practical Implications

This research offers important managerial implications. The first implication is regarding the evaluation of BDA initiative payoff. Companies need to take into account the industry environments they are operating in when assessing the payoff of their BDA initiatives. My study provides initial evidence that BDA initiative influences different dimensions of organizational performance in different industry environments. For example, the findings indicate that firms operating in a more dynamic environment will see growth in businesses but may not see any increase in the efficiency of operations. Therefore, at level of higher dynamism, BDA initiative will pay off through the measure of business growth but may not pay off if you look at operational efficiency. On the contrary, firms operating in a less dynamic environment will see improvement in the efficiency of operations but may not see growth in businesses. Thus, at low level of dynamism, BDA initiative will pay off according to the performance measure of operational efficiency but may not pay off if you look at the measure of business growth. Companies need to understand their industry environments and select the proper performance indicators to evaluate the payoff of BDA initiative.

Second, the moderating effects of industry environment on BDA-enabled value creation discovered in my study can guide practitioners to use BDA to generate value. Companies need to align BDA initiatives focusing on specific strategies with their industry contexts. In less dynamic and munificent but more complex environments, BDA initiatives support companies' cost-cutting strategy and yield better operational efficiency. In more dynamic and complex but less munificent environments, BDA initiatives support companies' growth strategy and generate higher Tobin's Q. Therefore, companies in less dynamic environment may adopt BDA to support their cost-cutting strategy. Companies in more dynamic environment may conduct BDA initiatives to support their growth strategy.

6. Limitation and Future Research

This study has several limitations. First, my measure for BDA initiative variable (i.e. the count of announcements regarding BDA related initiatives by organizations) is not perfect. The ideal measure for this variable should be the specific spending on BDA in dollars at firm level. However, I have searched different sources (e.g. firms' 10-K reports, international data corporation) but failed to obtain BDA spending data at firm level. BDA initiative announcements from the Nexus Uni database provide detailed descriptions regarding each BDA initiative. Therefore, I believe that the count of announcements can be a good proxy of BDA adoption. I collected data from 98 firms for the current study. In the future I will collect BDA related initiative announcements from more firms to have a larger size of sample firms.

Second, industry environment characteristics variables (i.e. complexity, dynamism, and munificence) are treated as exogenous variables in this study. Nevertheless, in some cases, the strategic use of IT may affect industry environments (Xue et al. (2011)). It is likely that the strategic use of BDA such as the adoption of NoSQL databases and advanced analytical software may also influence industry environments. Researchers in the future may explore how BDA spending at industry level will influence industry environments.

Third, my study focuses on the general BDA initiatives and does not consider the differences among the BDA initiatives. For examples, some initiatives are within certain functions and some are enterprise-wise. Some initiatives are related to day-to-day operations and some are concerning research and development activities. Future studies may take into account the differences among the BDA initiatives and investigate how the differences influence organizational performance differently. In addition, this research primarily concentrates on industry environment characteristics to understand the circumstances under which firms are likely to benefit from BDA initiatives. Future research may look into other contingency factors (e.g. information intensity and competitive strategy) and study their implications for organizational performance.

7. Conclusion

Grounded in the dynamic capability and contingency theories, this study empirically demonstrated the impact of BDA initiative on organizational performance and how industry environment characteristics moderate the BDA-performance relationship. Based on the 772 observations collected from Nexus Uni and COMPUSTAT databases, I constructed two dynamic panel data models. System GMM estimation was employed to analyze the data. BDA initiative was found to positively influence operational efficiency and business growth in the long term. Furthermore, BDA initiative is associated with a greater improvement in operational efficiency in less dynamic and munificent but more complex environments. BDA initiative is associated with greater business growth in more complex and dynamic environments. This research is among the first to provide a theory-centric understanding about BDA's economic benefits in the long term. The findings offer insights to firms about what actual benefits they may expect from BDA initiatives and how firms may realize the value of BDA by tailoring their BDA initiatives for the industry environments they are operating in.

References

Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(2).

Addo-Tenkorang, R., & Helo, P. T. (2016). Big data applications in operations/supply-chain management: A literature review. *Computers & Industrial Engineering*, *101*, 528-543.

Akter, S., Fosso Wamba, S., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment?. *International Journal of Production Economics*, *182*, 113-131.

Aldrich, E. Howard. (1979). Organizations and Environments. Englewood Cliffs, NJ: Prentice-Hall.

Altinkemer, K., Ozcelik, Y., & Ozdemir, Z. D. (2011). Productivity and performance effects of business process reengineering: A firm-level analysis. *Journal of Management Information Systems*, *27*(4), 129-162.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, *68*(1), 29-51.

Auschitzky, E., Hammer, M., & Rajagopaul, A. (2014). How big data can improve manufacturing. *McKinsey & Company*, *822*.

Azadegan, A., Patel, P. C., & Parida, V. (2013). Operational slack and venture survival. *Production and Operations Management*, *22*(1), 1-18.

Azadegan, A., Patel, P. C., Zangoueinezhad, A., & Linderman, K. (2013). The effect of environmental complexity and environmental dynamism on lean practices. *Journal of Operations Management*, *31*(4), 193-212.

Bain, J. S. (1968) Industry Organization. New York: John Wiley.

Bardhan, I., Krishnan, V., & Lin, S. (2013). Research note—Business value of information technology: Testing the interaction effect of IT and R&D on Tobin's Q. *Information Systems Research*, *24*(4), 1147-1161.

Barton, D., & Court, D. (2012). Making advanced analytics work for you. *Harvard business review*, *90*(10), 78-83.

Bean, R., 2017. How Companies Say They’re Using Big Data. *Harvard Business Review*. Available from: https://hbr.org/2017/04/how-companies-say-theyre-using-big-data (Retrieved on February 8th, 2018).

Bensinger, G., 2014. Amazon wants to ship your package before you buy it. *The Wall Street Journal*. Available from: https://blogs.wsj.com/digits/2014/01/17/amazon-wants-to-ship-your-package-before-you-buy-it/?KEYWORDS=amazon+anticipatory (Retrieved October 18th, 2018).

Bharadwaj, A. S., Bharadwaj, S. G., & Konsynski, B. R. (1999). Information technology effects on firm performance as measured by Tobin's q. *Management science*, *45*(7), 1008-1024.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, *87*(1), 115-143.

Brown, S. L., & Eisenhardt, K. M. (1997). The art of continuous change: Linking complexity theory and time-paced evolution in relentlessly shifting organizations. *Administrative science quarterly*, 1-34.

Brown, B., & Gottlieb, J. (2016). The need to lead in data and analytics. *McKinsey & Company*. Available from https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-need-to-lead-in-data-and-analytics (Retrieved October 5th, 2018).

Chae, B., Olson, D., & Sheu, C. (2014a). The impact of supply chain analytics on operational performance: a resource-based view. *International Journal of Production Research*, *52*(16), 4695-4710.

Chae, B. K., Yang, C., Olson, D., & Sheu, C. (2014b). The impact of advanced analytics and data accuracy on operational performance: A contingent resource-based theory (RBT) perspective. *Decision Support Systems*, *59*, 119-126.

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems*, *32*(4), 4-39.

Child, J. (1972). Organizational structure, environment and performance: The role of strategic choice. *sociology*, *6*(1), 1-22.

Chung, K. H., & Pruitt, S. W. (1994). A simple approximation of Tobin's q. *Financial management*, 70-74.

Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, *70*, 379-390.

Datameer, Inc., 2016. Top Five High-impact Use Cases for Big Data Analytics. Available from https://www.datameer.com/pdf/eBook-Top-Five-High-Impact-UseCases-for-Big-Data-Analytics.pdf (Retrieved February 8th, 2018).

Davenport, T. (2014). How strategists use “big data” to support internal business decisions, discovery and production. *Strategy & Leadership*, *42*(4), 45-50.

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2009). Optimal structure, market dynamism, and the strategy of simple rules. *Administrative Science Quarterly*, *54*(3), 413-452.

Dess, G. G., & Beard, D. W. (1984). Dimensions of organizational task environments. *Administrative science quarterly*, 52-73.

Dezsö, C. L., & Ross, D. G. (2012). Does female representation in top management improve firm performance? A panel data investigation. *Strategic Management Journal*, *33*(9), 1072-1089.

Dutta, S., Narasimhan, O. M., & Rajiv, S. (2005). Conceptualizing and measuring capabilities: Methodology and empirical application. *Strategic Management Journal*, *26*(3), 277-285.

Dyer, J. H., & Singh, H. (1998). The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of management review*, *23*(4), 660-679.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they?. *Strategic management journal*, 1105-1121.

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897-904.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data'can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, *165*, 234-246.

Fosso Wamba, S., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356-365.

Ghasemaghaei, M., Ebrahimi, S., & Hassanein, K. (2017a). Data analytics competency for improving firm decision making performance. *The Journal of Strategic Information Systems*.

Ghasemaghaei, M., Hassanein, K., & Turel, O. (2017b). Increasing firm agility through the use of data analytics: The role of fit. *Decision Support Systems*, *101*, 95-105.

Ghoshal, A., Larson, E., Subramanyam, R., & Shaw, M. (2014). The impact of business analytics strategy on social, mobile, and cloud computing adoption. In: Proceedings of the Thirty-fifth International Conference on Information Systems, Auckland, New Zealand December 14-17.

Goll, I., & Rasheed, A. A. (2004). The moderating effect of environmental munificence and dynamism on the relationship between discretionary social responsibility and firm performance. *Journal of Business Ethics*, *49*(1), 41-54.

Günther, W. A., Mehrizi, M. H. R., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems*.

Gunasekaran, A., Yusuf, Y. Y., Adeleye, E. O., & Papadopoulos, T. (2017). Agile manufacturing practices: the role of big data and business analytics with multiple case studies. *International Journal of Production Research*, 1-13.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049-1064.

Henke, N., Bughin, J., & Chui, M. (2016). Most industries are nowhere close to realizing the potential of analytics. *Harvard Business Review*. Available from https://hbr.org/2016/12/most-industries-are-nowhere-close-to-realizing-the-potential-of-analytics (Retrieved October 9th, 2018).

International Institution for Analytics, 2016. Advanced Analytics & Big Data Adoption Report 2016. Available from https://iianalytics.com/analytics-resources/advanced-analytics-big-data-adoption-report-2016 (Retrieved January 31th, 2018)

Jeseke, M., Grüner, M., & Weiß, F. (2013). Big data in logistics: A DHL perspective on how to move beyond the hype. White Paper; DHL Customer solutions & innovation, presented by

Martine Wegner, Vice President Solutions & Innovation, 53844 Troisdorf, Germany.

Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*, *55*(17), 5011-5026.

Julian, S. D., & Ofori-Dankwa, J. C. (2017). Context matters: Diversity's short-and long-term effects in fortune's “best companies to work for”. *Strategic Management Journal*, *38*(7), 1557-1565.

Keats, B. W., & Hitt, M. A. (1988). A causal model of linkages among environmental dimensions, macro organizational characteristics, and performance. *Academy of management journal*, *31*(3), 570-598.

Kiviet, J. F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of econometrics*, *68*(1), 53-78.

Koberg, C. S., Detienne, D. R., & Heppard, K. A. (2003). An empirical test of environmental, organizational, and process factors affecting incremental and radical innovation. *The Journal of High Technology Management Research*, *14*(1), 21-45.

Kortmann, S., Gelhard, C., Zimmermann, C., & Piller, F. T. (2014). Linking strategic flexibility and operational efficiency: The mediating role of ambidextrous operational capabilities. *Journal of Operations Management*, *32*(7-8), 475-490.

Krishnamoorthi, S., & Mathew, S. K. (2018). Business Analytics and Business Value: A Comparative Case Study. *Information & Management*.

Lam, H. K., Yeung, A. C., & Cheng, T. E. (2016). The impact of firms' social media initiatives on operational efficiency and innovativeness. *Journal of Operations Management*, *47*, 28-43.

Lau, T., Man, T. W., & Chow, I. (2004). Organizational capabilities and performance of SMEs in dynamic and stable environments. *The International Journal of Entrepreneurship and Innovation*, *5*(4), 221-229.

Lycett, M. (2013). 'Datafication': making sense of (big) data in a complex world. *European Journal of Information Systems*, *22*(4), 381-386.

Li, S., Shang, J., & Slaughter, S. A. (2010). Why do software firms fail? Capabilities, competitive actions, and firm survival in the software industry from 1995 to 2007. *Information Systems Research*, *21*(3), 631-654.

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems*, *24*(3), 149-157.

Marr B. (2015). A brief history of big data everyone should read. Available from https://www.weforum.org/agenda/2015/02/a-brief-history-of-big-data-everyone-should-read/ (Retrieved Decembe15th, 2017).

McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard business review*, *90*(10), 60-68.

Mukherji, P., Sorescu, A., Prabhu, J. C., & Chandy, R. K. (2011). Behemoths at the gate: how incumbents take on acquisitive entrants (and why some do better than others). *Journal of Marketing*, *75*(5), 53-70.

Najjar, M. S., & Kettinger, W. J. (2013). Data Monetization: Lessons from a Retailer's Journey. *MIS Quarterly Executive*, *12*(4).

NewVantage Partners LLC., 2017. Big Data Executive Survey 2017: Executive Summary of Findings. Available from: http://newvantage.com/wp-content/uploads/2017/01/Big-Data-Executive-Survey-2017-Executive-Summary.pdf (Retrieved December 29th, 2017).

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, 1417-1426.

Palmer, T. B., & Wiseman, R. M. (1999). Decoupling risk taking from income stream uncertainty: A holistic model of risk. *Strategic Management Journal*, *20*(11), 1037-1062.

Pigni, F., Piccoli, G., & Watson, R. (2016). Digital Data Streams: Creating value from the real-time flow of big data. *California Management Review*, *58*(3), 5-25.

Popovič, A., Hackney, R., Tassabehji, R., & Castelli, M. (2018). The impact of big data analytics on firms' high value business performance. *Information Systems Frontiers*, *20*(2), 209-222.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. Stata Journal, 9 (1), 86-136. *For Online Publication*.

Sabherwal, R., & Chan, Y. E. (2001). Alignment between business and IS strategies: A study of prospectors, analyzers, and defenders. *Information systems research*, *12*(1), 11-33.

Seddon, P. B., Constantinidis, D., Tamm, T., & Dod, H. (2017). How does business analytics contribute to business value?. *Information Systems Journal*, *27*(3), 237-269.

Subramani, M., & Walden, E. (2001). The impact of e-commerce announcements on the market value of firms. *Information Systems Research*, *12*(2), 135-154.

Tambe, P. (2014). Big data investment, skills, and firm value. *Management Science*, *60*(6), 1452-1469.

Tan, K. H., Zhan, Y., Ji, G., Ye, F., & Chang, C. (2015). Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. *International Journal of Production Economics*, *165*, 223-233.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic management journal*, 509-533.

Teece, D. J. (2009). *Dynamic capabilities and strategic management,* Oxford University Press, Oxford.

Teo, T. S., Nishant, R., & Koh, P. B. (2016). Do shareholders favor business analytics announcements?. *The Journal of Strategic Information Systems*, *25*(4), 259-276.

Tushman, M. L., & Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative science quarterly,* 439-465.

Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, *176,* 98-110.

Wang, Y., & Hajli, N. (2017). Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, *70,* 287-299.

Wang, Y., Kung, L., Wang, W. Y. C., & Cegielski, C. G. (2018). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management, 55*(1), 64-79.

Woerner, S. L., & Wixom, B. H. (2015). Big data: extending the business strategy toolbox. *Journal of Information Technology*, *30*(1), 60-62.

Xie, K., Wu, Y., Xiao, J., & Hu, Q. (2016). Value co-creation between firms and customers: The role of big data-based cooperative assets. *Information & Management*, *53*(8), 1034-1048.

Xue, L., Ray, G., & Sambamurthy, V. (2012). Efficiency or innovation: How do industry environments moderate the effects of firms' IT asset portfolios?. *MIS quarterly*, 509-528.

Xue, L., Ray, G., & Gu, B. (2011). Environmental uncertainty and IT infrastructure governance: A curvilinear relationship. *Information Systems Research*, *22*(2), 389-399.

Comprehensive References

Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(2), 3.

Abbott, G. (2017). Connected fleets leverage analytics for better business. *Material Handling & Logistics*, Retrieved from http://spot.lib.auburn.edu/login?url=https://search.proquest.com/docview/1877731616?accountid=8421 (Accessed on July 10, 2017).

Addo-Tenkorang, R., & Helo, P. T. (2016). Big data applications in operations/supply-chain management: A literature review. *Computers & Industrial Engineering*, *101*, 528-543.

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, *182*, 113-131.

Aldrich, E. Howard. (1979). Organizations and Environments. Englewood Cliffs, NJ: Prentice-Hall.

Altinkemer, K., Ozcelik, Y., & Ozdemir, Z. D. (2011). Productivity and performance effects of business process reengineering: A firm-level analysis. *Journal of Management Information Systems*, *27*(4), 129-162.

Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology*, *76*(5), 732.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, *68*(1), 29-51.

Aronow, S., Burkett, M., Romano, J., & Nilles, K. (2016). The 2016 supply chain TOP 25: Lessons from leaders. *Supply Chain Management Review, 20*(5), 10-21.

Arunachalam, D., Kumar, N., & Kawalek, J. P. (2017). Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. *Transportation Research Part E: Logistics and Transportation Review*.

Auschitzky, E., Hammer, M., & Rajagopaul, A. (2014). How big data can improve manufacturing. *McKinsey & Company*, *822*.

Azadegan, A., Patel, P. C., & Parida, V. (2013). Operational slack and venture survival. *Production and Operations Management*, *22*(1), 1-18.

Azadegan, A., Patel, P. C., Zangoueinezhad, A., & Linderman, K. (2013). The effect of environmental complexity and environmental dynamism on lean practices. *Journal of Operations Management*, *31*(4), 193-212.

Babiceanu, R. F., & Seker, R. (2016). Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook. *Computers in Industry*, *81*, 128-137.

Bain, J. S. (1968) Industry Organization. New York: John Wiley.

Bardhan, I., Krishnan, V., & Lin, S. (2013). Research note—Business value of information technology: Testing the interaction effect of IT and R&D on Tobin's Q. *Information Systems Research*, *24*(4), 1147-1161.

Barney, J. B. (2001). Resource-based theories of competitive advantage: A ten-year retrospective on the resource-based view. *Journal of management*, *27*(6), 643-650.

Barton, D., & Court, D. (2012). Making advanced analytics work for you. *Harvard business review*, *90*(10), 78-83.

Bean, R., 2017. How Companies Say They're Using Big Data. *Harvard Business Review*. Available from: https://hbr.org/2017/04/how-companies-say-theyre-using-big-data (Retrieved on February 8th, 2018).

Bensinger, G., 2014. Amazon wants to ship your package before you buy it. *The Wall Street Journal*. Available from: https://blogs.wsj.com/digits/2014/01/17/amazon-wants-to-ship-your-package-before-you-buy-it/?KEYWORDS=amazon+anticipatory (Retrieved October 18th, 2018).

Bentz, B. (2014). "Supply Chain Control Towers Help Organizations Respond to NEW PRESSURES." *Supply Chain Management Review* 18(4), 34-9.

Bharadwaj, A. S., Bharadwaj, S. G., & Konsynski, B. R. (1999). Information technology effects on firm performance as measured by Tobin's q. *Management science*, *45*(7), 1008-1024.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, *87*(1), 115-143.

Bollen, K. A. (2007). Interpretational confounding is due to misspecification, not to type of indicator: Comment on Howell, Breivik, and Wilcox (2007). Psychological Methods, 12(2), 219-228.

Boone, C. A., Skipper, J. B., & Hazen, B. T. (2017). A framework for investigating the role of big data in service parts management. *Journal of cleaner production*, *153*, 687-691.

Borkowsky, M., & Walther, B. (2014). The next big thing. *Air Transport World*, *51*(6), 34.

Briest, P., Dilda, V., & Somers, K. (2014). Taming manufacturing complexity with advanced analytics. *Mckinsey Quarterly*, (4), 24-25.

Briest, P., Dilda, V., & Somers, K. (2014). Taming manufacturing complexity with advanced analytics. *Mckinsey Quarterly*, (4), 24-25.

Brinch, M. (2018). Understanding the value of big data in supply chain management and its business processes: Towards a conceptual framework. *International Journal of Operations & Production Management*.

Brown, S. L., & Eisenhardt, K. M. (1997). The art of continuous change: Linking complexity theory and time-paced evolution in relentlessly shifting organizations. *Administrative science quarterly*, 1-34.

Brown, B., Chui, M., & Manyika, J. (2011). Are you ready for the era of 'big data'? (cover story). *Mckinsey Quarterly*, (4), 24-35.

Brown, B., & Gottlieb, J. (2016). The need to lead in data and analytics. *McKinsey & Company*. Available from https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-need-to-lead-in-data-and-analytics (Retrieved October 5th, 2018).

Bughin, J., Chui, M., & Manyika, J. (2015). An executive's guide to the Internet of Things. *Mckinsey Quarterly*, (4), 92-101.

Burt, R. S. (1976). Interpretational confounding of unobserved variables in structural equation models. *Sociological methods & research*, *5*(1), 3-52.

Catlin, T., Scanlan, J., & Willmott, P. (2015). Raising your digital quotient. *Mckinsey Quarterly*, (3), 30-43.

Chae, B., & Olson, D. L. (2013). Business analytics for supply chain: A dynamic-capabilities framework. *International Journal of Information Technology & Decision Making*, *12*(01), 9-26.

Chae, B., Olson, D., & Sheu, C. (2014 a). The impact of supply chain analytics on operational performance: A resource-based view. *International Journal of Production Research*, *52*(16), 4695-4710.

Chae, B. K., Yang, C., Olson, D., & Sheu, C. (2014 b). The impact of advanced analytics and data accuracy on operational performance: A contingent resource-based theory (RBT) perspective. *Decision Support Systems*, *59*, 119-126.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, *36*(4).

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. *Journal of Management Information Systems*, *32*(4), 4-39.

Chidambaram, V., Evans, H., & Etheredge, K. (2015). Big data: Is the energy industry starting to see real applications? *Supply Chain Management Review, 19*(1), 62-64.

Child, J. (1972). Organizational structure, environment and performance: The role of strategic choice. *sociology*, *6*(1), 1-22.

Chowdhury, M. M. H., & Quaddus, M. (2017). Supply chain resilience: Conceptualization and scale development using dynamic capability theory. *International Journal of Production Economics*, *188*, 185-204.

Chui, M., Loffler, M., & Roberts, R. (2010). The Internet of Things. *Mckinsey Quarterly*, (2), 70-79.

Chung, K. H., & Pruitt, S. W. (1994). A simple approximation of Tobin's q. *Financial management*, 70-74.

Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, *70*, 379-390.

Datameer, Inc., 2016. Top Five High-impact Use Cases for Big Data Analytics. Available from https://www.datameer.com/pdf/eBook-Top-Five-High-Impact-UseCases-for-Big-Data-Analytics.pdf (Retrieved February 8th, 2018).

Davenport, T. (2014). How strategists use "big data" to support internal business decisions, discovery and production. *Strategy & Leadership*, *42*(4), 45-50.

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2009). Optimal structure, market dynamism, and the strategy of simple rules. *Administrative Science Quarterly*, *54*(3), 413-452.

De Oliveira, M. P. V., McCormack, K., & Trkman, P. (2012). Business analytics in supply chains–The contingent effect of business process maturity. *Expert Systems with Applications*, *39*(5), 5488-5498.

DeVellis, R. 2011. *Scale Development: Theory and Applications* (3rd ed.), Thousand Oaks, CA: Sage Publications.

Demirkan, H., & Delen, D. (2013). Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems*, *55*(1), 412-421.

Dess, G. G., & Beard, D. W. (1984). Dimensions of organizational task environments. *Administrative science quarterly*, 52-73.

Dezsö, C. L., & Ross, D. G. (2012). Does female representation in top management improve firm performance? A panel data investigation. *Strategic Management Journal*, *33*(9), 1072-1089.

Dhawan, R., Singh, K., & Tuteja, A. (2014). When big data goes lean. *McKinsey Quarterly,* (2), 97-102.

Dirlea, V., Ng, H., Gutiérrez, M., & See, B. (2016). The power of modern supply chain modeling. *Supply Chain Management Review, 20*(5), 52-54.

Douglas, M. (2014). Leveraging Big Data. *Inbound Logistics*, Retrieved from http://www.inboundlogistics.com/cms/article/leveraging-big-data/ (Accessed on July 10, 2017).

Drazin, R., & Van de Ven, A. H. (1985). Alternative forms of fit in contingency theory. *Administrative science quarterly*, 514-539.

Dutta, S., Narasimhan, O. M., & Rajiv, S. (2005). Conceptualizing and measuring capabilities: Methodology and empirical application. *Strategic Management Journal*, *26*(3), 277-285.

Dyer, J. H., & Singh, H. (1998). The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of management review*, *23*(4), 660-679.

Edwards, J. R. (2001). Multidimensional constructs in organizational behavior research: An integrative analytical framework. *Organizational research methods*, *4*(2), 144-192.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they?. *Strategic management journal*, 1105-1121.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing*, *62*(1), 107-115.

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897-904.

Fathi, N. (2011). A Holistic View of Quality. *Supply & Demand Chain Executive 12*(3), 14-16.
Fedewa, D., Velarde, G. L., & O'Neill, B. (2016). Using buzz analytics to gain a product and marketing edge. *Mckinsey Quarterly,* (2), 14-15.

Field, A. M. (2014). Putting data to work. (cover story). *Journal of Commerce (1542-3867),* 4A-12A.

Finley, F., Blaeser, J., & Djavairian, A. (2015). Harnessing big data: Building & maintaining capabilities that deliver results. *Supply Chain Management Review, 19*(5), 24-28,30-31.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 39-50.

Fosso Wamba, S., Angappa, G., Papadopoulos, T., & Ngai, E. (2018). Big data analytics in logistics and supply chain management. *International Journal of Logistics Management, The*, (just-accepted), 00-00.

Fosso Wamba, S., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356-365.

Garrett, R. (2016a). Climbing the IoT Mountain. *Supply & Demand Chain Executive,* Retrieved from https://www.sdcexec.com/software-technology/article/12219117/climbing-the-iot-mountain (Accessed on July 17, 2017).

Garrett, R. (2016b). Set your sights on strategic sourcing. *Supply & Demand Chain Executive,* Retrieved from https://www.sdcexec.com/sourcing-procurement/article/12247791/set-your-sights-on-strategic-sourcing (Accessed on July 17, 2017).

Garrett, R. (2016c). Taming big data in the IoT. *Supply & Demand Chain Executive,* Retrieved from https://www.sdcexec.com/software-technology/article/12271216/taming-big-data-in-the-iot (Accessed on July 17, 2017).

Garrett, R. (2016d). Go big or go home. *Supply & Demand Chain Executive,* Retrieved from https://www.sdcexec.com/sourcing-procurement/article/12174233/go-big-or-go-home (Accessed on July 17, 2017).

Garrett, R. (2017). The benefits of digital transformation. *Supply & Demand Chain Executive,* Retrieved from https://www.sdcexec.com/sourcing-procurement/article/12328491/the-benefits-of-digital-transformation (Accessed on July 17, 2017).

Ghasemaghaei, M., Ebrahimi, S., & Hassanein, K. (2017a). Data analytics competency for improving firm decision making performance. *The Journal of Strategic Information Systems*.

Ghasemaghaei, M., Hassanein, K., & Turel, O. (2017b). Increasing firm agility through the use of data analytics: The role of fit. *Decision Support Systems*, *101*, 95-105.

Ghoshal, A., Larson, E., Subramanyam, R., & Shaw, M. (2014). The impact of business analytics strategy on social, mobile, and cloud computing adoption. In: Proceedings of the Thirty-fifth International Conference on Information Systems, Auckland, New Zealand December 14-17.

Giannakis, M., & Louis, M. (2016). A multi-agent based system with big data processing for enhanced supply chain agility. *Journal of Enterprise Information Management*, *29*(5), 706-727.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational research methods*, *16*(1), 15-31.

Godwin, K. (2016). How distributors can win with scan-based trading. *Industrial Distribution (Online),* Retrieved from https://www.inddist.com/article/2016/06/how-distributors-can-win-scan-based-trading (Accessed on July 14, 2017).

Goes, P. B. (2014). Big Data and IS Research. *MIS Quarterly*, *38*(3), iii-viii.

Goll, I., & Rasheed, A. A. (2004). The moderating effect of environmental munificence and dynamism on the relationship between discretionary social responsibility and firm performance. *Journal of Business Ethics*, *49*(1), 41-54.

Günther, W. A., Mehrizi, M. H. R., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems*.

Gunasekaran, A., Yusuf, Y. Y., Adeleye, E. O., & Papadopoulos, T. (2017). Agile manufacturing practices: the role of big data and business analytics with multiple case studies. *International Journal of Production Research*, 1-13.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049-1064.

Hadavi, C. (2016). Self-Improving supply chains have arrived. *Inbound Logistics,* Retrieved from http://www.inboundlogistics.com/cms/article/self-improving-supply-chains-have-arrived/ (Accessed on July 11, 2017).

Hahn, G. J., & Packowski, J. (2015). A perspective on applications of in-memory analytics in supply chain management. *Decision Support Systems*, *76*, 45-52.

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, *154*, 72-80.

Hazen, B. T., Skipper, J. B., Ezell, J. D., & Boone, C. A. (2016). Big Data and predictive analytics for supply chain sustainability: A theory-driven research agenda. *Computers & Industrial Engineering*, *101*, 592-598.

Henke, N., Bughin, J., & Chui, M. (2016). Most industries are nowhere close to realizing the potential of analytics. *Harvard Business Review*. Available from https://hbr.org/2016/12/most-industries-are-nowhere-close-to-realizing-the-potential-of-analytics (Retrieved October 9th, 2018).

Hinkin, T. R., & Tracey, J. B. (1999). An analysis of variance approach to content validation. *Organizational Research Methods*, *2*(2), 175-186.

Hoehle, H., & Venkatesh, V. (2015). Mobile Application Usability: Conceptualization and Instrument Development. *Mis Quarterly*, *39*(2).

Holsapple, C., Lee-Post, A., & Pakath, R. (2014). A unified foundation for business analytics. *Decision Support Systems*, *6Ex4*, 130-141.

Hurwitz, J., Nugent, A., Hapler, F., Kaufman, M., 2013. Big Data for Dummies. John Wiley & Sons, Hoboken, New Jersey.

International Institution for Analytics, 2016. Advanced Analytics & Big Data Adoption Report 2016. Available from https://iianalytics.com/analytics-resources/advanced-analytics-big-data-adoption-report-2016 (Retrieved January 31th, 2018)

Ittmann, H. W. (2015). The impact of big data and business analytics on supply chain management. *Journal of Transport and Supply Chain Management*, *9*(1), 1-9.

Jeseke, M., Grüner, M., & Weiß, F. (2013). Big data in logistics: A DHL perspective on how to move beyond the hype. White Paper; DHL Customer solutions & innovation, presented by Martine Wegner, Vice President Solutions & Innovation, 53844 Troisdorf, Germany.

Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*, *55*(17), 5011-5026.

Julian, S. D., & Ofori-Dankwa, J. C. (2017). Context matters: Diversity's short-and long-term effects in fortune's "best companies to work for". *Strategic Management Journal*, *38*(7), 1557-1565.

Kache, F., & Seuring, S. (2017). Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. *International Journal of Operations & Production Management*, *37*(1), 10-36.

Karimi, J., Somers, T. M., & Bhattacherjee, A. (2007). The role of information systems resources in ERP capability building and business process outcomes. *Journal of Management Information Systems*, *24*(2), 221-260.

Karlskind, M. (2014). How 5 companies use mobile apps to manage big data. *Material Handling & Logistics*, Retrieved from http://spot.lib.auburn.edu/login?url=https://search.proquest.com/docview/1637434936?accountid=8421 (Accessed on July 14, 2017).

Kartal, H., Oztekin, A., Gunasekaran, A., & Cebi, F. (2016). An integrated decision analytic framework of machine learning with multi-criteria decision making for multi-attribute inventory classification. *Computers & Industrial Engineering*, *101*, 599-613.

Keats, B. W., & Hitt, M. A. (1988). A causal model of linkages among environmental dimensions, macro organizational characteristics, and performance. *Academy of management journal*, *31*(3), 570-598.

Khan, K. (2013). The transformative power of advanced analytics. *Supply Chain Management Review*, *17*(3), 48-49.

Kim, S., Kim, H., & Park, Y. (2017). Early detection of vessel delays using combined historical and real-time information. *Journal of the Operational Research Society*, *68*(2), 182-191.

Kim, G., Shin, B., & Kwon, O. (2012). Investigating the value of sociomaterialism in conceptualizing IT capability of a firm. *Journal of Management Information Systems*, *29*(3), 327-362.

Kiviet, J. F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of econometrics*, *68*(1), 53-78.

Koberg, C. S., Detienne, D. R., & Heppard, K. A. (2003). An empirical test of environmental, organizational, and process factors affecting incremental and radical innovation. *The Journal of High Technology Management Research*, *14*(1), 21-45.

Kortmann, S., Gelhard, C., Zimmermann, C., & Piller, F. T. (2014). Linking strategic flexibility and operational efficiency: The mediating role of ambidextrous operational capabilities. *Journal of Operations Management*, *32*(7-8), 475-490.

Krippendorff, K. (2012). *Content analysis: An introduction to its methodology*. Sage Publications.

Krishnamoorthi, S., & Mathew, S. K. (2018). Business Analytics and Business Value: A Comparative Case Study. *Information & Management*.

Lade, P., Ghosh, R., & Srinivasan, S. (2017). Manufacturing Analytics and Industrial Internet of Things. *IEEE Intelligent Systems*, *32*(3), 74-79.

Lam, H. K., Yeung, A. C., & Cheng, T. E. (2016). The impact of firms' social media initiatives on operational efficiency and innovativeness. *Journal of Operations Management*, *47*, 28-43.

Lau, T., Man, T. W., & Chow, I. (2004). Organizational capabilities and performance of SMEs in dynamic and stable environments. *The International Journal of Entrepreneurship and Innovation*, *5*(4), 221-229.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT sloan management review*, *52*(2), 21.

Levans, M. (2011). 2011 technology roundtable: Putting data into context. *Logistics Management (2002)*, *50*(5), 42-44,46.

Li, S., Shang, J., & Slaughter, S. A. (2010). Why do software firms fail? Capabilities, competitive actions, and firm survival in the software industry from 1995 to 2007. *Information Systems Research*, *21*(3), 631-654.

Li, J., Tao, F., Cheng, Y., & Zhao, L. (2015). Big data in product lifecycle management. *The International Journal of Advanced Manufacturing Technology*, *81*(1-4), 667-684.

Liu, P., & Yi, S. P. (2017). A study on supply chain investment decision-making and coordination in the Big Data environment. *Annals of Operations Research*, 1-19.

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems*, *24*(3), 149-157.

Lustig, I., Dietrich, B., Johnson, C., & Dziekan, C. (2010). The analytics journey. *Analytics Magazine*, *3*(6), 11-13.

Lycett, M. (2013). 'Datafication': making sense of (big) data in a complex world. *European Journal of Information Systems*, *22*(4), 381-386.

MacKenzie, S. B. (2003). The dangers of poor construct conceptualization. *Journal of the Academy of Marketing Science*, *31*(3), 323-326.

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS quarterly*, *35*(2), 293-334.

Makadok, R. (2001). Toward a synthesis of the resource-based and dynamic-capability views of rent creation. *Strategic management journal*, *22*(5), 387-401.

Mani, V., Delgado, C., Hazen, B. T., & Patel, P. (2017). Mitigating supply chain risk via sustainability using big data analytics: evidence from the manufacturing supply chain. *Sustainability*, *9*(4), 608.

Marr B. (2015). A brief history of big data everyone should read. Available from https://www.weforum.org/agenda/2015/02/a-brief-history-of-big-data-everyone-should-read/ (Retrieved Decembe15th, 2017).

Matthias, O., Fouweather, I., Gregory, I., & Vernon, A. (2017). Making sense of Big Data–can it transform operations management? *International Journal of Operations & Production Management*, *37*(1), 37-55.

McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard business review*, *90*(10), 60-68.

McCrea, B. (2014). Overhead Handling: Raise Worker Safety. *Modern Materials Handling 69*(12), 34-7.

Michel, R. (2016). Turning data capture into decisions. *Modern Materials Handling*, *71*(7), 48-50,52,54-55.

Mishra, D., Gunasekaran, A., Papadopoulos, T., & Childe, S. J. (2016). Big Data and supply chain management: a review and bibliometric analysis. *Annals of Operations Research*, 1-24.

Mitchell, P. (2012). SUPPLY ANALYTICS: An overlooked opportunity. *Supply Chain Management Review, 16*(4), 26-33.

Mukherji, P., Sorescu, A., Prabhu, J. C., & Chandy, R. K. (2011). Behemoths at the gate: how incumbents take on acquisitive entrants (and why some do better than others). *Journal of Marketing*, *75*(5), 53-70.

Najjar, M. S., & Kettinger, W. J. (2013). Data Monetization: Lessons from a Retailer's Journey. *MIS Quarterly Executive, 12*(4).

NewVantage Partners LLC., 2017. Big Data Executive Survey 2017: Executive Summary of Findings. Available from: http://newvantage.com/wp-content/uploads/2017/01/Big-Data-Executive-Survey-2017-Executive-Summary.pdf (Retrieved December 29th, 2017).

Nguyen, T., Zhou, L., Spiegler, V., Ieromonachou, P., & Lin, Y. (2017). Big data analytics in supply chain management: A state-of-the-art literature review. *Computers & Operations Research*.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, 1417-1426.

Ou, T. Y., Cheng, C. Y., Chen, P. J., & Perng, C. (2016). Dynamic cost forecasting model based on extreme learning machine-A case study in steel plant. *Computers & Industrial Engineering*, *101*, 544-553.

Palmer, T. B., & Wiseman, R. M. (1999). Decoupling risk taking from income stream uncertainty: A holistic model of risk. *Strategic Management Journal*, *20*(11), 1037-1062.

Palmquist, D., Leal, T. (2016). The IoT supply chain: 3 things you should know. *Inbound Logistics*, Retrieved from http://www.inboundlogistics.com/cms/article/iot-in-the-supply-chain/ (Accessed on July 11, 2017).

Partridge, A. R. (2014). Materials Handling Meets Big Data. *Inbound Logistics*, Retrieved from http://www.inboundlogistics.com/cms/article/materials-handling-meets-big-data/ (Accessed on July 10, 2017).

Pavlou, P. A., & El Sawy, O. A. (2010). The “third hand”: IT-enabled competitive advantage in turbulence through improvisational capabilities. *Information systems research*, *21*(3), 443-471.

Pearson, M. (2011). Predictive analytics: Looking forward to better supply chain decisions. *Logistics Management (2002), 50*(9), 22-23.

Perreault, W. D., & Leigh, L. E. (1989). Reliability of nominal data based on qualitative judgments. *Journal of marketing research*, *26*(2), 135.

Pigni, F., Piccoli, G., & Watson, R. (2016). Digital Data Streams: Creating value from the real-time flow of big data. *California Management Review*, *58*(3), 5-25.

Popovič, A., Hackney, R., Tassabehji, R., & Castelli, M. (2018). The impact of big data analytics on firms' high value business performance. *Information Systems Frontiers*, *20*(2), 209-222.

Rai, A., Pavlou, P. A., Im, G., & Du, S. (2012). Interfirm IT capability profiles and communications for cocreating relational value: evidence from the logistics industry. *MIS quarterly*, *36*(1), 233-262.

Ravichandran, T., & Lertwongsatien, C. (2005). Effect of information systems resources and capabilities on firm performance: A resource-based perspective. *Journal of management information systems*, *21*(4), 237-276.

Richey Jr, R. G., Morgan, T. R., Lindsey-Hall, K., & Adams, F. G. (2016). A global exploration of big data in the supply chain. *International Journal of Physical Distribution & Logistics Management*, *46*(8), 710-739.

Roberts, N., & Grover, V. (2012). Leveraging information technology infrastructure to facilitate a firm's customer agility and competitive activity: An empirical investigation. *Journal of Management Information Systems*, *28*(4), 231-270.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. Stata Journal, 9 (1), 86-136. *For Online Publication*.

Ryan, G.W., & Bernard, H. R. (2000). Data management and analysis methods. N. Denzin, Y. Lincoln, eds. *Handbook of Qualitative Research,* 2nd ed. Thousand Oaks, CA, Sage, 769-802.

Sabherwal, R., & Chan, Y. E. (2001). Alignment between business and IS strategies: A study of prospectors, analyzers, and defenders. *Information systems research*, *12*(1), 11-33.

Sanders, N. R. (2016). How to use big data to drive your supply chain. *California Management Review*, *58*(3), 26-48.

Schoenherr, T., & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: Current state and future potential. *Journal of Business Logistics*, *36*(1), 120-132.

Seddon, P. B., Constantinidis, D., Tamm, T., & Dod, H. (2017). How does business analytics contribute to business value?. *Information Systems Journal*, *27*(3), 237-269.

Shah, S., Horne, A., & Capellá, J. (2012). Good data won't guarantee good decisions. *Harvard Business Review*, *90*(4).

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems*, *23*(4), 433-441.

Shenkman, D., Johnson, C., & Elliot, J. (2016). How they did it: Walgreens' talent strategy. *Supply Chain Management Review*, *20*(4), 35-39.

Souza, G. C. (2014). Supply chain analytics. *Business Horizons*, *57*(5), 595-605.
Steppeler, B. (2013). Make: Take One Tablet and Get Instant Relief. *Material Handling & Logistics*.

Straub, D., Boudreau, M. C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *The Communications of the Association for Information Systems*, *13*(1), 63.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research techniques*. Sage publications.

Subramani, M., & Walden, E. (2001). The impact of e-commerce announcements on the market value of firms. *Information Systems Research*, *12*(2), 135-154.

Suhas, S. (2017). Four ways the Internet of Things will transform the supply chain. *Material Handling & Logistics*, Retrieved from http://www.mhlnews.com/technology-automation/four-ways-internet-things-will-transform-supply-chain (Accessed on July 11, 2017).

Szakonyi, M. (2014). Big data to the rescue. *Journal of Commerce (1542-3867)*, *15*(10), 64-65.

Tambe, P. (2014). Big data investment, skills, and firm value. *Management Science*, *60*(6), 1452-1469.

Tan, K. H., Zhan, Y., Ji, G., Ye, F., & Chang, C. (2015). Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. *International Journal of Production Economics*, *165*, 223-233.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic management journal*, 509-533.

Teece, D. J. (2009). *Dynamic capabilities and strategic management,* Oxford University Press, Oxford.

Teo, T. S., Nishant, R., & Koh, P. B. (2016). Do shareholders favor business analytics announcements?. *The Journal of Strategic Information Systems*, *25*(4), 259-276.

Tiwari, S., Wee, H. M., & Daryanto, Y. (2018). Big data analytics in supply chain management between 2010 and 2016: Insights to industries. *Computers & Industrial Engineering*, *115*, 319-330.

Trebilcock, B. (2013). The big picture on BIG DATA. *Logistics Management*, *52*(10), 53-57.

Trebilcock, B. (2015). Cisco's warehouse of the future. *Modern Materials Handling, 70*, 24-28.

Trent, R. J. (2017). 5 supply chain trends happening now. *Supply Chain Management Review, 21*(2), 18-24.

Trkman, P., McCormack, K., De Oliveira, M. P. V., & Ladeira, M. B. (2010). The impact of business analytics on supply chain performance. *Decision Support Systems*, *49*(3), 318-327.

Tushman, M. L., & Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative science quarterly*, 439-465.

Urquhart, C., Lehmann, H., & Myers, M. D. (2010). Putting the 'theory' back into grounded theory: guidelines for grounded theory studies in information systems. *Information systems journal*, *20*(4), 357-381.

Viswanathan, N. (2018). Using Prescriptive Analytics to Decrease Supply Chain Disruptions. *Supply Chain Management Review.* Retrieved from http://www.scmr.com/article/are_you_using_prescriptive_analytics_to_decrease_supply_chain_disruptions_y (Accessed on July 29, 2018).

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77-84.

Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, *176*, 98-110.

Wang, N., Liang, H., Zhong, W., Xue, Y., & Xiao, J. (2012). Resource structuring or capability building? An empirical study of the business value of information technology. *Journal of Management Information Systems*, *29*(2), 325-367.

Wang, Y., & Hajli, N. (2017). Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, *70*, 287-299.

Wang, Y., *Kung*, L., Wang, W. Y. C., & Cegielski, C. G. (2017). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management*.

Ward, T., & Gopal, V. (2014). How they did it: Moving IBM's smarter supply chain to the cloud. *Supply Chain Management Review, 18*(2), 26-31.

Watson, M. (2012). Discovering the value of analytics. *Inbound Logistics*, Retrieved from http://www.inboundlogistics.com/cms/article/discovering-the-value-of-analytics/ (Accessed on July 11, 2017).

Weill, P., & Vitale, M. (2002). What IT infrastructure capabilities are needed to implement e-business models. *MIS quarterly Executive*, *1*(1), 17-34.
Wixom, B. H., Yen, B., & Relich, M. (2013). Maximizing Value from Business Analytics. *MIS Quarterly Executive*, *12*(2), 111-123.

Wetzels, M., Odekerken-Schröder, G., & Van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS quarterly*, 177-195.

Woerner, S. L., & Wixom, B. H. (2015). Big data: extending the business strategy toolbox. *Journal of Information Technology*, *30*(1), 60-62.

Woods, R. (2017). How smart ULDs are becoming essential Big-Data tools. *Air Cargo World, 107*(3), 22-24.

Wright, J. (2016). Turning data into insight. *Inbound Logistics,* Retrieved from http://www.inboundlogistics.com/cms/article/joanne-wright-turning-data-into-insight/ (Accessed on July 11, 2017).

Xie, K., Wu, Y., Xiao, J., & Hu, Q. (2016). Value co-creation between firms and customers: The role of big data-based cooperative assets. *Information & Management*, *53*(8), 1034-1048.

Xue, L., Ray, G., & Sambamurthy, V. (2012). Efficiency or innovation: How do industry environments moderate the effects of firms' IT asset portfolios?. *MIS quarterly*, 509-528.

Xue, L., Ray, G., & Gu, B. (2011). Environmental uncertainty and IT infrastructure governance: A curvilinear relationship. *Information Systems Research*, *22*(2), 389-399.

Yao, G., Wu, C. H., & Yang, C. T. (2008). Examining the content validity of the WHOQOL-BREF from respondents' perspective by quantitative methods. *Social Indicators Research*, *85*(3), 483-498.

Zhong, R. Y., Huang, G. Q., Lan, S., Dai, Q. Y., Chen, X., & Zhang, T. (2015). A big data approach for logistics trajectory discovery from RFID-enabled production data. *International Journal of Production Economics*, *165*, 260-272.

Zhong, R. Y., Lan, S., Xu, C., Dai, Q., & Huang, G. Q. (2016 a). Visualization of RFID-enabled shop floor logistics Big Data in Cloud Manufacturing. *The International Journal of Advanced Manufacturing Technology*, *84*(1-4), 5-16.

Zhong, R. Y., Newman, S. T., Huang, G. Q., & Lan, S. (2016 b). Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives. *Computers & Industrial Engineering, 101*, 572-591.

Zhong, R. Y., Xu, C., Chen, C., & Huang, G. Q. (2017). Big Data Analytics for Physical Internet-based intelligent manufacturing shop floors. *International Journal of Production Research*, *55*(9), 2610-2621.

Zhu, S., Song, J., Hazen, B. T., Lee, K., & Cegielski, C. (2018). How supply chain analytics enables operational supply chain transparency: An organizational information processing theory perspective. *International Journal of Physical Distribution & Logistics Management*, *48*(1), 47-68.

Zhu, S., Gupta, A., Paradice, D., & Cegielski, C. (2018). Understanding the Impact of Immersion and Authenticity on Satisfaction Behavior in Learning Analytics Tasks. *Information Systems Frontiers*, 1-24.