



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Ghasemaghaei, M;Turel, O

Title:

Possible negative effects of big data on decision quality in firms: The role of knowledge hiding behaviours

Date:

2021-03-01

Citation:

Ghasemaghaei, M. & Turel, O. (2021). Possible negative effects of big data on decision quality in firms: The role of knowledge hiding behaviours. *Information Systems Journal*, 31 (2), pp.268-293. <https://doi.org/10.1111/isj.12310>.

Persistent Link:

<https://hdl.handle.net/11343/286985>

Ghasemaghaei Maryam (Orcid ID: 0000-0001-7854-3177)
Turel Ofir (Orcid ID: 0000-0002-6374-6382)

Possible Negative Effects of Big Data on Decision Quality in Firms: The Role of Knowledge Hiding Behaviors

Dr. Maryam Ghasemaghaei
Email: ghasemm@mcmaster.ca

DeGroote School of Business
McMaster University
Hamilton, Ontario, Canada
(905) 525-9140

Dr. Ofir Turel
Email: oturel@fullerton.edu

Professor of Information Systems and Decision Sciences
California State University, Fullerton
(657) 278-5613

Corresponding author:
Dr. Maryam Ghasemaghaei
Phone: (905) 525-9140, ext 21721
Fax: (905) 521-8632

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as doi: [10.1111/isj.12310](https://doi.org/10.1111/isj.12310)

Acknowledgment

This work was supported by the Social Sciences and Humanities Research Council of Canada [grant number 20006922].

Abstract

While common wisdom suggests that big data facilitates better decisions, we posit that it may not always be the case, as big data aspects can also afford and motivate knowledge hiding. To examine this possibility, we integrate adaptive cost theory with the resource-based view of the firm. This integration suggests that the effect of big data characteristics (i.e., data variety, volume, and velocity) on firm decision quality can be explained, in part, by data analysts' perceived knowledge hiding behaviors, including evasive hiding, playing dumb, and rationalized hiding. We examined this model with survey data from 149 data analysts in firms that use big data to varying degrees. The findings show that big data characteristics have distinct effects on knowledge hiding behaviors. While data volume and velocity enhance knowledge hiding, data variety reduces it. Moreover, evasive hiding, playing dumb, and rationalized hiding have varying effects on firm decision quality. Whereas evasive hiding reduces firm decision-making quality, playing dumb

does not affect it, and rationalized hiding improves it. These results are further validated with applicability checks. Ultimately, these results can explain inconsistent past findings regarding the return on investment in big data and provide a unique look into the potential “dark sides” of big data.

Keywords: Big data, evasive hiding, playing dumb, rationalized hiding, decision quality, adaptive cost theory

Introduction

Big data (data that has the attributes of high volume, variety, and velocity (Ghasemaghaei, 2018a)) is an important aspect of modern organizations; it allows them to make better and more informed decisions (Clarke, 2016). Thus, many firms invest in big data initiatives (Agarwal & Dhar, 2014; Grover et al., 2018). However, there are mixed findings about the value of such investments (Zhang et al., 2017). Some studies suggest that big data initiatives improve firm decision-making quality (Chen et al., 2015; Wamba et al., 2017), which captures a firm’s ability to make optimal decisions (Sharma et al., 2014). Nevertheless, others argue that big data investment may quite often fail to improve decision quality and, indirectly, firm outcomes (Ghasemaghaei et al., 2017; Ghasemaghaei et al., 2018).

One potential reason for this inconsistency is the fact that decision making requires knowledge inputs, and these are often controlled by data analysts. Data analysts are often the key

gatekeepers and analysts of big data¹ (Ghasemaghaei & Calic, 2019). When knowledge insights stemming from big data use are not properly shared by data analysts, it follows that decision quality (and likely downstream outcomes such as performance) may not benefit from big data initiatives. In other words, one common way to impede the knowledge transfers needed for harvesting the value of big data, which has not been commonly discussed in the IS literature, is *knowledge hiding* (i.e., intentional attempts by individuals to conceal or withhold knowledge that has been requested by another person (Connelly et al., 2009)).

In general, employees often hide insights from peers (i.e., engage in knowledge hiding); 76% of employees reported withholding knowledge from other organizational members (Connelly et al., 2012). There are three prototypical knowledge hiding behaviors: evasive hiding, playing dumb, and rationalized hiding. Evasive hiding involves intentionally delaying knowledge delivery to the requester until the information is not useful anymore, and then promising to help the next time; playing dumb is defined as pretending to know nothing about the requested knowledge (Kang, 2016); and rationalized hiding is when the individual provides a legitimate justification for why the knowledge is not forthcoming (Connelly et al., 2012). The objective of this study is to

¹ With the advancement of technologies that enable firms to process big data, the role of data analysts has become critical (De Mauro et al., 2018) and jobs related to data analytics have become highly appealing and sought after (Davenport & Patil, 2012). According to PWC (2018), data analysts is among the most sought-after positions in North America. Whereas in the USA alone the number of open jobs asking for analytical skills was 2.3 million in 2015, this number increased to 2.9 million in 2018. There are still many companies that are seeking skillful data analysts who would be able to extract insights from the vast amounts of collected data (De Mauro et al., 2018).

assess how big data might influence knowledge hiding behaviors and can hence adversely affect decision-making quality in organizations².

To address this objective, we integrate two theoretical streams. First, we argue that knowledge hiding behaviors can disrupt the diffusion of knowledge stemming from big data and, as such, can reduce decision quality. This effect is rooted in the resource-based view (RBV) of the firm, which suggests that firms' resources (e.g., big data) serve to increase firm outcomes (e.g., quality of decisions) by enhancing firm capabilities (e.g., knowledge transformation) (Melville et al., 2004). In the context of big data, the knowledge that data analysts obtain by integrating, processing and analyzing the data is a vital input for decision making (Ghasemaghaei et al., 2017). Thus, if the knowledge that is obtained from processing big data is not shared with other employees, and therefore is not or is only partially accessible to decision-makers, it may hurt the quality of decisions. Hiding knowledge from other employees does not allow them to solve problems (Cross, 2001) and it limits their access to needed resources for decision making. This behavior may consequently and aggregately reduce the quality of decisions in the firm (Voelpel et al., 2005).

Second, we posit that the key attributes of big data (variety, volume, and velocity) may serve as factors that afford and motivate knowledge hiding behavior. This concern has been recently raised by several studies (e.g., Janssen et al., 2017; Sivarajah et al., 2017). For example,

² Note that our theory builds on the premise that both big data and data analysts play a key role in decision-making processes. This defines the contextual scope of the firms to which our theory applies.

Sivarajah et al. (2017) argue that one of the challenges of utilizing big data in firms is data analysts' difficulty in transferring the knowledge and their unwillingness to transfer the knowledge they obtained from big data. Ghasemaghaei (2018b) also found that big data utilization enhances task complexity perceptions among data analysts. These findings can be explained by the fact that by having access to big data, data analysts are often overwhelmed, which creates "tunnel vision" through which they see their needs and ignore others' requests (Webb, 2018). Indeed, data analysts often look for less demanding and lower-pressure work environments (Waters, 2017). Their pressure may be due, in part, to the nature of their job, which includes frequent changes in tools, data formats and demands, and increased reliance on their skills across business domains (Hass, 2005).

Such demanding jobs, especially ones that require high cognitive complexity, can motivate knowledge hiding (Cabrera & Cabrera, 2002). Therefore, we argue that data analysts can be motivated, to some extent, to hide knowledge from peers when feasible. Such knowledge hiding behaviors are rooted in adaptive cost theory, which describes a reduction in individuals' attentional capacity that is available for dealing with less relevant tasks when performing self-relevant complex tasks (Cohen, 1980). Applied to our context, it implies that when other employees seek knowledge from a data analyst, the complexity of obtaining and sharing knowledge from big data may affect the data analyst's motivation to comply with such knowledge sharing requests. Here, we used Connelly et al. (2012)'s framework in conjunction with adaptive cost theory to theorize

and examine how three key knowledge hiding behaviors (playing dumb, evasive hiding, and rationalized hiding) can be motivated by big data attributes.

To validate the proposed integrated model, we collected organizational-level data through the eyes of 149 data analysts in American firms who were highly familiar with big data use and decision making in their organizations. To further validate and enrich the findings, we performed applicability checks with 10 data analysts. The focus on knowledge hiding in the context of big data is not only unique in integrating RBV and adaptive cost theory but is also important for improving the understating of big data impacts on firms. Previous research has largely portrayed the positive impacts of big data (Ghasemaghaei et al., 2017). In this study, we show that under certain circumstances the availability of big data can promote undesirable behaviors such as knowledge hiding, which can ultimately hurt organizations. This work also extends insights regarding knowledge hiding, which has rarely been studied in information systems (IS) research and particularly in the context of big data. Ultimately, this study shows that IS, through the data attributes they support, can play a role in motivating and/or preventing knowledge hiding behaviors, and that through this mechanism, IS can affect decision quality in organizations. The findings point to guidelines that can help managers to better transform investments in big data initiatives into desirable firm outcomes.

Conceptual Background

Big Data Effects on Firm Outcomes

Big data often form through the integration of large amounts of unstructured and structured data, which can be accessed and processed in almost real-time; it is characterized by data variety, velocity, and volume (Ghasemaghaei et al., 2018). Big data plays an important role in creating knowledge within firms (Khan & Vorley, 2017; Müller et al., 2018; Müller et al., 2016). Specifically, it can affect the process of absorbing and transferring insights obtained from various sources (Chen et al., 2012; George et al., 2014; Khan & Vorley, 2017; Kitchens et al., 2018; Seddon et al., 2017). As business complexity and uncertainty increase, firms rely to a lesser extent on existing knowledge and instead attempt to rapidly create new knowledge (Chen et al., 2012; Ma & Agarwal, 2007; Majchrzak & Malhotra, 2016); this condition enhances the importance of and reliance on big data for firm success. Importantly, the new insights and knowledge obtained from big data can serve to improve firm decision-making quality (Boldsova, 2019; Brown et al., 2015; Ghasemaghaei et al., 2018) and the value extracted from big data (Cavaliere et al., 2015). Therefore, big data is an important input or resource for knowledge creation (Côte-Real et al., 2017). This knowledge can then be transferred among employees and be used for decision making (Zhang et al., 2017). Consequently, from an RBV standpoint, big data may be treated as an important organizational resource (Côte-Real et al., 2017; Kwon et al., 2014; Opresnik & Taisch, 2015; Wamba et al., 2017).

In line with this perspective, we use RBV to suggest that big data may influence firm decision-making quality by impacting employees' intra-firm knowledge transformations. RBV posits that firm resources serve as a basis for forming capabilities, which in turn influence firm

outcomes (e.g., decision-making quality) (Melville et al., 2004; Nwankpa & Datta, 2017; Zhang et al., 2008). Firm resources are “stocks of available factors that are owned or controlled by the firm” (Amit & Schoemaker, 2012; p. 35). According to RBV, firms need to effectively deploy their resources in order to enhance their outcomes (Chae et al., 2014). For example, having cash reserves (a resource) does not guarantee performance gains; only when it is distributed to work units and initiatives that create and sustain strategic advantage it will increase firms’ ability to improve outcomes.

Sprouting from RBV, the knowledge-based view emphasizes the role of knowledge in gaining competitive advantage in firms (Erickson & Rothberg, 2014). Based on this view, as knowledge-based resources are often difficult for other companies to imitate, the knowledge assets of firms are critical resources that enable firms to enhance their outcomes (Côte-Real et al., 2017). In firms, employees are considered the best source of knowledge because knowledge resides within individuals as they apply and generate knowledge in performing their tasks (Ghasemaghahi, 2019). Therefore, the movement of knowledge within firms depends on the dissemination of knowledge among employees. In the context of this study, the transformation of knowledge obtained from utilizing big data by data analysts and disseminating this knowledge to relevant work units can be viewed as critical firm capabilities that improve firm decisions (Ghasemaghahi et al., 2018). However, hiding knowledge from other employees (i.e., preventing resource dissemination) could interrupt this process, prevent the formation of capabilities based on these resources, and ultimately reduce decision-making quality in organizations. This is akin to having

cash reserves (see example in the above paragraph), but hiding them and therefore preventing the cash from reaching relevant units that need it for obtaining firm outcomes (e.g., buying raw materials for producing products).

Knowledge Hiding and Adaptive Cost Theory

The literature on workplace knowledge hiding behavior is an offspring of knowledge management research (Xiao & Cooke, 2019). The initial work mainly focused on data withholding by scientists (Campbell et al., 2000) because of the knowledge-intensive nature of academia. With the growing emphasis on information flows in organizations and the growth of the information economy, Davenport (1999) called for paying more attention to the inclination of the knowledge workers to disclose their knowledge. Since then, knowledge hiding has attracted an increasing amount of attention (e.g., Cabrera & Cabrera, 2002; Haas & Park, 2010) in different fields, such as organizational behavior (e.g., Connelly et al., 2012), human resource management (e.g., Černe et al., 2017), and tourism management (e.g., Zhao et al., 2016).

While disseminating knowledge obtained from big data is vital for enhancing firm decision-making quality, the dissemination task may not be easy and/or sufficiently motivated (Khan & Vorley, 2017). Dissemination can be done vertically, between subordinates to superiors, and it can also be done horizontally, between peers. Because vertical transfers are typically within-job roles and are considered a mandatory part of the job, consistent with Connelly et al. (2012), we focus here on horizontal transfers, which can also be highly instrumental for firm performance. In many cases, the analysts do not report to all knowledge requestors and therefore often (but not

exclusively) engage in horizontal transfers (Sonteya & Seymour, 2012). Importantly, because knowledge dissemination can be cognitively consuming, people often intentionally mitigate knowledge sharing requests and hide knowledge, especially if knowledge sharing is not a direct part of their job (Černe et al., 2014; Connelly & Zweig, 2015; Huo et al., 2016; Serenko & Bontis, 2016; Zhao et al., 2016). For example, an employee may ask a data analyst to disclose knowledge he or she obtained about consumers' opinions regarding a new product developed by the firm. The data analyst may agree to provide the knowledge but does not intend to (i.e., engage in evasive hiding), may pretend that she/he does not have this knowledge (i.e., engage in playing dumb), or may reply that the knowledge is confidential and it is only available to authorized employees in the firm (i.e., engage in rationalized hiding). These behaviors may help the analyst to focus on pressing tasks assigned to him/her (e.g., generating a different report).

The prerequisites of knowledge hiding are: 1) knowledge clearly requested by another firm member and 2) an intentional attempt (Xiao & Cooke, 2019). The motivation of employees to hide knowledge is often explained from the conservation of resources, adaptive cost theory, and psychological ownership perspectives. From the conservation of resources and adaptive cost theory perspectives, employees have a tunnel vision and focus on what they need to do well in their job; in this process, they may ignore requests from peers, the addressing of which is typically not a direct part of their job (Černe et al., 2014).

Based on the theory of psychological ownership, employees often develop a sense of ownership when they invest a substantial amount of effort, time, and attention into obtaining

specific knowledge (Pierce et al., 2001) and they may tend to consider it as personal intellectual property (Avey et al., 2009). Thus, transforming knowledge with co-workers may be seen as a threat to their knowledge ownership (Cabrera & Cabrera, 2002) and they may decide to hide knowledge instead of sharing it. In addition, territoriality, which is a type of psychological attachment to roles, and ideas also explains a mechanism underlying knowledge hiding behavior (Brown et al., 2005; Peng, 2013). Specifically, employees who have a strong feeling of psychological ownership of knowledge may prevent their co-workers from accessing to their knowledge territory.

Another explanation for knowledge hiding behavior is that individuals hide knowledge for political gain (Ipe, 2003). For example, when individuals have special knowledge they can enhance their bargaining power in their firms (Evans et al., 2015; Xiao & Cooke, 2019). One of the main reasons for not disclosing knowledge with other employees in the firm is that many employees may perceive potential personal costs (e.g., fear a loss of power or unique status) involved in disclosing the knowledge (Cress et al., 2005). For example, in the context of this paper, data analysts may hide the knowledge their analyses generate from other employees because they may fear they lose the power they have in being able to process big data and communicate the insights to top managers directly. As a result, many data analysts may intentionally attempt to conceal their knowledge or withhold it from other employees in the firm.

Although knowledge hiding and sharing may seem like the opposite extremes on a continuum, it has been shown that they are conceptually and empirically different constructs

(Connelly et al., 2012). While knowledge hiding typically means low sharing, it may also involve sharing of the wrong or partial information. For example, when data analysts process big data, they may be under pressure to finish their tasks on time. Thus, when an employee asks them to disclose the knowledge, the data analysts may just disclose partial information they obtained from processing the data. Not intending to hide knowledge does not mean there is high sharing; people may simply lack the requested knowledge, in which case they neither share nor hide knowledge. Further, note that while knowledge hiding has a negative connotation, it may have positive outcomes in some contexts (e.g., avoiding hurting someone's feelings). However, it typically has negative effects on firm performance metrics (Peng, 2013). According to Connelly & Zweig (2015), employees more often engage in knowledge hiding behavior when the knowledge is complex (e.g., insights from big data). Previous studies argued that while it is reasonable to assume that the factors that decrease unproductive knowledge behavior should enhance productive knowledge behavior, evidence suggests that as knowledge hiding and knowledge sharing are two distinct phenomena, they are motivated by different sources; thus, the antecedents of knowledge hiding may not necessarily apply to knowledge sharing (Connelly & Zweig, 2015; Serenko & Bontis, 2016).

When individuals work on complex job tasks, knowledge hiding can be explained via adaptive cost theory (Cohen, 1978). This theory argues that when individuals attempt to adapt to an environmental stressor, cognitive resources decrease, and the individual is forced to prioritize. It suggests that although humans are intrinsically adaptable to the environment, there is a cost and

limit to adapting to demands (Cohen, 1980). Consequently, most individuals focus attention to their own needs, attach a higher cost to supporting others' needs, and become less sensitive to the needs of others when facing high-task demands (Cohen, 1980; Cohen & Spacapan, 1978). This happens because individuals adapt to environmental demands by allocating capacity, effort, and attention to prioritized tasks, which are typically part of their direct job responsibilities (e.g., providing big data analytics findings for managers and not other colleagues who request such knowledge) (Cohen, 1980). As such, humans' limited cognitive resources force them to prioritize tasks that concern and benefit them, and this focus is stronger when people perform complex work tasks (Cohen, 1980; Cohen & Spacapan, 1978; Ford et al., 2015).

As employees prioritize tasks that are directly related to their job responsibilities, they often have no additional resources to spend on disseminating knowledge in the firm. Rooted in the stress literature, the adaptive cost theory argues that humans adapt to their environment (Cohen, 1980); however, adapting to job task stressors (e.g., task complexity) is accompanied by costs such as illness, injury, and fatigue. The resulting cost to individuals' behavior leads to having less motivation to engage in demands that are not directly related to their jobs. Adaptive cost theory has four main underlying assumptions: 1) individuals have limited attentional capacities; 2) a set of priorities is developed when the environmental demands exceed individuals' capacity. Thus, individuals spend their available capacity on the most relevant tasks at hand at the cost of those tasks that are less related to the job task performance; 3) the occurrence of environmental demands activates a monitoring process in which the individual assesses the significances and priority of

the stimulus and decides on the proper coping responses; and 4) high job task demands cause a temporary depletion in cognitive capacity and the recovery of the capacity happens with rest (Cohen & Spacapan, 1978; Ford et al., 2015).

Adaptive cost theory also justifies the reason for a reduction in interpersonal helping. Particularly, individuals' exposure to stressors (e.g., processing big data) will be accompanied by a reduction in helping (e.g., assisting an employee who is asking for the results of predictive analytics about customers' opinions about a new product). According to Cohen (1980, p. 95), "exposure to unpredictable and uncontrollable stress is followed by a decreased sensitivity to others." This is due to the fact that the cost of adapting to the complex tasks leaves the individual with less available resources to respond to other individuals' requests. In the work environment, transferring knowledge to colleagues is a type of interpersonal helping that may be considered as a non-directly-required organizational citizenship behavior (Kelloway & Barling, 2000). When individuals spend considerable attentional capacity on highly demanding tasks (e.g., processing large amounts of data and sending the results to the manager), there is a reduction in their interpersonal helping, including knowledge transfers to colleagues (Ford et al., 2015). This occurs because according to the adaptive cost theory, the individual is simply unable to transform the knowledge. Indeed, when a person is exposed to stressors (e.g., complex tasks with strict deadlines), he or she will be less likely to help someone who requested help or knowledge (Cohen, 1980; Cohen & Spacapan, 1978). Utilizing big data can be a complex task that could cause information overload for data analysts (Jones et al., 2004). Having too much data that takes much

of one's time to process, digest, and interpret causes stress (Edmunds & Morris, 2000) and can overwhelm data analysts (Edmunds & Morris, 2000; Tarafdar et al., 2013). After processing such data, the data analysts may not have enough cognitive resources to engage in knowledge transformation, especially if it is not their direct responsibility.

In adaptive cost theory, the time and effort individuals need to allocate for performing tasks that are less relevant to them are considered as psychological costs (Connelly et al., 2014; Sedighi et al., 2016). Cost perceptions refer to the concerns of putting in effort and taking time to share knowledge (Yan et al., 2016). In the context of this study, Khan and Vorley (2017) argue that making inferences based on big data requires data analysts to allocate considerable cognitive capacity; sharing such information also requires cognitive resources (e.g., finding ways to transfer a large file or report to another employee after removing identifiers; communicating and explaining the findings) and can distract data analysts from their current tasks. Given the costs associated with interpersonal-helping behavior relative to its limited relevance to their job performance, data analysts will be more motivated to decrease interpersonal-helping behavior (Van Dyne & LePine, 1998; Vandyne et al., 1995). Simply put, adaptive cost theory suggests that enhancing the perceived costs of acts of knowledge transfer increases individuals' motivation to withhold the knowledge they produce (Sedighi et al., 2016).

Hypotheses Development

Figure 1 shows the proposed research model. The dashed-line boxes map associations onto theories. Table 1 presents construct definitions.

----- insert Figure 1 about here -----

----- insert Table 1 about here -----

We first argue that when knowledge is hidden by analysts, regardless of the form of knowledge hiding, it will likely adversely impact firm outcomes (Connelly et al., 2012). This idea stems from the RBV of the firm, according to which firms need to properly leverage their resources to create capabilities that will help them improve firm outcomes (here decision-making quality) (Chen et al., 2014; Zhang et al., 2008). Based on the knowledge-based view, the knowledge obtained by analyzing big data is an important organizational resource (Côte-Real et al., 2017; Kwon et al., 2014; Opresnik & Taisch, 2015; Wamba et al., 2017), but its translation into capabilities and the resultant effect and value are network dependent; it is more valuable if it diffuses to an increasing number of relevant employees (Ghasemaghaei et al., 2017). Therefore, if big-data-based knowledge is hidden from other employees in need of such resources, it may hurt decision quality in firms (Voelpel et al., 2005). The logic is that hiding knowledge from other employees reduces the ability of the organization to leverage its knowledge resources (Cross, 2001), wastes time by requiring other employees to ‘reinvent the wheel’; and prevents learning (Connelly et al., 2014). In contrast, addressing co-workers’ knowledge sharing requests by transferring the required knowledge enables employees to find solutions in a timely and less costly manner (Miranda & Saunders, 2003; Sher & Lee, 2004).

There are different decision-makers at different levels of organization (from executives, to middle managers, to operators) who need to make a variety of decisions in their job tasks (Liu et al., 2009). Thus, although managers make strategic firm decisions, other employees in the firm

need to also make decisions as part of their job. To make effective and efficient decisions, all decision-makers, at all levels, need to have access to sufficient knowledge and resources. In firms that employ data analytics, data analysts are often a good source of knowledge – they have unique capabilities that allow them data gathering and analyses that most employees cannot easily perform. Here we focus on whether data analysts' knowledge hiding behavior would reduce the quality of decisions made by all types of employees in the firm. Because data analysts often help multiple stakeholders, not just one supervisor they report to (Sonteya & Seymour, 2012), they often engage in horizontal knowledge transfers (Liu et al., 2010), which afford knowledge hiding as such transfers can be controlled by the data analyst and are not always a direct part of their job.

Knowledge hiding behaviors take three forms: evasive hiding, playing dumb, and rationalized hiding (Connelly et al., 2012), all of which can adversely and aggregately affect decision-making quality. This is because regardless of the rationale and strategies for the hiding behaviors, the quality of decisions depends on the knowledge inputs used to make the decisions (Ryan et al., 2010). Because all forms of knowledge hiding may limit employees' access to needed resources (Grover et al., 2009; Kearns & Lederer, 2003), it is reasonable to expect that knowledge hiding decreases the ability of employees to make timely optimal decisions. Overall, knowledge hiding can be captured with a formative knowledge hiding concept, which accounts for the weighted summation of all forms of knowledge hiding (Fang, 2017). When these behaviors are aggregated to the firm level, it means that the more knowledge hiding there is in the firm, the lower the decision-making quality in the firm will be. Hence,

H1: *Knowledge hiding behaviors will decrease decision-making quality in firms.*

Increasing the volume of data elevates knowledge generation potential (Ghasemaghaei, 2019; Intezari & Gressel, 2017; Pauleen, 2017). This is especially true if firms integrate multiple data sources (Arora & Rahman, 2016; Baboo & Kumar, 2013; Zhang et al., 2017). It also means that more effort is likely invested in generating knowledge from voluminous data compared to lower-volume data and that sharing such knowledge may be more effortful (Jagadish et al., 2014; Shollo & Galliers, 2016). The resultant complexity of obtaining and sharing knowledge from voluminous data may motivate knowledge hiding (Ko et al., 2005). This may be explained via adaptive cost theory (Cohen, 1980), according to which hiding knowledge may simply reflect the stronger emphasis on one's own tasks compared to others' needs; one's own task demands may increase as a function of the volume of the data a person needs to process (Fan et al., 2014). That is, if data analysts spend large amounts of their available cognitive capacity and time on analyzing voluminous data, they will avoid cognitive demands from others, which may often be judged as counterproductive (Černe et al., 2014; Huo et al., 2016). Moreover, based on adaptive cost theory, data analysts may engage in knowledge hiding if the requested knowledge requires effort and time to explain; complex knowledge transfers have a greater imposition compared to simpler ones (Connelly et al., 2012). Such efforts may relate to the volume of data. For example, explaining, deleting, and transforming a large number of fields is more effortful compared to the same actions applied to a smaller number of fields. As extracting knowledge from large volumes of data often takes substantial time and effort, data analysts may tend to use different knowledge hiding

strategies to avoid disclosing the knowledge to others. Thus, based on adaptive cost theory, we argue that the volume of data the data analysts analyze for knowledge generation can drive knowledge hiding behavior. Hence,

H2: *Data volume will increase knowledge hiding behaviors.*

Organizational data have expanded from numerical and structured data to include also unstructured data (e.g., text and video), from outside and inside the firm (Intezari & Gressel, 2017). This increased data variety enables firms to discover hidden insights and generate new knowledge (Khan & Vorley, 2017). Nevertheless, it increases the complexity of data processing and analyses (Intezari & Gressel, 2017). For example, data analysts may understand consumers' preferences about a specific product by spending considerable effort to integrate data from consumer transactions and postings on social media. Even the integration of numerical data from different sources (e.g., different databases) is effortful and requires specialized skills, more than the use of a single and simple source of data. Arguably, this complexity may increase data analysts' knowledge hiding motivations (Ko et al., 2005) through the adaptive cost consideration described for H2. Additionally, according to cognitive load theory, the mental resources individuals have to solve complex tasks are limited (Oviatt, 2006). Hence, if the amount of information processing exceeds an individual's cognitive capacity, her or his attention to tasks (e.g., transferring knowledge to other colleagues) may get diluted (Ghasemaghahi, 2019). Processing multiple cues (e.g., processing structured and unstructured data) enhances the complexity of a task as individuals have limitations in their memory capacity (Wood, 1986). Consequently, data analysts may prefer

hiding the knowledge they obtained from utilizing various data sources, given that sharing and explaining data from multiple sources can be effortful and interrupt their own work tasks (Issac & Baral, 2018; Totterdell et al., 1995). Thus, as data variety increases, the complexity of data processing, analysis, and sharing increases too. Consequently, data analysts may be increasingly motivated to hide the knowledge they obtain from such data. Hence,

H3: *Data variety will increase knowledge hiding behaviors.*

Firms are attempting to enhance the speed of data integration by analyzing and processing data in almost real time so that they can understand and respond fast to events as they unfold (Intezari & Gressel, 2017). The problem is that the timeliness of the data and the speed at which they arrive can also, like data volume and variety, operate to increase the perceived complexity of knowledge generated from such data, enhance perceived effort to share such knowledge, and increase expected inability to cater to one's own needs. Specifically, given the increased complexity and effort that manifest from the increased velocity of data, adaptive cost theory, as described for H2 and H3, applies here. Additionally, according to the job-demand theory (Bakker et al., 2003), enhancing job demands creates perceptions of complexity and overload. In the context of this study, processing big data in real time can enhance the mental demand from data analysts and may tax their mental resources (Van Knippenberg et al., 2015), which makes processing big data a demanding task. Particularly, when data analysts need to process data in almost real time to provide new knowledge, they need to search for the most useful and relevant data and analyze the data before it becomes obsolete (Ghasemaghaei et al., 2018). Under these

increased time-pressure circumstances, they will further emphasize their own needs over knowledge requestors' needs (Connelly et al., 2014). Indeed, as per adaptive cost theory, time is an important barrier to disclose knowledge in firms (Haldin-Herrgard, 2000). As transferring the knowledge is often time consuming (Connelly et al., 2014; Hew & Hara, 2007), employees whose tasks are time sensitive may focus solely on their work demands to which they are mainly accountable (Connelly et al., 2009). Thus, if data analysts perceive that interpersonal helping decreases their available cognitive resources and the available time they have to accomplish their task, they will be less likely to engage in helping behaviors (Connelly et al., 2014). For example, when a data analyst is under pressure to meet a deadline but is asked by a co-worker to disclose knowledge, the data analyst may simply use his or her time constraint as an excuse (i.e., engage in rationalized hiding), may pretend that he or she does not have that specific knowledge (i.e., engage in playing dumb), or may agree to help the co-worker later but never intend to (i.e., engage in evasive hiding). Consequently, data velocity can create a sense of rush, complexity, and scarcity in one's work environment (Škerlavaj et al., 2018), and this can motivate knowledge hiding behaviors. Hence,

H4: *Data velocity will increase knowledge hiding behaviors.*

Methodology

Procedure

Data were collected with an online survey, which captured organizational-level data. The survey instrument was first pilot tested with 50 data analysts to ensure reliability of the constructs.

All scales were valid and reliable. Next, with the help of a market research firm, an invitation was sent to 785 data analysts. We assumed that they have reasonable knowledge about information flows in the organization and firm performance because they are typically a focal point for information analysis and dissemination. After all, it is their main role to integrate data from various sources and share insights with stakeholders. Thus, it is reasonable to expect that they are at least somewhat aware of whether the decisions made in their firm fit with the insights they obtain from data, and whether insights based on big data analyses flow well within the organization.

For compensation, participants had a chance to win 1 of the 5 monthly \$1000 prizes that were awarded by the research firm. To control for the potential impact of culture, survey participants were limited to data analysts in the United States. To participate in the study, participants had to have sufficient knowledge regarding big data utilization and knowledge transfers in their organizations. Thus, at the beginning of the survey, we asked participants about the extent of their awareness of the exchanges of insights from the use of big data analytics in their firm. Those participants that were not aware were excluded from our sample. In addition, we removed responses that 1) did not complete the survey, 2) completed the survey in less than 10 minutes (the estimated time to complete the survey was about 20 minutes), and 3) those participants that provided the same answer to most of the questions (e.g., all 7's). In order to reduce biases in self-reports, respondents were informed that there are no correct or incorrect answers and they were assured that their anonymity would be retained (Gable, 1994; Podsakoff et al., 2003).

Sample

A total of 149 responses (19% response rate) was retained. This response rate is in the upper half of Churchill and Iacobucci's (2006) guidance of 12–20% for acceptability in cross-sectional surveys, and it is consistent with existing research (e.g., Johnson et al., 2017; Joshi, 2016; Reid et al., 2015; Schleimer & Faems, 2016). We also assessed nonresponse bias following the guideline provided by Armstrong & Overton (1977). We compared the last and first quartiles of respondents in terms of demographic characteristics and key study variables. The results of this analysis did not show significant differences between late and early respondents, which indicates that nonresponse bias is unlikely an issue in the data.

Respondents were gender-balanced (49.7% female); 13.4% of participants were between 20 and 29 years old, 27.5% were between 30 and 39 years old, 22.1% were between 40 and 49 years old, and 37% were over 50 years old. Moreover, 36.9% of the participants had a graduate degree, and the rest had lower levels of education. In addition, participants represented multiple industries (manufacturing 25.5%, services 59.1%, finance 12.7%, and utilities 2.7%). Furthermore, 39.6% of the participants worked in firms that had lower than \$5 million in revenues, 51.7% of them worked in firms that had revenues between \$5 million and \$1 billion, and 8.7% of them worked in firms that had more than \$1 billion in revenues. Lastly, 58.4% of the participants worked in firms with less than 500 employees, 32.2% of them worked in firms with 500 to 5000 employees, and 9.4% of them worked in firms that had more than 5000 employees.

Measures

As shown in Table A1 in Appendix A, we used validated scales for all constructs. We also included control variables (firm size, revenue, and industry type) for accounting for inter-firm differences that may confound the results.

Results

Data were analyzed with Partial Least Square (PLS) version 3.0 (Ringle et al., 2005). Descriptive statistics, intra-construct correlations, composite reliabilities, average variance extracted (AVE), and Cronbach's alphas are given in Table 2³. All constructs were sufficiently valid and reliable. In addition, the loadings of the items on their proposed factor (Table 3) were larger than cross-loadings by at least 0.10, which is the recommended threshold (Gefen & Straub, 2005).

----- insert Table 2 about here -----

----- insert Table 3 about here -----

The measurement properties for knowledge hiding behavior (the second-order formative construct) were assessed using common guidelines (Bagozzi & Fornell, 1982; Cenfetelli & Bassellier, 2009). First-order construct composites (i.e., evasive hiding, playing dumb, and rationalized hiding) were generated by weighing the indicators. Next, we used the weighted sum

³ Table 2 demonstrates high correlations among playing dumb, evasive hiding, and rationalized hiding. The correlation between these variables could vary based on the context of the study. Similar to our findings, there are many papers (e.g., Pavlou & Fygenson, 2006; Sun, 2012) that used second-order formative constructs (for theoretical reasons) and still observed high correlations between their first-order constructs. Based on the guideline provided by Shin & Kim (2011), it is reasonable to view knowledge hiding behavior as being formed by evasive hiding, playing dumb, and rationalized hiding because 1) the three first-order constructs of knowledge hiding theoretically form knowledge hiding behavior, and 2) the deletion or addition of the sub-constructs changes the construct's (here knowledge hiding) conceptual domain. All the first-order constructs shape knowledge hiding behavior, but not the reverse. According to Tay (2017), the authors need to keep all the theoretically important variables in the model even in the case of having high correlations among them.

of the first-order variables to generate the composite index. The composite index was used as the measure for knowledge hiding behavior. Multicollinearity assessments showed that the Variance Inflation Factor (VIF) values were below the threshold of 3.3 (Diamantopoulos & Siguaw, 2006). In order to further confirm the validity of the second-order construct, we assessed the outer model loadings and weights. The results showed that the outer model weights of evasive hiding, playing dumb, and rationalized hiding (0.265, 0.265, and 0.272) are significant; this demonstrated the importance of all knowledge hiding first-order constructs in forming an overall knowledge hiding behavior concept. Additionally, the results showed that the outer model loading of all first-order variables (evasive hiding, playing dumb, and rationalized hiding) are higher than 0.70, which again show that each variable contributes to the formation of the second-order variable (knowledge hiding behavior) (Dwivedi et al., 2006). To examine common method variance (CMV) risk, we performed a marker-variable analysis (Lindell & Whitney, 2001) using distributive justice (i.e., employee perception that the rewards appropriately reflect their contributions to the organization) as the marker variable. The average correlation among the main constructs and the marker variable was 0.10. This suggests that CMV is unlikely to be pertinent in the data.

The results of the structural model estimation are presented in Figure 2. They supported most of the hypotheses but also revealed some surprising findings. Specifically, findings did not support H1 by showing a significantly *positive* (rather than negative) effect of knowledge hiding behavior on the quality of firm decisions ($\beta = 0.195; p < 0.01$). In addition, while as hypothesized,

data volume and data velocity increased knowledge hiding behavior ($\beta = 0.298$; $p < 0.01$, and $\beta = 0.378$; $p < 0.001$, respectively) and data variety, interestingly, *reduced* it ($\beta = -0.310$; $p < 0.01$).

We next calculated effect sizes (Chin, 2010) to examine the effect of each big data characteristic on knowledge hiding behavior. The findings showed similar effect sizes for the effect of each big data characteristic on knowledge hiding behavior. Specifically, the effect sizes of volume, variety, and velocity on evasive hiding were 0.05, 0.05, and 0.09, respectively.

----- insert Figure 2 about here -----

Post Hoc Analyses

First, in the structural model shown in Figure 2, we noticed the low R^2 value of the effect of knowledge hiding behavior on firm decision-making quality, which may be indicative that the separate knowledge hiding behaviors counteract. Thus, we post-hoc tested the unique impact of each first-order construct (evasive hiding, playing dumb, and rationalized hiding) on decision-making quality. As can be seen in Figure 3, evasive hiding reduced decision quality ($\beta = -0.352$; $p < 0.05$), playing dumb did not significantly influence decision quality ($\beta = -0.026$; $p > 0.05$), and rationalized hiding *increased* decision quality ($\beta = 0.576$; $p < 0.001$). We calculated effect sizes (Chin, 2010) to compare the impact of each knowledge hiding dimension on firm decision-making quality. The results showed that the effect sizes of evasive hiding, playing dumb, and rationalized hiding on firm decision-making quality were markedly different: 0.04, 0, and 0.12, respectively.

Second, we also examined the effect of each big data characteristic on evasive hiding, playing dumb, and rationalized hiding. Figure 3 illustrated that data volume increased evasive hiding, playing dumb, and rationalized hiding ($\beta = 0.240$; $p < 0.05$; $\beta = 0.258$; $p < 0.01$; $\beta = 0.335$;

$p < 0.001$, respectively). Data variety, interestingly, *decreased* evasive hiding and playing dumb ($\beta = -0.331$; $p < 0.01$; $\beta = -0.363$; $p < 0.001$), and had no significant effect on rationalized hiding ($\beta = -0.165$; $p > 0.05$). Data velocity significantly increased evasive hiding, playing dumb, and rationalized hiding ($\beta = 0.308$; $p < 0.01$; $\beta = 0.405$; $p < 0.001$; $\beta = 0.342$; $p < 0.001$, respectively). The results show that big data characteristics explain about 10% of the variance in evasive hiding, 16% in playing dumb, and 24% in rationalized hiding. Evasive hiding, playing dumb, and rationalized hiding explain 14% of the variance in decision-making quality.

----- insert Figure 3 about here -----

Third, while there were low correlations among data variety and other variables in the model (Table 2), the data variety effects were negative and significant (Figures 2 and 3). To understand the possible reason for this, we first examined the path coefficient between data variety and the dependent variables (evasive hiding, playing dumb, and rationalized hiding). The results showed that the impacts of variety on evasive hiding ($\beta = 0.045$; $p > 0.05$) and playing dumb ($\beta = 0.096$; $p > 0.05$) were not significant, but were significant on rationalized hiding ($\beta = 0.299$; $p < 0.05$). We then added data volume to the model. The findings showed the impacts of variety on evasive hiding ($\beta = -0.183$; $p > 0.05$), playing dumb ($\beta = -0.163$; $p > 0.05$), and rationalized hiding ($\beta = 0.003$; $p > 0.05$) were not significant. We then added data velocity to the model. The findings showed a significant impact of variety on evasive hiding ($\beta = -0.331$; $p < 0.05$) and on playing dumb ($\beta = -0.363$; $p < 0.05$), while the impact on rationalized hiding was not significant ($\beta = -0.165$; $p > 0.05$). These are the marginal effects of variety after controlling for the variance, which is explained by the other attributes.

A possible reason for such findings is that looking at the zero-order correlation is deceiving because such correlations do not account for the simultaneous roles of volume, velocity, and variety. In reality, data cannot have just one of these attributes, and all of these attributes operate in parallel. It is, therefore, more practical and theoretically precise to examine the effect of variety after accounting for volume and velocity effects. This is expressed in the research model but not in the correlation matrix (which reports zero-order, rather than partial correlations). The interpretation is that only after accounting for differences in all three big data characteristics, their effects on knowledge hiding become pronounced. Looking, for instance, just at data variety may produce non-significant results because variety may be in some cases associated with large volume and/or high velocity. According to Tay (2017), if the authors exclude one of the theoretically important variables from the model, the estimates of the coefficients of the remaining variables are very biased and could even predict the opposite effect. This is a serious mistake if the correct estimate of the coefficients between variables in the model is important for the purpose of the research. Thus, important correlated explanatory variables should not be excluded from the model even if they have very high or very low correlation with another explanatory variable, as the estimated coefficients would become very biased.

Lastly, we examined the impact of each control variable (i.e., industry type, revenue, and firm size) on the different types of knowledge hiding. The findings showed that the control variables did not significantly associate with any of the knowledge hiding types. Previous studies have argued that individual differences (e.g., age, gender, education) may impact employees'

knowledge hiding behaviors (Hershcovis et al., 2007; Peng, 2013; Wang & Noe, 2010). Thus, we also controlled for these factors to investigate whether they impact any of the knowledge hiding behavior types in the model. The findings showed that while gender did not impact any of the knowledge hiding behaviors ($\beta = 0.058$; $p > 0.05$, $\beta = 0.124$; $p > 0.05$, $\beta = 0.056$; $p > 0.05$ on evasive hiding, playing dumb, and rationalized hiding, respectively), age significantly reduced all knowledge hiding types ($\beta = -0.468$; $p < 0.05$, $\beta = -0.388$; $p < 0.05$, $\beta = -0.393$; $p < 0.05$ on evasive hiding, playing dumb, and rationalized hiding, respectively). The findings showed that education increased all knowledge hiding types ($\beta = 0.171$; $p < 0.05$, $\beta = 0.138$; $p < 0.05$, $\beta = 0.174$; $p < 0.05$ on evasive hiding, playing dumb, and rationalized hiding, respectively). These findings showed that as individuals become older, they report less knowledge hiding behaviors. The findings also showed that individuals with higher education levels report hiding more knowledge compared to those with less education level. However, gender did not impact knowledge hiding behaviors.

Applicability Checks

In order to further validate and enrich the findings, we performed applicability checks (Rosemann & Vessey, 2008). We communicated the results of the study without including any logic to 10 data analysts who did not participate in the main study and who were knowledgeable about the utilization of big data within their firms. We specifically recruited ten MBA students who had at least three years of experience being data analysts and we asked them to respond to specific questions shown in Table 4. Sample informative responses that describe whether the

findings make sense based on their own experiences are given in Table 4. The comments support the results we obtained from the survey.

----- insert Table 4 about here -----

Discussion

This study assessed one possible explanation for the mixed findings about the return on investment in big data analytics. We built on the resource-based view to suggest that knowledge hiding, an emergent and relatively unexplored phenomenon in IS research, prevents the translation of big data utilization and dissemination into improvements in decision-making quality. We then turned to the adaptive cost theory to suggest that big data characteristics can motivate knowledge hiding behavior. Hence, we suggest that big data can indirectly, through motivating knowledge hiding, have adverse consequences for firms. This is a unique perspective, as the extant literature has largely focused on the “bright side” of big data and the positive impact of big data on decision-making quality (Chen et al., 2015; Sharma et al., 2014; Wamba et al., 2017). The positive effect of big data on firm decision-making quality is mainly due to the fact that by analyzing large sizes of different types of data, firms are able to provide deep insights about their business, market, and environment. However, there are possible difficulties in extracting value from big data and there are only a few papers that have highlighted this (Ghasemaghaei et al., 2017; Ghasemaghaei et al., 2018). This study provides a novel contribution that points to one possible reason (certainly not the only one) for not being able to fully harvest the benefits from big data investments. Particularly, this study focuses on the fact that decision making requires knowledge inputs from data analysts, and such analysts can be motivated to conceal their insights for various reasons.

Building on this idea, the findings show that big data characteristics can indirectly contribute to a reduction in decision quality in firms (Janssen et al., 2017). They further provide a nuanced understanding of the varying effects of each big data characteristic on knowledge hiding behavior. Specifically, the results showed that data volume and velocity significantly increased knowledge hiding. This means that in line with adaptive cost theory, a data analyst who performs complex data gathering, processing and analysis tasks in a timely manner may perceive higher time pressure compared to others⁴ and thus, they have less attentional capacity and motivation available for focusing on less relevant tasks, such as information requests. That is, when data analysts need to process high-volume and fast-arriving data for generating knowledge, they will be more motivated to hide knowledge stemming from these data and increasingly focus on their own tasks. However, in contrast to our expectations, data variety significantly decreased knowledge hiding. One potential explanation is that working with a high variety of data is a specialized task, as opposed to working with data that are characterized by volume and velocity. Under such circumstances, it is more difficult for data analysts to “hide”, because such specialized skills are expected from them and are likely part of their job duties; in other words, sharing such knowledge may be more likely to be an in-role rather than an out-of-role duty. Thus, adaptive cost considerations may be less relevant to these circumstances. These propositions, though, merit

⁴ To measure if data analysts who process big data perceive higher time pressure compared to others, we classified our sample into 3 categories (low, medium, high) based on the level of big data they process. We used ANOVA to examine whether there is any significant difference between the perception of time pressure among these three categories. The findings showed that the data analysts who analyze high levels of big data perceived higher time pressure compared to those who process lower levels of big data. This provides additional evidence, beyond anecdotal statements from prior research, that the “bigger” the data that one analyzes, the more time pressure one feels.

future research. The findings of the post-hoc analysis also showed similar results; while data volume and velocity increase all knowledge hiding behaviors, data variety reduces them.

Another important contribution of this study relates to the nuanced explanation of the impacts of unique knowledge hiding behaviors on firm decision-making quality. In contrast to our expectations, the findings showed a positive effect of knowledge hiding on decision-making quality. To further explore this relationship, we conducted a post-hoc analysis that accounts for the effect of each knowledge hiding dimension on decision making quality. The results indicated that knowledge hiding behaviors have different effects on decision-making quality: evasive hiding reduces it, playing dumb does not significantly impact it, and rationalized hiding improves it. This means that evasive hiding prevents the knowledge flows that make knowledge a valuable resource, as per the knowledge-based view (Evangelou & Karacapilidis, 2007). It prevents peers from having access to information that they need and is valuable for them. However, if data analysts dishonestly claim that they do not have the specific knowledge their co-workers request (i.e., playing dumb), the co-workers may ask other employees in the firm for the knowledge until they obtain that knowledge. Therefore, playing dumb may slightly slow down decisions but does not significantly decrease decision-making quality. Moreover, data analysts may provide an explanation for why they are not able to disclose knowledge to their co-workers (rationalized hiding) (Zhao et al., 2016). Employees may be unauthorized to access such knowledge for a reason; these data may be irrelevant for them.

Thus, the positive impact of rationalized hiding on decision-making quality could be due to the fact that the knowledge the co-workers are seeking to obtain could be, in fact, unrelated to the decisions they want to make. Alternatively, knowledge requestors may be forced to process their own data which may lead them to gain more insights. This type of knowledge hiding is different in that the analyst does not delay the delivery of knowledge. Instead, he or she says to employees “I cannot help you”. This may propel employees to seek solutions on their own, which may result in better decisions. This idea, again, merits further research. Together, these findings are theoretically and practically important as there is a need to have a deeper understanding of the effect of big data on firms and a need to focus on knowledge hiding in the context of IS, as this is a growing yet underexplored problem that may be influenced by IS, its affordances, and attributes.

Theoretical Contributions

The results extend several theoretical perspectives on big data, knowledge hiding, and adaptive cost theory. First, the results extend the big data literature by showing the necessity to conceptually and operationally differentiate among key big data characteristics, rather than treating big data as a holistic concept. Specifically, this study shows interesting and unique findings regarding the direct effect of each big data characteristic on data analysts’ behaviors, and the indirect effect on firm outcomes (here decision-making quality). Thus, the findings extend the understanding of big data implications from organizational and information systems perspectives by showing how big data can adversely affect firm outcomes.

In addition, the findings point to the potential of different big data characteristics to produce different theory-driven effects. Specifically, whereas the positive effect of data velocity and volume on knowledge hiding dimensions reinforces adaptive cost theory considerations, the negative impact of data variety on knowledge hiding dimensions may need different theorizing. These findings show that perhaps the unique expertise related to processing data characterized by variety makes knowledge sharing perceived to be an in-role responsibility (Srivastava et al., 2006) and also makes it difficult to hide and avert requests for knowledge sharing. As such, these findings extend adaptive cost theory by showing that although there is a cost and limit to adapting to knowledge demands, if individuals find requested knowledge to be part of his/her specialized tasks, they would be less likely to hide knowledge from their peers. For example, if a data analyst has done a sentiment analysis about the attitude of customers about their product/services and a colleague asks for specific knowledge, the data analyst may disclose the knowledge even though he/she has spent considerable time and effort to collect and analyze the data. As suggested by Connelly et al. (2012), when the requested knowledge is straightforward, individuals may be less motivated to hide the knowledge. We make here first strides toward understanding the effects of big data characteristics on knowledge hiding and call for future research to open the black boxes we revealed here by assessing the specific mechanisms through which big data characteristics impact knowledge hiding behaviors. Overall, our findings offer an opportunity for other scholars to evaluate the impact of different big data characteristics in various contexts.

The results also extend the big data literature by theorizing on and testing the role of knowledge hiding in translating big data characteristics into firm outcomes (here decision quality). This perspective represents a unique integration of adaptive cost theory with the RBV of the firm. This perspective can resolve debates in the big data literature regarding why big data does not always enhance firm outcomes (Ghasemaghaei, 2018a; Ghasemaghaei et al., 2017; Ghasemaghaei et al., 2018). Our results demonstrate that big data characteristics have a similar magnitude of total effects on the dimensions of knowledge hiding. This shows that future studies need to consider not only the important positive influences of big data on the outcomes of the firms but also its potential adverse effects. That is, big data should join the myriad of IT that present both “dark” and “bright” sides in their effects on firms (Tarafdar et al., 2015). While most scholars focused on the positive impact of big data on firm outcomes, our findings highlight the possible negative impact of big data characteristics on firm capabilities. Thus, scholars need to also examine the potential pitfalls and risks of implementing big data within organizations. Future research can employ the theoretical integration we present here to examine possible effects of big data characteristics on other out-of-role behaviors, such as organizational citizenship behaviors, volunteering to help, and innovating and training others.

While calls have been issued to examine the effect of knowledge hiding dimensions on firm outcomes (Connelly et al., 2012, 2014), little empirical research in this domain was conducted. Similar to Wang et al. (2019), who argue the possible positive impact of knowledge hiding in the sales context, our findings show the possible *positive* impact of aggregated knowledge

hiding behavior on firm decision-making quality. Nevertheless, this may be deceiving. When looking at separate dimensions, our findings show that while evasive hiding reduces decision-making quality, playing dumb does not affect it, and rationalized hiding improves decision-making quality. These results emphasize the need to understand the effect of different dimensions of knowledge hiding on firm outcomes because their impacts can vary and counteract. While most scholars have focused on the negative influence of knowledge hiding on firm outcomes (Butt, 2019; Fong et al., 2018), the findings of our study are in line with the argument Connelly et al. (2012) had regarding the potential positive impact of knowledge hiding on firm outcomes. Particularly, they argue that in the same way that knowledge sharing is not always a beneficial behavior, knowledge hiding is also not necessarily a deviant or bad behavior. Similar to any other behavior (e.g., lying), knowledge hiding may have unintended, negative, or positive consequences. While studying knowledge hiding in organizations is still novel and most scholars have focused on the negative consequences of knowledge hidings, our research provides novel insights for scholars who are interested in understanding the consequences of knowledge hiding in the organizational context. Future studies could examine whether the use of other technologies in firms would have similar impacts on knowledge transformation and flow in firms and the consequent decision quality. By showing how each type of knowledge hiding impacts decision quality, we have provided an important foundation for future research on the positive and negative impacts of knowledge hiding in firms.

Lastly, the findings extend previous research that emphasizes the importance of interpersonal helping among employees within firms (Durach & Machuca, 2018; Wang et al., 2018) to the domain of decision-making quality; this study is the first to empirically show the importance of knowledge hiding dimensions for firm decision-making quality. This study also provides an extension of adaptive cost theory to the new context of big data. This theory is often used in organizational settings (Marques et al., 2019; Totterdell et al., 1995), but rarely in IS and specifically big data contexts. This study shows the importance of integrating IS and big data research with organizational-behavior literature on undesirable employee behaviors. It moreover illuminates the need to further study the direct impact of IS on each knowledge hiding dimension and its indirect impact on firm outcomes.

Practical Contributions

The findings suggest that in order to mitigate knowledge hiding, companies can change their big data attributes. However, this is unrealistic to expect and execute. Instead, relying on the theoretical basis for this research (adaptive cost theory), the findings imply that companies can strengthen the perception of knowledge transformation as an in-role behavior, either formally (through policies and reward and punishment mechanisms) or informally (through creating a sharing culture). This will reduce motivation for knowledge hiding because it will make knowledge sharing more related to one's in-role responsibilities. This need is exacerbated when firms are high in (at least some) big data attributes. Thus, companies should examine their big data attributes, and when these are high, consider formulating policies and strategies for mitigating

knowledge hiding. The finding that data variety can reduce evasive hiding and playing dumb is also noteworthy. It implies that firms may want to make these tasks a clear responsibility of a few experts; such a move will demotivate and reduce opportunities for knowledge hiding. For example, when employees ask data analysts about specific knowledge (e.g., consumers' comments on social media), as data analysts are expected to have this knowledge, analysts may not be able to hide that knowledge from other employees, and thus data variety may reduce knowledge hiding. However, future studies should examine this explanation (or others) to fully understand how data variety might reduce knowledge hiding. For example, future studies could examine the role of advanced analytical tools in this association and whether the analysts are the only ones who have access to unstructured data (in which case, they cannot "hide" it or avoid requests related to it).

The findings also indicate that firms that employ high-volume and high-velocity data unknowingly promote knowledge hiding. Although many firms have access to advanced big data technologies, the task of processing large volumes of near real-time data can still be very complex and requires effort and time. This decreases analysts' available cognitive capacity to deal with knowledge sharing requests. In general, depending on firms' policies and goals, they may mainly focus on those big data characteristics that could enhance firm outcomes. If a firm is willing to reduce data analysts' time pressure, it can purchase some advanced technologies to facilitate the work of data analysts or the firm can have a number of data analysts who work in a team and each person would be responsible for a specific task in processing big data. Firms should also train data

analysts to be able to handle the processing of big data. This will reduce the time pressure and the heavy workload a data analyst could perceive when he/she analyzes big data.

Based on the results, if data analysts provide useless information or agree to help the knowledge requester but never really do, then decision-making quality will diminish. However, if data analysts just pretend that they do not have the information, the knowledge seeker may try to find the knowledge from other sources. Most importantly, if data analysts provide a good explanation for why they cannot provide knowledge to their co-worker, the knowledge requester may perceive that the knowledge he or she is looking for is unlikely to be relevant to the decision she or he is making. As such, in line with Connelly et al. (2012), we show that knowledge hiding is not always deceptive and that it may have a positive impact on firm outcomes when it is intended to preserve confidentiality, protect other party's feeling, or serves the interests of the requestor (i.e., it saves them time and saves them from "barking up the wrong tree"). Thus, companies can measure knowledge hiding behaviors through direct tests or surveys, and train data analysts to avoid evasive hiding and/or playing dumb. They should simultaneously consider that rationalized hiding may sometimes be good, at least for decision-making quality, as it likely shelters employees from unnecessary and less relevant information for their decisions.

Limitations and Future Research

Several limitations of this study are noteworthy. First, our focus on data analysts, while reasonable, limits the generalizability of findings because analysts represent just one perspective in the organization. While they are expected to be aware of the focal constructs in our model, their

knowledge is imperfect. Perspectives of other members of the organization, such as CEOs or line workers, may also be imperfect. Thus, future research would benefit from using multiple perspectives and enriching the single-perspective approach we took here. Second, we investigated the impact of knowledge hiding on decision-making quality. Future studies could explore the effect of knowledge hiding on more distal outcomes (e.g., financial performance). Third, in this study, we investigated the impact of the three main big data characteristics. Recent studies provide a more nuanced view of big data characteristics (including 42 dimensions) (Shafer, 2018). Future studies could explore these characteristics on firm outcomes. Fourth, the effect of big data characteristics on firm decision making could be influenced by factors beyond the scope of the current study (e.g., innovation, leadership). These should be considered in future research. Fifth, the findings of this study show that each type of knowledge hiding behavior (i.e., evasive hiding, playing dumb, rationalized hiding) has different impacts on decision-making quality. Future studies could explore the mechanisms of the impact of each knowledge hiding behavior on firm outcomes. Sixth, the scale we used to measure decision quality focuses on decisions made at different levels of the organization (from executives, to middle managers, to line workers). Although managers make strategic firm decisions, other employees in the firm also need to make decisions as part of their jobs. Our focus is, therefore, on all employees in general, and all types of decisions. A key limitation of this scale is that it does not measure decision-making quality related to vertical transfers of knowledge in the firm. It would be interesting in future research to investigate whether there is a difference in decision-making quality when employees hide the knowledge they transfer

vertically compared to horizontally. Seventh, the findings point to conflicting insights regarding data variety using the correlation matrix vs. the structural model. While we provide an explanation for this, future studies could more systematically examine when and how low zero-order correlations manifest in significant relationships in structural models when adding other important variables to the model. Lastly, the R^2 values for knowledge hiding ($R^2=0.174$) and decision-making quality ($R^2 =0.043$) were not high. This implies that there are likely other variables (e.g., knowledge sharing, data analytics insight generation, analytical skills) that could impact these constructs. Future studies could extend our model with such variables.

Conclusion

Big data is important in modern firms. Nevertheless, its effects on firm outcomes have not been consistent. Here, we integrated adaptive cost theory with the resource-based view of the firm and showed that prior inconsistent results may be due to the fact that big data characteristics can directly promote knowledge hiding behaviors and indirectly adversely affect the quality of firm decision making. Thus, future research should carefully consider, while considering big data benefits, possible adverse effects of big data on firms. It should also consider the unique role of each big data attribute in influencing firm outcomes. In addition, it should better integrate knowledge hiding into IS research models, as information systems can afford (or block) and motivate (or demotivate) such typically undesirable behaviors. Ultimately, the findings of this study provide useful insights for researchers and practitioners and pave the way for deeper research of when and how big data use translates into improved firm performance.

References

- Agarwal, R., & Dhar, V. (2014). *Big data, data science, and analytics: The opportunity and challenge for IS research*. *Information Systems Research*, 3(25), 443–448.
- Amit, R., & Schoemaker, P. J. (2012). Strategic assets and organizational rent. *Strategic Management Journal*, 14, 33–46.
- Armstrong, J. S., & Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *Journal of Marketing Research*, 396–402.
- Arora, B., & Rahman, Z. (2016). Using Big Data Analytics for Competitive Advantage. *International Journal of Innovative Research and Development*, 5(2).
- Avey, J. B., Avolio, B. J., Crossley, C. D., & Luthans, F. (2009). Psychological ownership: Theoretical extensions, measurement and relation to work outcomes. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 30(2), 173–191.
- Baboo, L. D. S., & Kumar, P. R. (2013). Next generation data warehouse design with big data for big analytics and better insights. *Global Journal of Computer Science and Technology*.7.
- Bagozzi, R. P., & Fornell, C. (1982). Theoretical concepts, measurements, and meaning. *A Second Generation of Multivariate Analysis*, 2(2), 5–23.
- Bakker, A. B., Demerouti, E., De Boer, E., & Schaufeli, W. B. (2003). Job demands and job resources as predictors of absence duration and frequency. *Journal of Vocational Behavior*, 62(2), 341–356.
- Boldosova, V. (2019). Deliberate storytelling in big data analytics adoption. *Information Systems Journal*.
- Brown, D. E., Abbasi, A., & Lau, R. Y. K. (2015). Predictive Analytics: Predictive Modeling at the Micro Level. *IEEE Intelligent Systems*, 30(3), 6–8.
- Butt, A. S. (2019). Consequences of top-down knowledge hiding in firms: A pilot study. *Heliyon*, 5(12), e03000.
- Cabrera, A., & Cabrera, E. F. (2002). Knowledge-sharing dilemmas. *Organization Studies*, 23(5), 687–710.
- Campbell, E. G., Weissman, J. S., Causino, N., & Blumenthal, D. (2000). Data withholding in academic medicine: Characteristics of faculty denied access to research results and biomaterials. *Research Policy*, 29(2), 303–312.
- Cavaliere, V., Lombardi, S., & Giustiniano, L. (2015). Knowledge sharing in knowledge-intensive manufacturing firms. An empirical study of its enablers. *Journal of Knowledge Management*, 19(6), 1124–1145.
- Cenfetelli, R. T., & Bassellier, G. (2009). Interpretation of formative measurement in information systems research. *MIS Quarterly*, 689–707.
- Černe, M., Nerstad, C. G., Dysvik, A., & Škerlavaj, M. (2014). What goes around comes around: Knowledge hiding, perceived motivational climate, and creativity. *Academy of Management Journal*, 57(1), 172–192.

- Chae, B. K., Yang, C., Olson, D., & Sheu, C. (2014). 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.
- Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 1165–1188.
- Chen, Y., Wang, Y., Nevo, S., Jin, J., Wang, L., & Chow, W. S. (2014). IT capability and organizational performance: The roles of business process agility and environmental factors. *European Journal of Information Systems*, 23(3), 326–342.
- Chin, W. W. (2010). How to write up and report PLS analyses. In *Handbook of partial least squares* (pp. 655–690). Springer. http://link.springer.com/chapter/10.1007/978-3-540-32827-8_29
- Churchill, G. A., & Iacobucci, D. (2006). *Marketing research: Methodological foundations*. Dryden Press New York.
- Clarke, R. (2016). Big data, big risks. *Information Systems Journal*, 26(1), 77–90.
- Cohen, S. (1978). Environmental load and the allocation of attention. *Advances in Environmental Psychology*, 1, 1–29.
- Cohen, S. (1980). Aftereffects of stress on human performance and social behavior: A review of research and theory. *Psychological Bulletin*, 88(1), 82–108.
- Cohen, S., & Spacapan, S. (1978). The aftereffects of stress: An attentional interpretation. *Environmental Psychology and Nonverbal Behavior*, 3(1), 43–57.
- Connelly CE, D Zweig and J Webster. (2006). Knowledge hiding in organisations. *Paper Presented in the Symposium 'Don't Say a Word: Explaining Employees' Withholding of Knowledge from Coworkers', the Society for Industrial and Organizational Psychology Conference. Dallas, TX.*
- Connelly, C. E., Ford, D. P., Gallupe, B., Turel, O., & Zweig, D. (2009). The effects of competition and time constraints on knowledge transfer: Exploratory findings from two experiments. *2009 42nd Hawaii International Conference on System Sciences*, 1–10.
- Connelly, C. E., Ford, D. P., Turel, O., Gallupe, B., & Zweig, D. (2014). 'I'm busy (and competitive)!' Antecedents of knowledge sharing under pressure. *Knowledge Management Research & Practice*, 12(1), 74–85.
- Connelly, C. E., & Zweig, D. (2015). How perpetrators and targets construe knowledge hiding in organizations. *European Journal of Work and Organizational Psychology*, 24(3), 479–489.
- Connelly, C. E., Zweig, D., Webster, J., & Trougakos, J. P. (2012). Knowledge hiding in organizations. *Journal of Organizational Behavior*, 33(1), 64–88.
- 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.

- Cress, U., Barquero, B., Buder, J., & Hesse, F. W. (2005). Social dilemma in knowledge communication via shared databases. In *Barriers and biases in computer-mediated knowledge communication* (pp. 143–167). Springer.
- Cross, P. (2001). Knowingwhatweknow: Supporting knowledge creation and sharing in social networks. *Organizational Dynamics*, 30(2), 100–120.
- Davenport, T. O. (1999). *Human capital: What it is and why people invest it*. Jossey-Bass Inc Pub.
- Davenport, T. H., & Patil, D. J. (2012). Data scientist. *Harvard Business Review*, 90(5), 70–76.
- De Mauro, A., Greco, M., Grimaldi, M., & Ritala, P. (2018). Human resources for Big Data professions: A systematic classification of job roles and required skill sets. *Information Processing & Management*, 54(5), 807–817.
- Diamantopoulos, A., & Siguaw, J. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management*, 17(4), 263–282.
- Durach, C. F., & Machuca, J. A. (2018). A matter of perspective—the role of interpersonal relationships in supply chain risk management. *International Journal of Operations & Production Management*, 38 (10), 1866-1887.
- Dwivedi, Y. K., Choudrie, J., & Brinkman, W.-P. (2006). Development of a survey instrument to examine consumer adoption of broadband. *Industrial Management & Data Systems*, 106(5), 700–718.
- Edmunds, A., & Morris, A. (2000). The problem of information overload in business organisations: A review of the literature. *International Journal of Information Management*, 20(1), 17–28.
- Erickson, S., & Rothberg, H. (2014). Big data and knowledge management: Establishing a conceptual foundation. *Electronic Journal of Knowledge Management*, 12(2), 101.
- Evangelou, C. E., & Karacapilidis, N. (2007). A multidisciplinary approach for supporting knowledge-based decision making in collaborative settings. *International Journal on Artificial Intelligence Tools*, 16(06), 1069–1092.
- Evans, J. M., Hendron, M. G., & Oldroyd, J. B. (2015). Withholding the ace: The individual- and unit-level performance effects of self-reported and perceived knowledge hoarding. *Organization Science*, 26(2), 494–510.
- Fan, J., Han, F., & Liu, H. (2014). Challenges of big data analysis. *National Science Review*, 1(2), 293–314.
- Fang, Y.-H. (2017). Coping with fear and guilt using mobile social networking applications: Knowledge hiding, loafing, and sharing. *Telematics and Informatics*, 34(5), 779–797.
- Fong, P. S., Men, C., Luo, J., & Jia, R. (2018). Knowledge hiding and team creativity: The contingent role of task interdependence. *Management Decision*, 56(2), 329–343.
- Ford, D., Myrden, S. E., & Jones, T. D. (2015). Understanding “disengagement from knowledge sharing”: Engagement theory versus adaptive cost theory. *Journal of Knowledge Management*, 19(3), 476–496.

- Gable, G. G. (1994). Integrating case study and survey research methods: An example in information systems. *European Journal of Information Systems*, 3(2), 112–126.
- Gefen, D., & Straub, D. (2005). A practical guide to factorial validity using PLS-Graph: Tutorial and annotated example. *Communications of the Association for Information Systems*, 16(1), 5–26.
- George, G., Haas, M. R., & Pentland, A. (2014). *Big data and management*. Academy of Management Briarcliff Manor, NY.
- Ghasemaghaei, M. (2018a). Improving Organizational Performance Through the Use of Big Data. *Journal of Computer Information Systems*, 1–14.
- Ghasemaghaei, M. (2018b). The role of positive and negative valence factors on the impact of bigness of data on big data analytics usage. *International Journal of Information Management*, 50, 395–404.
- Ghasemaghaei, M. (2019). Does data analytics use improve firm decision making quality? The role of knowledge sharing and data analytics competency. *Decision Support Systems*, 120, 14–24.
- Ghasemaghaei, M., & Calic, G. (2019). Does big data enhance firm innovation competency? The mediating role of data-driven insights. *Journal of Business Research*, 104, 69–84.
- Ghasemaghaei, M., Ebrahimi, S., & Hassanein, K. (2018). Data analytics competency for improving firm decision making performance. *The Journal of Strategic Information Systems*, 27(1), 101–113.
- Ghasemaghaei, M., Hassanein, K., & Benbasat, I. (2019). Assessing the Design Choices for Online Recommendation Agents for Older Adults: Older Does Not Always Mean Simpler Information Technology. *MIS Quarterly*, 43(1), 329–346.
- Ghasemaghaei, M., Hassanein, K., & Turel, O. (2017). Increasing firm agility through the use of data analytics: The role of fit. *Decision Support Systems*, 101, 95–105.
- Grover, V., Chiang, R. H., Liang, T.-P., & Zhang, D. (2018). Creating Strategic Business Value from Big Data Analytics: A Research Framework. *Journal of Management Information Systems*, 35(2), 388–423.
- Grover, V., Gokhale, R., & Narayanswamy, R. S. (2009). Resource-based framework for IS research: Knowledge firms and sustainability in knowledge markets. *Journal of the Association for Information Systems*, 10(4), 306–332.
- Haas, M. R., & Park, S. (2010). To share or not to share? Professional norms, reference groups, and information withholding among life scientists. *Organization Science*, 21(4), 873–891.
- Haldin-Herrgard, T. (2000). Difficulties in diffusion of tacit knowledge in organizations. *Journal of Intellectual Capital*, 1(4), 357–365.
- Hass, K. B. (2005). The business analyst: The pivotal IT role of the future. *White Paper*.
- Herscovis, M. S., Turner, N., Barling, J., Arnold, K. A., Dupré, K. E., Inness, M., LeBlanc, M. M., & Sivanathan, N. (2007). Predicting workplace aggression: A meta-analysis. *Journal of Applied Psychology*, 92(1), 228–238.

- Hew, K. F., & Hara, N. (2007). Knowledge sharing in online environments: A qualitative case study. *Journal of the American Society for Information Science and Technology*, 58(14), 2310–2324.
- Huo, W., Cai, Z., Luo, J., Men, C., & Jia, R. (2016). Antecedents and intervention mechanisms: A multi-level study of R&D team's knowledge hiding behavior. *Journal of Knowledge Management*, 20(5), 880–897.
- Intezari, A., & Gressel, S. (2017). Information and reformation in KM systems: Big data and strategic decision-making. *Journal of Knowledge Management*, 21(1), 71–91.
- Ipe, M. (2003). Knowledge sharing in organizations: A conceptual framework. *Human Resource Development Review*, 2(4), 337–359.
- Issac, A. C., & Baral, R. (2018). Dissecting knowledge hiding: A note on what it is and what it is not. *Human Resource Management International Digest*, 26(7), 20–24.
- Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. (2014). Big data and its technical challenges. *Communications of the ACM*, 57(7), 86–94.
- Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, 70, 338–345.
- Johnson, J. S., Friend, S. B., & Lee, H. S. (2017). Big data facilitation, utilization, and monetization: Exploring the 3Vs in a new product development process. *Journal of Product Innovation Management*, 34(5), 640–658.
- Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. *Information Systems Research*, 15(2), 194–210.
- Joshi, A. W. (2016). When does customer orientation hinder (help) radical product innovation? The role of organizational rewards. *Journal of Product Innovation Management*, 33(4), 435–454.
- Kang, S.-W. (2016). Knowledge withholding: Psychological hindrance to the innovation diffusion within an organisation. *Knowledge Management Research & Practice*, 14(1), 144–149.
- Kearns, G. S., & Lederer, A. L. (2003). A resource-based view of strategic IT alignment: How knowledge sharing creates competitive advantage. *Decision Sciences*, 34(1), 1–29.
- Kelloway, E. K., & Barling, J. (2000). Knowledge work as organizational behavior. *International Journal of Management Reviews*, 2(3), 287–304.
- Khan, Z., & Vorley, T. (2017). Big data text analytics: An enabler of knowledge management. *Journal of Knowledge Management*, 21(1), 18–34.
- Kitchens, B., Dobolyi, D., Li, J., & Abbasi, A. (2018). Advanced Customer Analytics: Strategic value through integration of relationship-oriented big data. *Journal Of Management Information Systems*, 35(2), 540–574.
- Ko, D.-G., Kirsch, L. J., & King, W. R. (2005). Antecedents of knowledge transfer from consultants to clients in enterprise system implementations. *MIS Quarterly*, 29(1), 59–85.

- Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, 34(3), 387–394.
- Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology*, 86(1), 114–121.
- Liu, S., Duffy, A. H., Whitfield, R. I., & Boyle, I. M. (2010). Integration of decision support systems to improve decision support performance. *Knowledge and Information Systems*, 22(3), 261–286.
- Liu, S., Duffy, A. H., Whitfield, R. I., Boyle, I. M., & McKenna, I. (2009). Towards the realization of an integrated decision support environment for organizational decision making. *International Journal of Decision Support System Technology (IJDSST)*, 1(4), 38–58.
- Ma, M., & Agarwal, R. (2007). Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities. *Information Systems Research*, 18(1), 42–67.
- Majchrzak, A., & Malhotra, A. (2016). Effect of knowledge-sharing trajectories on innovative outcomes in temporary online crowds. *Information Systems Research*, 27(4), 685–703.
- Marques, F. M. F. R., La Falce, J. L., Marques, J. M. R., & De Muylder, C. F. (2019). The relationship between stress and maturity in knowledge management. *International Journal of Organizational Analysis*.
- Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Information technology and organizational performance: An integrative model of IT business value. *MIS Quarterly*, 28(2), 283–322.
- Miranda, S. M., & Saunders, C. S. (2003). The social construction of meaning: An alternative perspective on information sharing. *Information Systems Research*, 14(1), 87–106.
- Müller, O., Fay, M., & vom Brocke, J. (2018). The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics. *Journal of Management Information Systems*, 35(2), 488–509.
- Müller, O., Junglas, I., vom Brocke, J., & Debortoli, S. (2016). Utilizing big data analytics for information systems research: Challenges, promises and guidelines. *European Journal of Information Systems*, 25(4), 289–302.
- Nwankpa, J. K., & Datta, P. (2017). Balancing exploration and exploitation of IT resources: The influence of Digital Business Intensity on perceived organizational performance. *European Journal of Information Systems*, 26(5), 469–488.
- Opresnik, D., & Taisch, M. (2015). The value of big data in servitization. *International Journal of Production Economics*, 165, 174–184.
- Oviatt, S. (2006). Human-centered design meets cognitive load theory: Designing interfaces that help people think. *Proceedings of the 14th Annual ACM International Conference on Multimedia*, 871–880. <http://dl.acm.org/citation.cfm?id=1180831>
- Pauleen, D. J. (2017). Davenport and Prusak on KM and big data/analytics: Interview with David J. Pauleen. *Journal of Knowledge Management*, 21(1), 7–11.

- Pavlou, P. A., & Fygenson, M. (2006). Understanding and predicting electronic commerce adoption: An extension of the theory of planned behavior. *MIS Quarterly*, 30(1), 115–143.
- Peng, H. (2013). Why and when do people hide knowledge? *Journal of Knowledge Management*, 17(3), 398–415.
- Pierce, J. L., Kostova, T., & Dirks, K. T. (2001). Toward a theory of psychological ownership in organizations. *Academy of Management Review*, 26(2), 298–310.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879.
- PWC, PricewaterhouseCoopers. (2018). *What's next for the data science and analytics job market?* PwC. Retrieved May 25, 2020, from <https://www.pwc.com/us/en/library/data-science-and-analytics.html>
- Reid, S. E., Roberts, D., & Moore, K. (2015). Technology vision for radical innovation and its impact on early success. *Journal of Product Innovation Management*, 32(4), 593–609.
- Ringle, C. M., Wende, S., & Will, A. (2005). SmartPLS (Release 2.0 M3) [http://www. Smartpls. De](http://www.smartpls.de). University of Hamburg. *Hamburg: Germany*.
- Rosemann, M., & Vessey, I. (2008). Toward improving the relevance of information systems research to practice: The role of applicability checks. *Mis Quarterly*, 1–22.
- Ryan, S. D., Windsor, J. C., Ibragimova, B., & Prybutok, V. R. (2010). Organizational Practices That Foster Knowledge Sharing: Validation across Distinct National Cultures. *Informing Science*, 13, 26.
- Schleimer, S. C., & Faems, D. (2016). Connecting interfirm and intrafirm collaboration in NPD projects: Does innovation context matter? *Journal of Product Innovation Management*, 33(2), 154–165.
- 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.
- Sedighi, M., van Splunter, S., Brazier, F., van Beers, C., & Lukosch, S. (2016). Exploration of multi-layered knowledge sharing participation: The roles of perceived benefits and costs. *Journal of Knowledge Management*, 20(6), 1247–1267.
- Serenko, A., & Bontis, N. (2016). Understanding counterproductive knowledge behavior: Antecedents and consequences of intra-organizational knowledge hiding. *Journal of Knowledge Management*, 20(6), 1199–1224.
- Shafer, T. (2018). *The 42 V's of Big Data and Data Science*. Retrieved June 4, 2018, from <https://www.kdnuggets.com/2017/04/42-vs-big-data-data-science.html>
- 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.
- Sher, P. J., & Lee, V. C. (2004). Information technology as a facilitator for enhancing dynamic capabilities through knowledge management. *Information & Management*, 41(8), 933–945.

- Shin, B., & Kim, G. (2011). Investigating the reliability of second-order formative measurement in information systems research. *European Journal of Information Systems*, 20(5), 608–623.
- Shollo, A., & Galliers, R. D. (2016). Towards an understanding of the role of business intelligence systems in organisational knowing. *Information Systems Journal*, 26(4), 339–367.
- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263–286.
- Škerlavaj, M., Connelly, C. E., Cerne, M., & Dysvik, A. (2018). Tell me if you can: Time pressure, prosocial motivation, perspective taking, and knowledge hiding. *Journal of Knowledge Management*, 22(7), 1489–1509.
- Sonteya, T., & Seymour, L. F. (2012). Towards an understanding of the business process analyst: An analysis of competencies. *Journal of Information Technology Education: Research*, 11(1), 43–63.
- Srivastava, A., Bartol, K. M., & Locke, E. A. (2006). Empowering leadership in management teams: Effects on knowledge sharing, efficacy, and performance. *Academy of Management Journal*, 49(6), 1239–1251.
- Sun, H. (2012). Understanding user revisions when using information system features: Adaptive system use and triggers. *MIS Quarterly*, 36(2), 453–478.
- Tarafdar, M., D'Arcy, J., Turel, O., & Gupta, A. (2015). The dark side of information technology. *MIT Sloan Management Review*, 56(2), 61–70.
- Tarafdar, M., Gupta, A., & Turel, O. (2013). The dark side of information technology use. *Information Systems Journal*, 23(3), 269–275.
- Tay, R. (2017). Correlation, variance inflation and multicollinearity in regression model. *Journal of the Eastern Asia Society for Transportation Studies*, 12, 2006–2015.
- Totterdell, P., Spelten, E., Smith, L., Barton, J., & Folkard, S. (1995). Recovery from work shifts: How long does it take? *Journal of Applied Psychology*, 80(1), 43–57.
- Van Dyne, L., & LePine, J. A. (1998). Helping and voice extra-role behaviors: Evidence of construct and predictive validity. *Academy of Management Journal*, 41(1), 108–119.
- Van Knippenberg, D., Dahlander, L., Haas, M. R., & George, G. (2015). *Information, attention, and decision making*. Academy of Management Briarcliff Manor, NY.
- Vandyne, L., Cummings, L. L., & Parks, J. M. (1995). Extra-role behaviors-in pursuit of construct and definitional clarity (a bridge over muddied waters). *Research in Organizational Behavior: AN Annual Series of Analytics Essays and Critical Review*, 17, 215–285.
- Voelpel, S. C., Dous, M., & Davenport, T. H. (2005). Five steps to creating a global knowledge-sharing system: Siemens' ShareNet. *Academy of Management Perspectives*, 19(2), 9–23.
- Wamba, S. F., 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.

- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70, 356–365.
- Wang, B., Kang, Y., Childerhouse, P., & Huo, B. (2018). Service supply chain integration: The role of interpersonal relationships. *Industrial Management & Data Systems*, 118(4), 828–849.
- Wang, S., & Noe, R. A. (2010). Knowledge sharing: A review and directions for future research. *Human Resource Management Review*, 20(2), 115–131.
- Wang, Y., Han, M. S., Xiang, D., & Hampson, D. P. (2019). The double-edged effects of perceived knowledge hiding: Empirical evidence from the sales context. *Journal of Knowledge Management*, 32(2), 279–296.
- Waters, R. (2017, November 29). *How machine learning creates new professions—And problems*. <https://www.ft.com/content/49e81ebe-cbc3-11e7-8536-d321d0d897a3>
- Webb, R. (2018). *12 Challenges of Data Analytics and How to Fix Them*. Retrieved May 22, 2020, from <https://www.clearrisk.com/risk-management-blog/challenges-of-data-analytics>
- Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes*, 37(1), 60–82.
- Xiao, M., & Cooke, F. L. (2019). Why and when knowledge hiding in the workplace is harmful: A review of the literature and directions for future research in the Chinese context. *Asia Pacific Journal of Human Resources*, 57(4), 470–502.
- Yan, Z., Wang, T., Chen, Y., & Zhang, H. (2016). Knowledge sharing in online health communities: A social exchange theory perspective. *Information & Management*, 53(5), 643–653.
- Zhang, M., Sarker, S., & Sarker, S. (2008). Unpacking the effect of IT capability on the performance of export-focused SMEs: A report from China. *Information Systems Journal*, 18(4), 357–380. <https://doi.org/10.1111/j.1365-2575.2008.00303.x>
- Zhang, T., Wang, W. Y. C., & Pauleen, D. J. (2017). Big data investments in knowledge and non-knowledge intensive firms: What the market tells us. *Journal of Knowledge Management*, 21(3), 623–639.
- Zhao, H., Xia, Q., He, P., Sheard, G., & Wan, P. (2016). Workplace ostracism and knowledge hiding in service organizations. *International Journal of Hospitality Management*, 59, 84–94.

Appendix A

----- insert Table A1 about here -----

Tables

Table 1. Construct definitions

	Construct	Definition
--	-----------	------------

Knowledge Hiding	Evasive hiding	Intentionally delaying knowledge delivery to the requester until the information is not useful anymore and then promising to help next time (Connelly et al., 2012).
	Playing dumb	Pretending to know nothing about the requested knowledge (Connelly et al., 2012).
	Rationalized hiding	When the individual provides a justification for why the knowledge is not forthcoming (Connelly et al., 2012).
Big Data Characteristics	Variety	Different types of data (Ghasemaghaei et al., 2018).
	Volume	Increased amount of the data (Wamba et al., 2015).
	Velocity	Speed of processing and analyzing data (Ghasemaghaei, et al., 2018).
Decision-Making Quality	Decision outcomes that are high in terms of accuracy, precision, and reliability (Ghasemaghaei et al., 2018).	

Table 2. Correlation matrix and descriptive statistics

	Cron	Comp	Evas	Play	Ration	DQ	Var	Vel	Vol
Evasive hiding	0.96	0.96	0.94						
Playing dumb	0.96	0.97	0.78	0.95					
Rationalized hiding	0.94	0.96	0.70	0.75	0.92				
Decision quality	0.95	0.96	0.07	0.16	0.29	0.89			
Data Variety	0.93	0.96	0.04	0.09	0.3	0.48	0.94		
Data Velocity	0.94	0.95	0.23	0.32	0.43	0.63	0.66	0.92	
Data Volume	0.95	0.97	0.19	0.24	0.42	0.49	0.68	0.59	0.93
<i>Mean</i>	----	----	4.04	3.54	3.83	5.20	5.23	4.96	4.90
<i>Standard Deviation</i>	----	----	1.92	1.89	1.89	1.29	1.41	1.37	1.54

Note: Cron: Cronbach's alphas; Comp: composite reliability; Evas: evasive hiding; Play: playing dumb; Ration: rationalized hiding; DQ: decision quality; Vel: data velocity; Var: data variety; Vol: data volume

Table 3. Loadings and cross-loadings

	Var	Vel	Vol	DQ	Evas	Play	Ration
Data variety1	0.94	0.6	0.62	0.44	0.04	0.09	0.27
Data variety2	0.96	0.66	0.69	0.47	0.05	0.08	0.3
Data variety3	0.92	0.59	0.67	0.44	0.03	0.08	0.27
Data velocity1	0.57	0.89	0.55	0.56	0.23	0.29	0.39
Data velocity2	0.62	0.92	0.52	0.56	0.28	0.38	0.46
Data velocity3	0.62	0.93	0.54	0.62	0.14	0.21	0.34
Data velocity4	0.61	0.93	0.53	0.59	0.17	0.23	0.36
Data vulme1	0.64	0.55	0.96	0.44	0.19	0.23	0.41

Data vulme2	0.69	0.56	0.93	0.47	0.17	0.21	0.4
Data vulme3	0.67	0.55	0.91	0.47	0.1	0.15	0.32
Data vulme4	0.67	0.55	0.93	0.45	0.22	0.28	0.42
De cision-Making Quality1	0.41	0.55	0.4	0.91	0.08	0.16	0.26
De cision-Making Quality2	0.42	0.54	0.46	0.9	0.07	0.14	0.26
De cision-Making Quality3	0.45	0.57	0.49	0.91	-0.02	0.05	0.19
De cision-Making Quality4	0.46	0.62	0.48	0.89	0.15	0.23	0.34
De cision-Making Quality5	0.45	0.58	0.42	0.86	0.06	0.14	0.26
De cision-Making Quality6	0.37	0.51	0.34	0.87	0.03	0.12	0.22
Evasive hiding1	0.05	0.23	0.17	0.07	0.94	0.79	0.73
Evasive hiding2	0.03	0.22	0.17	0.03	0.97	0.79	0.71
Evasive hiding3	0.04	0.18	0.2	0.08	0.95	0.76	0.73
Evasive hiding4	0.04	0.24	0.16	0.09	0.92	0.73	0.7
Playing dumb1	0.1	0.3	0.24	0.15	0.75	0.94	0.76
Playing dumb2	0.13	0.33	0.25	0.18	0.81	0.96	0.79
Playing dumb3	0.07	0.31	0.22	0.16	0.82	0.97	0.78
Playing dumb4	0.03	0.25	0.19	0.12	0.74	0.93	0.75
Rationalized hiding1	0.28	0.47	0.4	0.28	0.74	0.78	0.91
Rationalized hiding2	0.32	0.43	0.44	0.3	0.69	0.7	0.94
Rationalized hiding 3	0.28	0.35	0.37	0.25	0.68	0.73	0.94
Rationalized hiding4	0.21	0.31	0.31	0.22	0.67	0.78	0.89

Note: Vel: data velocity; Var: data variety; Vol: data volume; Evas: evasive hiding; Play: playing dumb; Ration: rationalized hiding; DQ: decision quality

Table 4. Sample applicability check responses

Is it reasonable to expect that when firms process large amounts of data quickly in almost real time, data analysts will be more motivated to hide knowledge stemming from such data?⁵
<i>“When large amounts of data are derived automatically, employees tend to be ignorant about sharing this knowledge. There have been instances where data that could’ve helped us solve major business problems were not brought to our attention, as other employees felt the data generated could be irrelevant. They were not willing to share the data, thinking that their job could be in danger.”</i>

⁵ In this question, we asked a general question of the positive impact of knowledge hiding on decision quality because 1) we wanted to test whether the results we obtained for each of our hypotheses were reasonable through the eyes of data analysts. H1 specifically focuses on the impact of knowledge hiding (in general) on decision quality and not the impact of each type of knowledge hiding on decision quality, and 2) data analysts understand the meaning of knowledge hiding in general. However, they may not understand the meaning of each type of knowledge hiding very well, which could bias their responses.

<p><i>“Yes, from my experience I would say that knowledge hiding is prevalent and it does happen with large amounts of data available. During my work experience I would have access to data from all the departments and often would get requests from other departments to analyze data for them. I would play dumb sometimes or sometimes ask them to take my manager's approval for me to work on their data because my performance and metrics were getting affected.”</i></p>
<p><i>“When I was working at company X, part of my job was to extract a huge amount of data in a relevant format as requested by the Business team and share it with them on a daily basis. As a result, there were times when I would avoid sharing extra knowledge from the data on the basis of my experience to avoid the extra workload. I only replied and shared what they specifically asked from me.”</i></p>
<p>Is it reasonable to expect that when firms process different types of data, the data analysts would be less motivated to hide knowledge stemming from such data?</p>
<p><i>“Yes, in my work, part of my job is to extract a huge amount of data in a relevant format as requested by the Business team and share it with them on a daily basis. But when data is varied, I have to ask the Business team about exactly what they want and share my own experience about what exactly they might need. I cannot hide data this way.”</i></p>
<p><i>“Yes, processing different types of data ensures classification or characterization of data by need and hence motivates employees to share knowledge in their own datasets.”</i></p>
<p><i>“Yes, amongst the different types some data will seem more relevant than others, which employees may be more willing to share with others.”</i></p>
<p>One would expect that hiding knowledge would negatively affect decision-making quality. However, our findings show the opposite: that knowledge hiding increases firm decision-making quality. Is this finding reasonable?</p>
<p><i>“Yes, if knowledge is hidden to some level, the employee will be forced to get and process their own data. During the process, their individual ideas and insights will be generated. The number of insights increases and as a result, the quality of decision making will increase.”</i></p>
<p><i>“This is True. More data would not mean more accurate decision making. In most cases, it just adds to more analysis with very little to gain in return and at times even affecting negatively.”</i></p>
<p><i>“Yeah, data hiding increases firm decision-making quality because in case of large data or different types of data, employees will transfer only those data which are necessary for decision making and this filtration helps managers to make decisions effectively and efficiently.”</i></p>

Table A1. Measurement items of the variables

Construct Names & Resource	Measurement Items (7-point Likert scales ranging from “strongly disagree” to “strongly agree”)
<p>Knowledge Hiding. Connelly et al. (2012)</p>	<p>In my firm, when asked by a colleague to share data analytics-driven insights:</p> <p>Evasive Hiding</p> <ul style="list-style-type: none"> • People agree to help him/her but never really intend to. • People agree to help him/her but instead give him/her shallow information (excluding important insights). • People tell him/her that they would help him/her out later but stall as much as possible.

	<ul style="list-style-type: none"> • People offer him/her some other information instead of what he/she really wants. <p>Playing Dumb</p> <ul style="list-style-type: none"> • People pretend that they do not know the information. • People say that they do not know, even though they do. • People pretend they do not know what she/he was talking about. • People say that they are not very knowledgeable about the topic. <p>Rationalized Hiding</p> <ul style="list-style-type: none"> • People explain that they would like to tell him/her but are not supposed to. • People explain that the information is only available to people on a particular project. • People tell him/her that their boss would not let anyone share this knowledge. • People say that they would not answer his/her questions because he or she is not part of the project team.
<p>Big Data Characteristics Ghasemaghahi & Calic (2019)</p>	<p>Data Volume</p> <ul style="list-style-type: none"> • In my firm, people analyze large amounts of data. • In my firm, the quantity of data people explore is substantial. • In my firm, people use a great deal of data. • In my firm, people scrutinize copious volumes of data. <p>Data Variety</p> <ul style="list-style-type: none"> • In my firm, people use several different sources of data to gain insights. • In my firm, people analyze many types of data. • In my firm, people examine data from a multitude of sources. <p>Data Velocity</p> <ul style="list-style-type: none"> • In my firm, people analyze data as soon as they receive it. • In my firm, the time period between when people get new data and when they analyze it is short. • In my firm, people are fast in exploring data. • In my firm, people analyze data speedily.
<p>Decision Quality Ghasemaghahi et al. (2018)</p>	<p>In my firm, decisions we make are often:</p> <ul style="list-style-type: none"> • Accurate • Correct • Precise • Flawless • Error-free • Reliable

Figures

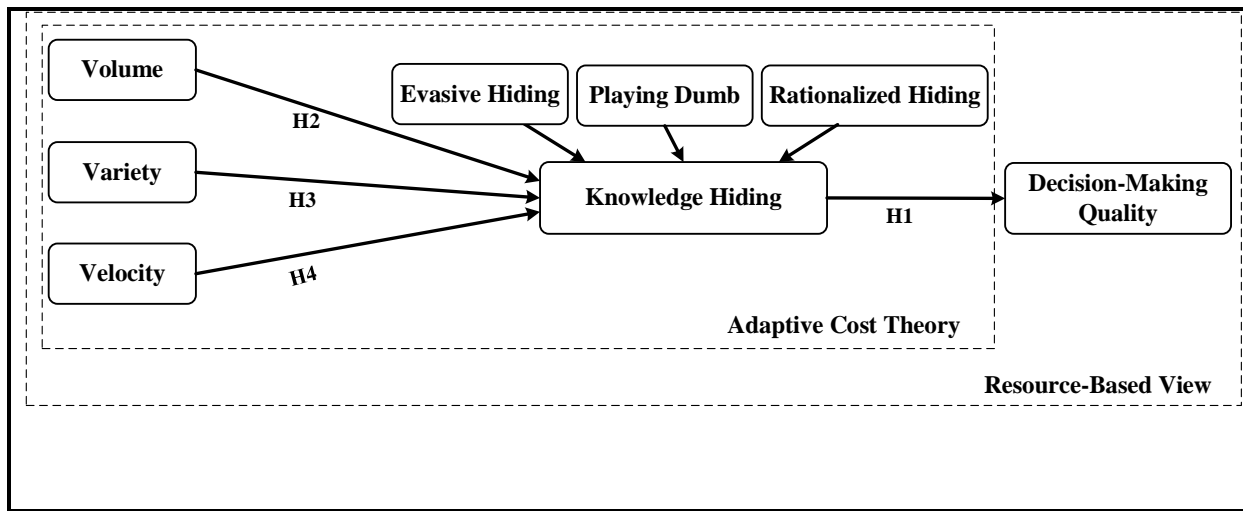


Figure 1. Research model

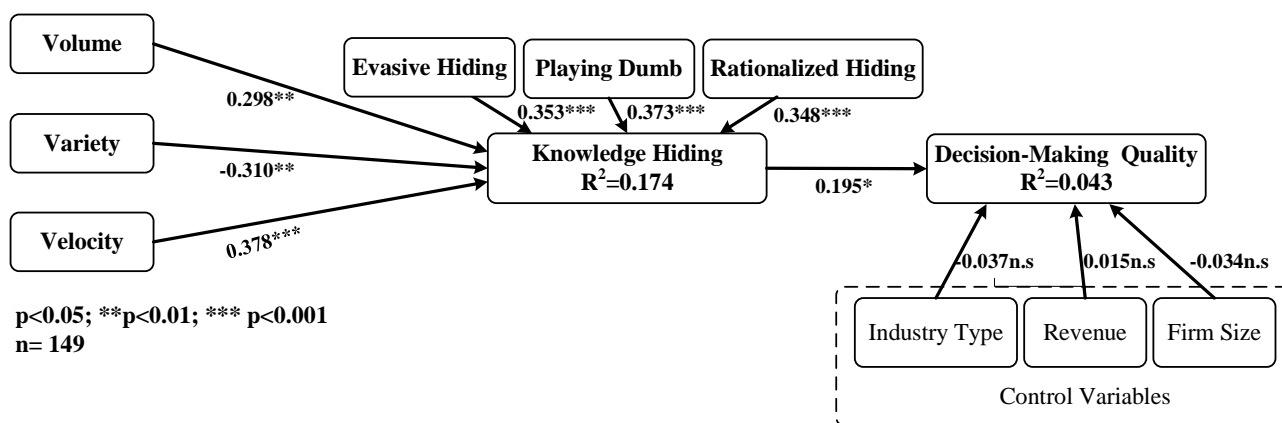


Figure 2. Structural model

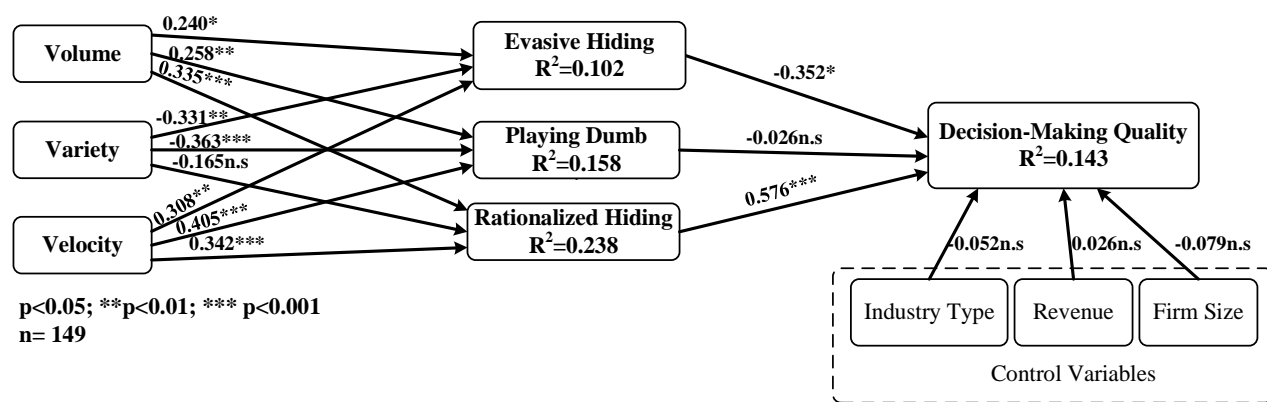


Figure 3. Structural model (considering all constructs as first-order variables)

**Possible Negative Effects of Big Data on Decision Quality in Firms: The Role
of Knowledge Hiding Behaviors**

Dr. Maryam Ghasemaghæi

Email: ghasemm@mcmaster.ca

DeGroote School of Business
McMaster University
Hamilton, Ontario, Canada
(905) 525-9140

Dr. Ofir Turel

Email: oturel@fullerton.edu

Professor of Information Systems and Decision Sciences
California State University, Fullerton
(657) 278-5613

Corresponding author:

Dr. Maryam Ghasemaghæi

Phone: (905) 525-9140, ext 21721

Fax: (905) 521-8632

Acknowledgment

This work was supported by the Social Sciences and Humanities Research Council of Canada [grant number 20006922].