UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

INSTITUTO COPPEAD DE ADMINISTRAÇÃO

GUILHERME CALDERUCIO DUQUE ESTRADA

Can data collection technologies increase information processing capacity and help to reduce transaction costs?

Rio de Janeiro

2020

GUILHERME CALDERUCIO DUQUE ESTRADA

Can data collection technologies increase information processing capacity and help to reduce transaction costs?

Master's dissertation presented to the COPPEAD Graduate School of Business, Universidade Federal do Rio de Janeiro, as part of the mandatory requirements in order to obtain the title of Master in Business Administration (M.Sc.).

Supervisor: Prof. Leonardo Marques, Ph.D.

Rio de Janeiro

CIP - Catalogação na Publicação

C146c

Calderucio Duque Estrada, Guilherme
Can data collection technologies increase information processing capacity and help to reduce transaction costs? / Guilherme Calderucio Duque Estrada. -- Rio de Janeiro, 2020.
61 f.

Orientador: Leonardo Marques Gomes.
Dissertação (mestrado) - Universidade Federal do Rio de Janeiro, Instituto COPPEAD de Administração, Programa de Pós-Graduação em Administração, 2020.

1. supply chain management. 2. transaction cost theory. 3. information processing theory. 4. data collection technologies. 5. exploratory factor analysis. I. Marques Gomes, Leonardo, orient. II. Título.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a), sob a responsabilidade de Miguel Romeu Amorim Neto - CRB-7/6283.

GUILHERME CALDERUCIO DUQUE ESTRADA

Can data collection technologies increase information processing capacity and help to reduce transaction costs?

Master's dissertation presented to the COPPEAD Graduate School of Business, Universidade Federal do Rio de Janeiro, as part of the mandatory requirements in order to obtain the title of Master in Business Administration (M.Sc.).

Approved on February 18th, 2020

Leonardo Marques Gomes, Ph.D - Advisor
COPPEAD

Peter Fernandes Wanke, D.Sc
COPPEAD

p/

Susana Carla Farias Pereira, D.Sc
FGV-SP

Rio de Janeiro

2020

ACKNOWLEDGEMENTS

I'd like to thank God for giving me physical and mental health, financial conditions and all the other requisites needed for the conclusion of this Master that transformed my career and allowed me to feel happy again at work and have a good perspective for my future.

I'd also like to thank my girlfriend Carol for all the support and patience in the often occasions when I couldn't see her because of my duties in the Master and for the happy moments we had in my spare time that allowed me to recover my energies for one more day in the Master.

I must also thank my parents who were the first that made me believe I should enroll in the Master, helped me to study during the first 27 years of my life and have all the skills I needed to have to pursue my studies and also gave me financial and emotional support in addition to showing me the right values that will allow me to be not only a competent professional but also an ethical person.

I should also thank my advisor Leo and the whole Coppead faculty for investing so much energy in this work as well as during the Master classes.

I'd also like to thank E-Create for betting in my potential during the Summer Job and giving me autonomy to put in practice many of the concepts I learnt in the Master.

Finally, I should thank Cici, Luiz Felipe, Dário, Hugo, Tuninho, Marmota, Natario and all my friends and Master colleagues that made my life lighter and funnier during the last 2 years.

ABSTRACT

ESTRADA, Guilherme. **Data collection technologies for supply chain management:** a dual perspective of transaction costs and information processing theories. Rio de Janeiro, 2020. 63pp. Dissertation (Master's Degree in Business Administration) - COPPEAD Graduate School of Business, Federal University of Rio de Janeiro, Rio de Janeiro, 2020.

Uncertainty has been in the center of both transaction cost theory and information processing theory. Although the term uncertainty holds some differences in these two contexts, in both lenses the term is related to external dependence and a resulting need for adaptation. Data collection technologies have been previously identified as possible ways of increasing information processing capacity and reducing uncertainties in supply chain management. Studies have shown a direct link between the use of data collection technologies for supply chain management and the reduction of transaction costs. In this study, besides mapping adoption of data collection technologies, other related mechanisms of increasing information processing capacity were also explored: the number of tiers involved in data collection, the frequency of data collection, the availability of human resources skills to collect data and the use of supply chain collaboration. Supply chain professionals answered a survey about these five constructs and their effects on transaction costs. Results were analyzed through exploratory factor analysis and linear regression models. A positive correlation with the reduction of transaction costs was found for: the use of QR Code, shared systems and cookies, frequency of data collection, the availability of managers' skills to collect supply chain data and one supply chain collaboration practice (meetings between members of different supply chain participants). This study might help to fill gaps in the existing literature on information processing and transaction cost theories and drive managerial decisions on projects aiming the reduction of transaction costs through the use data collection technologies.

Keywords: supply chain management, transaction cost theory, information processing theory, data collection technologies, exploratory factor analysis, technology, supply chain tiers, frequency of data collection, skills for data collection, supply chain collaboration, monitoring costs, risks of opportunism.

LIST OF ILUSTRATIONS

LIST OF TABLES

LIST OF ABREVIATIONS

EDI	Electronic Data Interchange
IoT	Internet of Things
IP	Information Processing
IPT	Information Processing Theory
KMO	Kaiser-Meyer-Olkin
MSA	Measure of Sampling Adequacy
QR	Quick Response
RFID	Radio-frequency Identification
SCM	Supply Chain Management
TCE	Transaction Cost Economics
VMI	Vendor Managed Inventory

CONTENTS

1. INTRODUCTION

1.1. CONTEXT

Supply chain management (SCM) is defined as how firms utilize suppliers' processes, capabilities and technology to enhance its competitive advantage, manufacturing, logistics and materials management (Tan, 2001). The role played by technology in the field is to enable firms to share information throughout the supply chain (Kearns & Lederer, 2003).

Significant growth in information technology capabilities observed since the 1990s has changed SCM (Mentzer et al., 2001). More recently, technologies such as the internet of things (IoT), big data, blockchain and additive manufacturing, have become relevant to SCM (Min, Zacharia, & Smith, 2019).

Special regard has been given on how the use of data collection technologies in SCM such as blockchain (Tokar & Swink, 2019) and RFID (Lee & Lee, 2010) help better manage transactions and might reduce transaction costs.

Information Processing Theory (IPT) posits that uncertainty comes as a result of information processing (IP) misfits (Tushman & Nadler, 1978). And uncertainty is one of the elements identified as a source of transaction costs (McIvor, 2009).

The use of different data collection technologies (Ngai, Cheung, Lam, & Ng, 2014; Qrunfleh & Tarafdar, 2014; Srinivasan & Swink, 2018), number of tiers involved in data collection (Monczka, Petersen, Handfield, & Ragatz, 1998; V. Roth, 1996), frequency of data collection (Zhou & Benton Jr, 2007), human resources skills employed on collecting supply chain data and analyzing it (Waller & Fawcett, 2013) and supply chain collaboration (Flynn, Koufteros, & Lu, 2016) have also been identified as important aspects that can help to reduce IP misfits.

1.2. THEORETICAL BACKGROUND & RESEARCH QUESTION

The theoretical background was first built upon a literature review relating SCM and data collection technologies. Transaction cost economics was identified in the literature as a theory that could link all these concepts, a deeper literature review about this theory was conducted. Then, IPT emerged as the theory that would allow this study to explain the relationship between transaction costs reduction and data collection technologies. Finally, the conceptual model was presented relating the key aspects of this research through the hypotheses that will be tested.

Some studies have already investigated the impact of technology in reducing IP misfit. Others have explored the impact of specific individual data collection technologies in the reduction of transaction costs. However, in the literature review, no study relating data collection technologies as a group of technologies capable of reducing IP misfit and their impact on transaction costs reduction was found. This research aims to fill this theoretical gap. So, the research question is:

Can data collection technologies that decrease IP misfits contribute to reduce transaction costs?

1.3. RESEARCH METHOD

The research method chosen for this research was a survey. This survey tests all the hypotheses and sub-hypotheses raised in the conceptual model and control for variables related to company characteristics. The target of the survey is a group of subscribers to a Brazilian supply chain magazine. The results were analyzed through an exploratory factor analysis and linear regression models.

1.4. DISSERTATION STRUCTURE

In Chapter 2, a literature review about transaction cost economics, IPT, data collection technologies, transaction costs assessment and the conceptual model will be presented. In Chapter 3, the details about the survey sample, the questionnaire and operationalization of variables and the data analysis will be provided. In Chapter 4, the findings of this research will be presented. In Chapter 5, the conclusions about these findings will be discussed.

2. LITERATURE REVIEW

2.1. LITERATURE SEARCH METHOD

This chapter is structured as follows. First, an explanation of how the papers used in this thesis were selected will be given. Moreover, articles related to transaction costs found in the literature will be outlined. Then, papers that served as a reference for this study in terms of the IPT, IP fit and mechanisms that add IP capacity will be cited. Finally, the focus will move to applications and previous studies of data collection technologies in SCM, indicating a gap in the literature that will be fulfilled by this research, connecting the two former concepts.

The articles that served as a basis for this research were found in multiple databases: Wiley, Taylor and Francis, Scopus, Science Direct, Emerald and Web of Science. The keywords selected for the search through these databases were "('supply chain' OR 'transaction cost*' OR 'information processing theory') AND 'data collection technolog*'" and no filter was employed initially. Then, duplicated results and articles published in journals pertaining to a classification in the QUALIS System of CAPES lower than B2 were not considered. To guarantee the adherence of this literature review to the main focus of this thesis, an abstract review allowed also a selection of the papers that are related to the goal of this study. Other articles specifically related to transaction cost economics and the IPT were added to the study.

2.2. TRANSACTION COST ECONOMICS

2.2.1. DEFINITIONS OF TRANSACTION COSTS

According to Coase (1937), the growth of one firm is limited to its capacity of keeping the costs of carrying an additional transaction internally lower than doing it through another firm. This is explained by the difference between transaction costs and production costs.

Transaction costs are “the costs of running the economic system” according to Arrow (1969) as cited by Oliver E. Williamson (1985) and Transaction Cost Theory is based on four concepts that drive transaction costs: bounded rationality, opportunism, small numbers bargaining and information impactedness (McIvor, 2009). Besides those four elements, McIvor (2009) points other three causes for transaction costs: asset specificity, uncertainty and infrequency. These concepts are defined on Table 1.

Table 1 - Definitions of the main concepts related to transaction costs

Concept	Definition
Bounded rationality	Bounded rationality refers to the incapacity of humans to predict all possibilities while taking decisions (McIvor, 2009). As a consequence to this human condition, all kinds of contracts are incomplete and unanticipated disturbances end up arising (Oliver E. Williamson, 2008), increasing transaction costs of negotiating specifications and prices (Grover & Malhotra, 2003).
Opportunism	The risk of opportunism is caused by the tendency of human beings to act in their own self-interest (Grover & Malhotra, 2003). This provokes practices associated to “cheating, lying and subtle forms of violation of agreements”, rising costs linked to preventing this kind of behavior and protect assets (Grover & Malhotra, 2003).
Small numbers bargaining	Small numbers bargaining refers to the amount of alternatives available to the buyer in a transactional relationship. When faced with few alternatives, the buyer is vulnerable to worse conditions as the supplier has a more limited concurrence to beat in terms of number of players (McIvor, 2009).
Information impactedness	Information impactedness occurs when there is some information asymmetry between the parties involved in a transaction (McIvor, 2009). As explained by Oliver E Williamson (1983), when one of the parties retains a piece of information unknown by the other party, the organization that lacks that information will not take related decisions in a way that is as rational (although still limited by bounded rationality) as it could be otherwise.
Asset specificity	Asset specificity can be understood as the degree of customization embedded in a transaction (McIvor, 2009). In other words, if an investment has a high value in one specific transaction and the same investment would have little or no value at all in another one. Klein, Crawford, and Alchian (1978) define the concepts of quasi-rent that helps to explain asset specificity. They compare the difference of the value of an asset to its renter and the value the same asset would have in its “next best use” to another renter. This difference is called quasi-rent. The higher the quasi-rent of an investment is the higher the specificity of the associated asset. In this kind of relationship, the potential for opportunistic behavior is higher because organizations tend to avoid cooperation to prevent the other party to gain more bargaining power in an asset specific transaction.
Uncertainty	Oliver E. Williamson (2008) refers to uncertainty as the factor that produces disturbances, requiring adaptations from the organizations.
Infrequency	Infrequency is related to the number of times a transaction is performed. Infrequency in a transaction leads to higher setup costs and higher reputation costs. Setup costs are the ones that incur when the transaction is started and reputation costs are those required to create trust between the parties involved. More infrequent transactions are expected not to be worth the required coordination and control efforts (Maltz, 1994).

Data collection technologies, such as EDI, have been used in companies with the intent to reduce transaction costs (Johnson, Klassen, Leenders, & Awaysheh, 2007). Different data collection technologies might be useful to reduce different types of transaction costs. The two components of transaction costs will be presented on the next session.

2.2.2. COMPONENTS OF TRANSACTION COSTS

Clemons, Reddi, and Row (1993) split transaction costs in two different sources of costs: coordination costs and transaction risk.

Coordination costs are resulted by the need to exchange information with the other party and consider that piece of information into the decision process (Grover & Malhotra, 2003). They include costs of negotiating prices, establishing scope, specifications or sharing other kinds of information about products, demand or availability, and aligning design changes (Grover & Malhotra, 2003).

Transaction risks come in two different ways: the risk of opportunistic behavior and the risk of changes in transaction circumstances (Wever, Wognum, Trienekens, & Omta, 2012). As explained in the topic 2.2.1 of this thesis, opportunistic behavior is highly affected by asset specificity and information impactedness while the risk of changes in transaction circumstances is mainly due to the level of uncertainty (Wever et al., 2012).

In an attempt to manage both coordination costs and transaction risks, companies have used different data collection technologies (Antonucci et al., 2019; Rosemberg, 2019). The next paragraph explains in further details how they are expected to bring this type of outcome in supply chains.

2.2.3. THE ROLE OF DATA COLLECTION TECHNOLOGIES TO REDUCE TRANSACTION COSTS

Different data collection technologies have been associated in previous studies to reductions in transaction costs. Each of them has particular characteristics that influence how they function and how they end up contributing to the reduction of transaction costs. However, as it will be explained in the next paragraphs of this session, data collection technologies often share some of these characteristics such as the reliability of the data collected, which reduces the risk of opportunism, and the reduction of efforts to guarantee high levels of supply chain coordination, which reduces coordination costs. As the transaction costs literature doesn't explore data collection technologies as a whole, the specific roles of some examples of them were selected to be presented in this study: RFID, blockchain and QR Code's roles in the reduction of transaction costs will be presented.

Blockchain applications can not only protect supply chain data, increasing trust and thus reducing transaction costs, but also reduce the effort to verify historical data about suppliers, reduce information asymmetry in supply chains and, reduce the effort to run real-time monitoring systems and reduce the risk of opportunism through smart contracts (Rosemberg, 2019).

Additional security provided by QR Code has been seen as a way to increase security and improve trust in suppliers and then reduce transaction costs (Yuen, Wang, Ma, & Wong, 2019). QR Code are also often used in conjunction with blockchain to ease the verification of historical data about suppliers, reducing the risk of opportunism (Antonucci et al., 2019).

RFID is another data collection technology that might provide real-time supply chain visibility. It allows companies to improve supply chain traceability (Deichmann, Goyal, & Mishra, 2016) by keeping records of products' physical information so it reduces the cost of real-time monitoring in supply chains (Wang, Hu, & Zhou, 2017).

Some studies have related transaction costs and some important concepts that are part of the IPT: IP capacity and IP misfit. For instance, one of the mechanisms (cross-functional integration) used to increase IP capacity brings the reduction of coordination costs as a consequence of reduced IP misfit (Swink & Schoenherr, 2015). In the next chapter, IPT will be presented along with mechanisms taken to achieve IP fit and a different perspective relating the two theories will be discussed.

2.3. INFORMATION PROCESSING THEORY (IPT)

IPT can be explained in three different parts using the reasoning established by Tushman and Nadler (1978). These parts are presented on Table 2.

Table 2 – IPT explained in three parts

Part	Explanation
Part 1: relationship between uncertainty and IP needs	Uncertainty is defined as "the difference between information possessed and information required to complete a task". As the tasks performed by different organizational subunits vary, so does the information needed and the degree of uncertainty for each of them.
Part 2: relationship between IP needs and IP capacity	As uncertainty increases, so does the need for processing more information and, as a result, there is an increasing need for IP capacity as well.
Part 3: relationship between IP capacity and effective IP	Effective IP includes the collection of the right data, the movement of that data at the right pace and its transmission without any losses in the quality of the information, as well as the ability to adequately use the amounts of information obtained. So IP capacity consists on the level of effectiveness of IP that an organizational unit or subunit possess.

On the definition of uncertainty proposed by Tushman and Nadler (1978) for the IPT and the one proposed by Oliver E. Williamson (2008) for Transaction Cost Economics, there are some differences. While the former is focused on the information gap in tasks performed by organizations, the latter is focused on production of disturbances as the emergent need for adaptations. However, there is also a link between the two definitions. On IPT, one of the sources of uncertainty is task environment, defined as the dependence on external actors. When focal external actors are suppliers, dependence is also a source of uncertainty for TCE. In both definitions, uncertainty demands adaptations, but different propositions are presented as adaptations - adequate governance structures for TCE and mechanisms to achieve IP fit for IPT. The next paragraphs explain these mechanisms in further detail.

IP mechanisms, the strategies adopted by companies to reduce IP misfits (the gaps between IP needs and IP capacity), may vary from company to company. Busse, Meinlschmidt, and Foerstl (2017) describe three different categories of IP mechanisms: reducing IP needs required to coordinate activities, increasing IP capacity or combining both strategies.

IP mechanisms' effectiveness vary according to the level of existing uncertainty (Foerstl, Meinlschmidt, & Busse, 2018). In environments where uncertainty is low, rules and programs, hierarchical referral and goal setting have proven to be effective according to Galbraith (1974) as cited by Foerstl et al. (2018). However, high uncertainty environments demand different IP mechanisms.

The IP mechanisms capable of reducing IP needs are creating slack resources and creating self-contained tasks. The former consists on decreasing the required level of performance (for instance, through longer delivery times and higher levels of buffer inventories) and the latter consists on spreading the geographical area where the organization acts (Foerstl et al., 2018).

IP mechanisms able to increase IP capacity are lateral relations and vertical integration systems. Creating lateral relations means decentralizing the source of information by conducting organizational redesign practices. Investing in vertical integration systems aims to acquire or use more information through the increase of IP capacity of either employees or technology (Foerstl et al., 2018). Alt (2017) also recognizes the impact that information technology, such as data collection technologies, has on providing more market information.

When faced with transaction costs associated with IP misfits, companies can adopt strategies to reduce IP needs or strategies to increase IP capacity. As the scope of this research is focused on data collection technologies, only IP capacity enhancement was investigated in the literature review. Through its capabilities of storing multiple types of information about transactions in a reliable manner (Foerstl, Schleper, & Henke, 2017), some authors suggest that data collection technologies such as blockchain can be used to dramatically increase an organization's IP capacity (Massimino, Gray, & Lan, 2018), reducing uncertainties (Tushman & Nadler, 1978) and thus transaction costs (McIvor, 2009).

2.4. CONCEPTUAL MODEL

This chapter will develop on the notion previously cited that transaction costs are the sum of two components: coordination costs and transaction risk. (Grover & Malhotra, 2003). On the following paragraphs, the literature on how these components have been assessed is reviewed and how they evolved until the measurement of transaction costs was broken down into four types of transaction costs by Grover and Malhotra (2003). These four types will serve as a basis for the assessment of transaction costs in this research.

One of the first authors who tried to measure transaction costs was Strassmann (1997). She considered selling and general administrative costs reported in financial statements a proxy for coordination costs. However, she did not consider any type of cost due to transaction risks.

Pilling, Crosby, and Jackson (1994) measured transaction costs through a broader perspective. They ran an experiment with groups of midlevel purchasing executives who had to perform tasks similar to the ones they did in their jobs and answer questions related to these tasks regarding three concepts defining transaction costs: the costs of preparing an exchange relationship, those related to monitor supplier performance and those aimed at dealing with opportunism. They were able to assess transaction costs through a wider lens, but their analysis lacked validation due to the size of the sample: 51 people.

Grover and Malhotra (2003) based their research partially on the work of Pilling et al. (1994) to build scale items and chose a survey method to add scale validation to the measure. They also chose to measure transaction costs using slightly different dimensions: effort, monitor, problem and advantage. Effort means the effort needed to develop the buyer-supplier relationship. Monitor refers to monitoring the performance of the focal supplier. Problem relates to managing the problems that arise in the relationship. Advantage is the likelihood of opportunism in the relationship. Grover and Malhotra's (2013) approach allowed them to assess both coordination costs and transaction risks but the scale items used still miss part of what constitutes these elements.

Handley and Benton (2013) have introduced a different angle to transaction costs. They divided them in control costs – "the time, effort and resources associated with achieving adequate levels of provider control" – and coordination costs. To measure control costs, they used two constructs: contract management and monitoring. These constructs were assessed through surveys using a sample of buyers and another sample of suppliers. They developed scale items to assess coordination costs and contract management and adapted the scale items developed by Grover and Malhotra (2003) to assess monitoring costs. Strassmann (1997), Pilling et al. (1994), Grover and Malhotra (2003) and Handley and Benton (2013) developed general assessments of transaction costs but they could not focus on specific drivers of transaction costs.

Other authors focused just on transaction risks. Kraljic (1983) analyzed transaction risks through availability, number of suppliers, competitive demand, make-or-buy opportunities, and storage risks and substitution possibilities. Hallikas, Virolainen, and Tuominen (2002) suggested a framework for risk analysis and assessment in supply networks that splits risks in demand related factors and value chain positioning, delivery performance ability, financial factors and pricing. Zsidisin (2003) shed light on the origin of risk and proposed risk assessments based on item characteristics, market characteristics and supplier characteristics. Manuj and Mentzer (2008) listed specific sources of transaction risk: disruption of supply, inventory, schedules and technology access; price escalation; quality issues; technology uncertainty; product complexity; frequency of material design changes. All these contributions will enable a more accurate assessment of transaction risk.

The references cited above constitute the basis of the constructs that will be used in the formation of the hypotheses – pictured in Figure 1 and further explained in the following paragraphs.

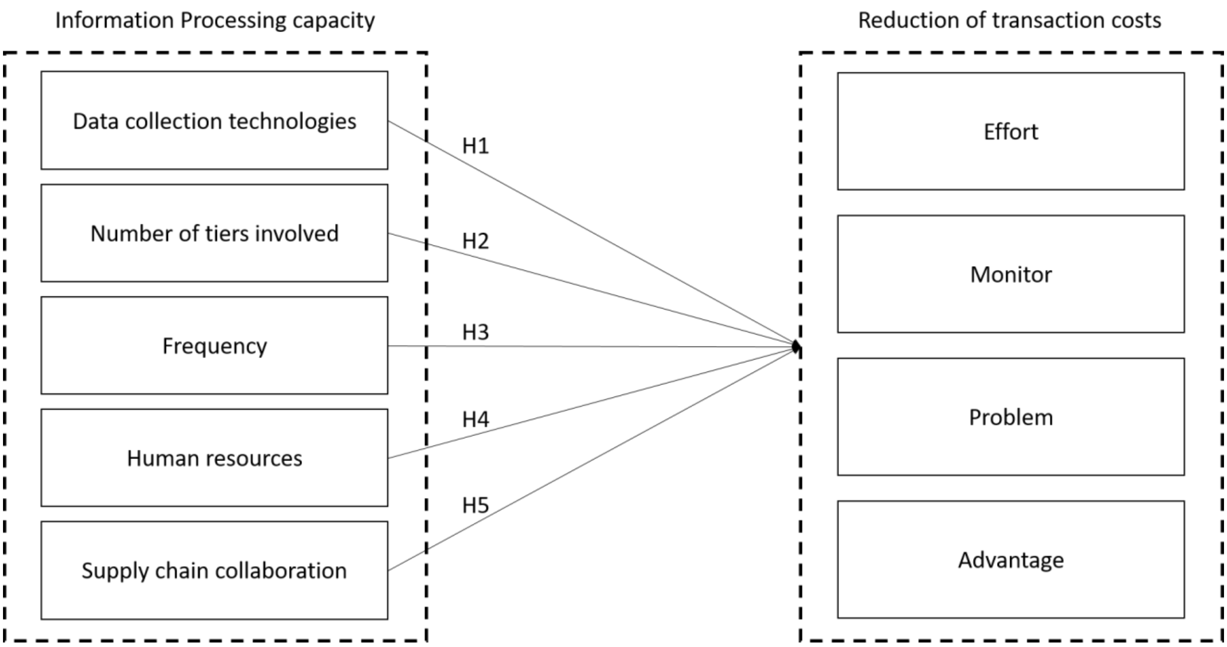


Figure 1 - Initial Conceptual Model

The first construct, Data collection technologies, is explored by several authors. Blockchain applications to verify data about suppliers, protect supply chain data, allow easier real-time monitoring and reduce opportunism risk through smart contracts are potential ways of reducing transaction costs (Rosemberg, 2019). RFID tracking applications' potential to generate relevant information in supply chains has been the focus of some studies (Ngai et al., 2014; Qrunfleh & Tarafdar, 2014; Wowak, Craighead, & Ketchen, 2016). Data generated in supply chains from unstructured data technology sources (e.g. camera images or forum discussions) have also been studied (Wamba, Gunasekaran, Dubey, & Ngai, 2018) as well as SCM systems like Electronic Data Interchange (EDI) systems or Vendor Managed Inventory (VMI) systems (Srinivasan & Swink, 2018).

Data collection technologies are especially associated with decreases in transaction costs as their power of creating trust among supply chain partners have been studied. Examples of how this occurs include: the use of blockchain to create trust among supply chain members and then reduce the risk of opportunistic behavior (Kumar, Liu, & Shan, 2019) and the use of e-business technologies (internet, intranets, extranets, and web-based applications) to enhance inter-organizational collaboration, which, in turn, help to create trust among supply chain partners and, as a result, help to reduce transaction costs (Sanders, 2007). The hypothesis related to this construct is:

H1: there is a positive correlation between the use of data collection technologies to monitor the supply chain and reduction in transaction costs.

The second construct of the conceptual model is number of tiers involved. The number of tiers involved in data collection is associated with the concept of visibility as the latter is defined as high quality information sharing through supply linkages (Barratt & Oke, 2007). The number of levels of visibility is cited as a factor that might contribute to a reduction of transaction costs (Barratt & Oke, 2007).

In addition to this, one more time trust is seen in the literature as a mediation effect to reduce transaction costs. In this case, a relationship between supply chain visibility and perceived trustworthiness and transaction costs was observed in a previous study (Dyer & Chu, 2003). These statements lead to the second hypothesis:

H2: there is a positive correlation between the number of tiers involved in data collection and reduction of transaction costs.

The third construct presented in Figure 1 is frequency. Frequency refers to the frequency of data collection about supply chain partners. It is considered one of the aspects that form information quality (Zhou & Benton Jr, 2007). It is also identified as a key aspect needed to guarantee an effective IP (Tushman & Nadler, 1978).

As the maximum frequency of supply chain data collection that can be achieved, real-time monitoring has been a special theme cited in previous studies as a possible way of reducing transaction costs. For example, Aron, Bandyopadhyay, Jayanty, and Pathak (2008) study the effects of real-time monitoring and show how it can reduce uncertainties and, as a consequence, transaction costs. The hypothesis that corresponds to this construct is:

H3: there is a positive correlation between the frequency of data collection in a supply chain and reduction in transaction costs.

Human resources (HR) skills will also be considered in this research. This construct refers to the skill sets (internal or acquired through a third-party) that a firm needs to collect data in the supply chain. These skill sets allow the use of data collection technologies, which enables the firm to manage transaction costs (Sanders, 2008).

Aron et al. (2008) explain into further details how the lack of the required skills sets for supply chain data collection undermines knowledge sharing in supply chains and increase transaction costs. They state that less skilled managers might end up spending more time and making more efforts (for example, through the need of policy readjustments more often) in decision-making processes related to supply chain data collection than more skilled managers would. This forms the following hypothesis:

H4: there is a positive correlation between the availability of human resources skills to collect data in the supply chain and reduction in transaction costs.

Finally, the last construct is supply chain collaboration. An important concept in the domain of supply chain collaboration that has been associated with lower transaction costs is supply chain connectivity, referred to as the ability to use technology to “synchronize decision making across value-added activities” (Fawcett, Wallin, Allred, Fawcett, & Magnan, 2011). This ability is expected to decrease transaction costs as it eases monitoring (Fawcett et al., 2011). So the last hypothesis is:

H5: supply chain collaboration regarding data collection technologies is positively correlated with reduction in transaction costs.

The next section explains how these constructs will be operationalized and how the hypotheses will be tested.

3. RESEARCH DESIGN

3.1. SAMPLE

A survey will be run to test the five hypotheses exposed in the conceptual model. Each of the constructs that form the hypotheses will be measured through scale items that will be further detailed in the next topics. To preserve the reliability of the measures, only supply chain professionals will be targeted to answer this survey as the questions explore concepts that require specific supply chain knowledge and work experience to be answered. Subscribers to a supply chain magazine called Tecnologística were chosen as the main target as they fit the targeted respondent profile.

An invitation to answer the survey was sent by e-mail to more than 2000 subscribers and they were given 15 days to fill the sent form with their answers. Then, a second round of invitation was sent to the same subscribers so that more observations could be collected in the following seven days. A total of 113 supply chain professionals answered to the survey, what represents 5% of the audience that received the invitation to answer the survey.

As all the hypotheses rely on some data collection about suppliers, the 58 respondents who stated that their companies don't collect any data from its supply chain didn't have to answer questions about the reduction of transaction costs. As a consequence, their answers will not be considered in the data analysis part. The answers from the resulting 55 participants constitute the basis for the analyses that will be run. A summary containing discriminant data about these participants is presented on Appendix B.

3.2. QUESTIONNAIRE AND OPERATIONALIZATION OF VARIABLES

Besides a selection of questions aimed at operationalizing each variable that is part of the conceptual model, the following control variables were added to the survey: current role level in the company, length of time in the role, time working for the current company, number of employees, company’s last year revenue and industry. The question and answer options corresponding to each of these variables is offered in Appendix B, those corresponding to independent variables are exposed in Appendix C and those corresponding to the dependent variable (reduction of transaction costs) are in Appendix D.

The data collection technology construct was assessed through a question that asks participants to answer the degree of implementation of each one among a set of technologies that were found in the literature to be used in SCM. A scale varying from “We don’t use it” (equals 1 point) to “We use it in a complete and mature way” (equals 5 points) was used to allow each data collection technology item to be measured quantitatively.

The number of tiers involved will be assessed through a question that asks respondents how much visibility (as a percentage of the number of suppliers) the company has over each level in the supply chain. Percentage ranges were used in a scale from one being equal to “up to 20%” and five points being represented by “between 81% and 100%”.

In order to assess frequency, just one scale item was chosen as it reflects the pace of data collection cited as a trigger to increase IP capacity (Tushman & Nadler, 1978; Zhou & Benton Jr, 2007). A 1 to 7 point scale will be used to measure how often data about the supply chain is collected. In this case, 1 point means on demand data collection and 7 points mean real time data collection.

Human resources Skills were assessed through two different items. The first one is measures through a five-point scale the percentage of the supply chain/purchasing/logistics/operations department has the needed skills to manipulate data collection about the supply chain. For the second item, a similar scale was used to measure the percentage of managers (or executives) associated with the same departments who possess the required skills to use this information to make decisions.

Supply chain collaboration is also measured through a five-point scale that is similar to the one used to measure the data collection technology construct in the sense that it evaluates the degree of implementation each item. The items, in this case, represent collaboration practices used with other companies in the supply chain to collect data, like external audits and joint planning.

Finally, the 4 constructs that compose the reduction of transaction costs and were adapted from Grover and Malhotra (2003) are measured through 5 point Likert scales that assess how much the participant agrees with each statement related to each one of the 4 constructs. The constructs are *effort, monitor, problem and advantage*.

3.3. DATA ANALYSIS

In this part, all the steps of analysis of the data will be explained. Each construct was treated separately and then each regression models were developed to test correlations between independent (related to *data collection technologies, number of tiers involved, frequency, human resources and supply chain collaboration*) and dependent variables (related to the reduction of transaction costs) and then the respective hypotheses. Although each hypothesis was tested separately, as H1, H2 and H5 have each more than 3 or more items associated to the respective independent construct, they were all analyzed in the same way: using exploratory factor analysis and then a linear regression model. H3 and H4 were analyzed simply through a linear regression model. All the analyses were run on R Studio software.

3.3.1. ANALYSIS OF H1, H2 AND H5

To test each of these hypotheses, with the aims of reducing data size and being able to compare it with results related to the reduction of transaction costs, the exploratory factor analysis method was chosen (Hair, Black, Babin, Anderson, & Tatham, 2006). Before running the analysis itself, some steps were taken to prepare for it. First, a correlation table relating all the different items belonging to the independent construct was created. Then, a heat map was drawn to expose higher and lower correlations among the items. Thus, a Bartlett sphericity test was run to test for the homogeneity of variances. After verifying the significance of the test for a 0.05 significance level (Hair et al., 2006) to ensure that the factor analysis could be done, a Kaiser-Meyer-Olkin test was used to measure the sampling adequacy for the complete model. Having achieved a Measure of Sampling Adequacy (MSA) higher than 0.50, factor analysis could be run (Hair et al., 2006). In the cases where overall MSA was lower than 0.50, the item with the lowest MSA was deleted from the database and the new MSA value was calculated and this process was repeated until the MSA value was at least 0.50. Then, in order to define an adequate number of factors, two different criteria were used, and their results were compared: the latent root criterion and the percentage of variance criterion.

Finally, the factor matrix was calculated with the number of factors defined through the 2 methods cited above. When there were components with communalities lower than 0.50, the one with the lowest communality value was deleted from the database and all the preparatory steps were redone as well as the factor matrix itself until all remaining items had communalities of 0.50 or higher. Then, in some cases there were components with loadings lower than the minimum threshold of 0.70 for a sample size of 55 observations (Hair et al., 2006) across all factors. So, the VARIMAX method was used to rotate the factor matrix with the objective of maximizing factor loadings. Still, in some cases some components had loadings lower than 0.70 across all factors. To solve this issue, the component with the lowest maximum loading across all factors was deleted from the database and then all the preparatory steps and the factor matrix itself were redone until all remaining components had one loading of 0.70 or more. Then, similarities among the components of each factor were used to give them a name and allow the interpretation of the results.

Having defined the factors, the value for each observation of each second order variable was calculated as the arithmetic average of the components that are part of the factor. Then, a linear regression model was made having as independent variables all the factors associated to the construct and as dependent variable one of the factors related to the reduction of transaction costs at a time. Then, the results of the models were interpreted through the comparison of the p-value of each linear regression model and a significance level 10%, which is commonly used in exploratory studies as this one (Hair Jr, Hult, Ringle, & Sarstedt, 2016).

3.3.2. ANALYSIS OF H3 AND H4

As in H3, there is just one item and in H4 there are only 2 items, linear regression models were used in both cases. For H3, a linear regression model was run for each factor of reduction of transaction costs as the dependent variable at a time and for H4 the same was done but there 2 independent variables in the latter.

4. DATA ANALYSIS

4.1. INTRODUCTION TO THE FINDINGS

In this chapter, the findings for each hypothesis will be presented. For the reduction of transaction costs and the independent variables in H1, H2 and H5, the final correlation table, heat map, Bartlett test of sphericity results, the MSA results, the criteria that drove the choice of the number of factors, the factor matrix and the names of the factors will be presented and interpreted. For all hypotheses, the results of the linear regression models will also be presented and interpreted.

4.2. REDUCTION OF TRANSACTION COSTS

The first set of data that was analyzed was the one regarding the reduction on transaction costs as it is needed to test all hypothesis. After removing items with factor loadings lower than 0.70 one by one, the final data set is correlated as shown in Table 3. A heat map was also drawn to enable the interpretation of the correlation table, as shown in Figure 2. Strong correlations are noted especially between TC1 (effort) and TC2 (monitor) items with TC4a and TC4b.

Table 3 – Correlation table of remaining items related to the reduction of transaction costs

	TC1a	TC1b	TC1c	TC1d	TC2a	TC2b	TC2c	TC2d	TC2e	TC2f	TC2g	TC3a	TC3b	TC3d	TC4a	TC4b	TC4c	TC4d	TC4e
TC1a	1.000																		
TC1b	0.966	1.000																	
TC1c	0.951	0.930	1.000																
TC1d	0.919	0.967	0.925	1.000															
TC2a	0.884	0.892	0.860	0.858	1.000														
TC2b	0.941	0.936	0.949	0.906	0.927	1.000													
TC2c	0.899	0.907	0.905	0.871	0.922	0.929	1.000												
TC2d	0.882	0.903	0.887	0.854	0.891	0.938	0.934	1.000											
TC2e	0.923	0.903	0.915	0.858	0.917	0.947	0.920	0.918	1.000										
TC2f	0.949	0.983	0.903	0.940	0.895	0.930	0.882	0.905	0.902	1.000									
TC2g	0.918	0.939	0.896	0.945	0.870	0.933	0.928	0.909	0.884	0.913	1.000								
TC3a	0.937	0.931	0.914	0.899	0.849	0.905	0.851	0.855	0.876	0.916	0.881	1.000							
TC3b	0.894	0.887	0.857	0.882	0.797	0.839	0.812	0.800	0.811	0.862	0.866	0.896	1.000						
TC3d	0.712	0.752	0.700	0.716	0.673	0.706	0.754	0.731	0.741	0.721	0.757	0.810	0.774	1.000					
TC4a	0.651	0.689	0.653	0.677	0.593	0.669	0.685	0.725	0.638	0.682	0.702	0.801	0.709	0.788	1.000				
TC4b	0.679	0.719	0.666	0.708	0.605	0.664	0.633	0.631	0.679	0.693	0.687	0.830	0.739	0.839	0.89	1.000			
TC4c	0.756	0.779	0.758	0.751	0.686	0.741	0.746	0.737	0.778	0.753	0.745	0.815	0.801	0.767	0.846	0.906	1.000		
TC4d	0.858	0.866	0.878	0.871	0.800	0.850	0.876	0.875	0.896	0.832	0.883	0.859	0.830	0.738	0.774	0.760	0.870	1.000	
TC4e	0.739	0.775	0.740	0.761	0.688	0.783	0.761	0.779	0.746	0.750	0.816	0.798	0.782	0.783	0.870	0.863	0.935	0.830	1.000

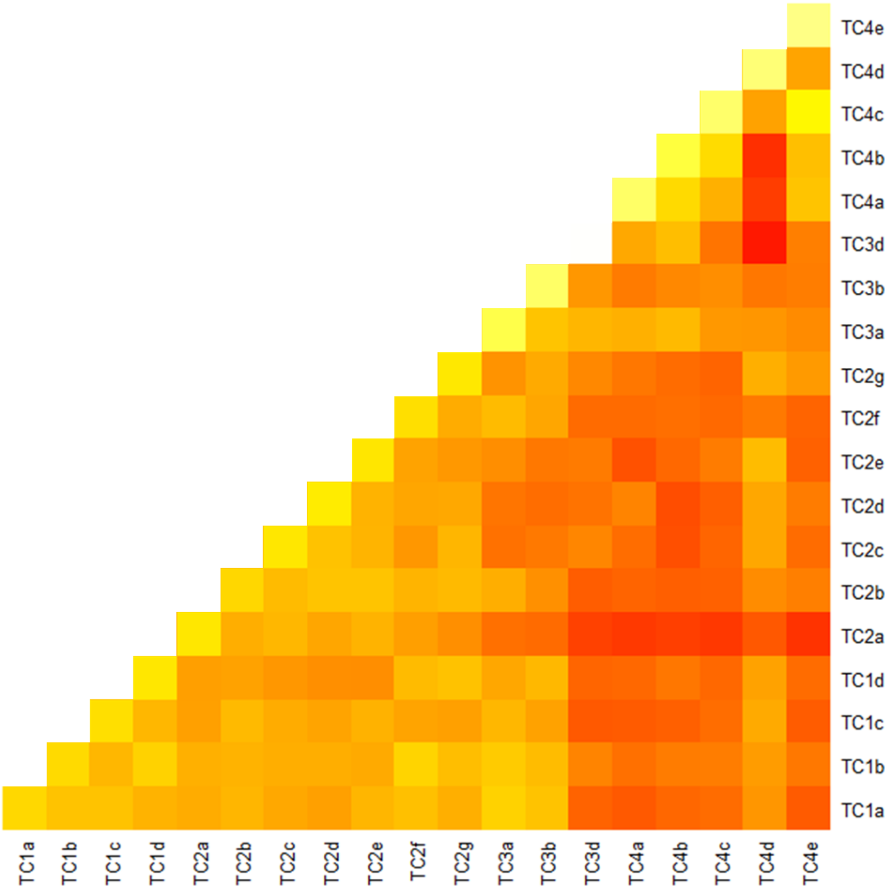


Figure 2 – Heat Map of items related to the reduction of transaction costs

Then, the Bartlett test of sphericity was run, and the p-value found was 2.02×10^{-137} which is higher than the chosen significance level of 0.05 so the data set is suitable for a data reduction method. Then, the Kaiser-Meyer-Olkin (KMO) test was run to measure the sampling adequacy. The MSA found is 0.61 which is considered mediocre (Hair et al., 2006) but still higher than 0.50 so high enough for factor analysis.

The eigenvalues were calculated, and 2 components have eigenvalues higher than 1. So the latent root criterion points to 2 factors. Using the percentage of variance criterion with a minimum cumulated proportion of variance of 60 percent as a minimum satisfactory for social sciences (Hair et al., 2006), we should have a minimum of 1 component since the cumulative proportion of variance of the main component is 75%. So, the number of 2 factors was chosen as it fits both methods.

Then the factor matrix was rotated using the VARIMAX method and is shown in Table 4 (the values in green represent significant loadings – those higher than 0.70).

Table 4 – Factor matrix of items related to the reduction of transaction costs

Items	Factor1	Factor2
TC1a	0.882	0.414
TC1b	0.868	0.459
TC1c	0.864	0.415
TC1d	0.838	0.456
TC2a	0.864	0.339
TC2b	0.884	0.404
TC2c	0.848	0.41
TC2d	0.836	0.421
TC2e	0.841	0.431
TC2f	0.867	0.431
TC2g	0.838	0.458
TC3a	0.746	0.607
TC3b	0.722	0.552
TC3d	0.487	0.714
TC4a	0.36	0.849
TC4b	0.339	0.907
TC4c	0.465	0.825
TC4d	0.701	0.6
TC4e	0.488	0.791

The final factors solution is then composed by two factors: Factor 1 will be is composed by all the effort and monitor items as well as 2 items from the problem set which are TC3a ("It became clearer how to solve problems in the relationship with this supplier") and TC3b ("We started to adopt standards of solutions or strategies for problems that might occur with this supplier") and one item from the advantage set which is TC4d ("This supplier rarely fails to deliver what had been promised"). As most of the factors with high loadings are related to monitoring costs, this factor will be named monitoring costs. Factor 2 is composed by all the other advantage items as well as TC3d ("We started to be able to find solutions for the problems and they are usually less customized now"). As all the items in this factor are related to opportunism, this Factor 2 will be called risks of opportunism.

Having defined the 2 factors, the two respective second order variables (TCF1 for monitoring costs and TCF2 for risks of opportunism) were created as the average of all the items belonging to the factor. These 2 variables will be used to test the 5 hypothesis that are further explored in the next subchapters.

4.3.HYPOTHESIS 1

The data collection technology items will also be reduced to a certain number of factors using exploratory factor analysis and then their correlations with monitoring costs and risks of opportunism will be tested separately. After deleting the items with communalities lower than 0.50 and loadings lower than 0.70 one by one, the final correlation matrix of data collection technology items was obtained (Table 5) and the corresponding heat map was created to allow the interpretation of the results (Figure 3). Strong correlations are perceived especially between T8 (QR Code) and T9 (Point of Sales Hardware) with T1 (Blockchain).

Table 5 – Correlation table of data collection technology items

	T1	T3	T4	T8	T9	T12
T1	1.000					
T3	0.618	1.000				
T4	0.609	0.605	1.000			
T8	0.587	0.780	0.647	1.000		
T9	0.393	0.616	0.793	0.667	1.000	
T12	0.645	0.714	0.562	0.786	0.558	1.000

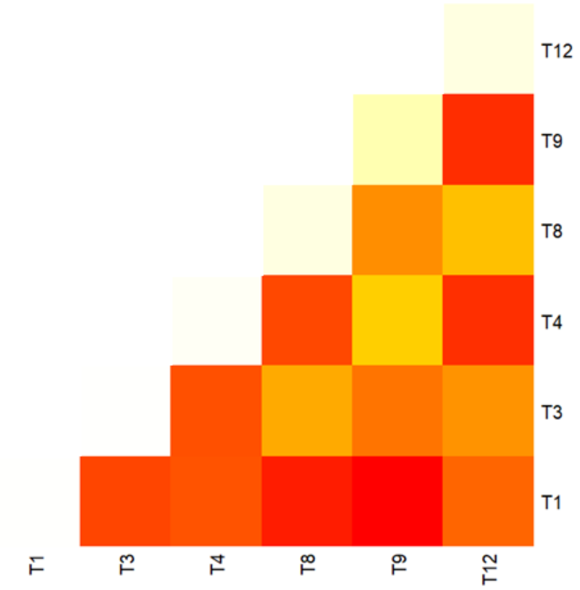


Figure 3 – Heat Map of data collection technology items

Once again, the Bartlett test was run, and the p-value found was 2.45×10^{-41} so we can apply a data reduction method. The KMO test was also run and the overall MSA value found is 0.81 so the sample is considered meritorious and thus appropriate for factor analysis.

To find the ideal number of factors, the latent root criterion was applied, and it pointed to only 1 factor (1 component with eigenvalue over 1). Then the percentage of variance criterion was applied considering the 60 percent threshold for social sciences and it pointed to a minimum of 2 factors with 80.6% cumulative proportion of variance. So, 2 factors were used but they produced loadings that were lower than the ones found in 3 factors solutions. In this case, probably none of the methods was able to predict the ideal number of factors. Then the 3 factors solution was applied, and it was rotated using the VARIMAX method to maximize factor loadings and then the final factor matrix was created (Table 6).

Table 6 – Factor matrix of data collection technology items

Items	Factor1	Factor2	Factor3
T1	0.360	0.224	0.903
T3	0.706	0.345	0.317
T4	0.276	0.877	0.347
T8	0.810	0.391	0.230
T9	0.458	0.732	0.000
T12	0.735	0.269	0.355

Three factors are identified. Factor 1 is composed by Cookies (T3), QR Code (T8) and Systems shared with suppliers and partners (CPFR, VMI, ECR and others) (T12). As the most significant loadings are on Shared systems and QR Code, this factor was names Shared code. Factor 2 is composed by RFID (T4) and Point of Sales Hardware (T9). As both are related to hardware solutions, Factor 2 will be named Hardware-related. The last factor is composed only by blockchain so it will keep the name of Blockchain.

The second order variable for each factor was created: TF1 is the average of the values of the items of Shared Code, TF2 is the average of the values of the items of Hardware-related and TF3 holds the Blockchain values.

Then 2 linear regression models were made: the first one has TCF1 as dependent variable (Table 7) and the second one has TCF2 as dependent variable (Table 8).

Table 7 – Summary of the linear regression model of monitoring costs and the data collection technology items

	Coefficient	Standard Error	p-value
(Intercept)	2.433	0.260	0.000
TF1	0.486	0.190	0.013
TF2	-0.151	0.163	0.361
TF3	-0.033	0.158	0.837

Table 8 - Summary of the linear regression model of risks of opportunism and the data collection technology items

	Coefficient	Standard Error	p-value
(Intercept)	2.393	0.233	0.000
TF1	0.511	0.170	0.004
TF2	-0.193	0.147	0.193
TF3	-0.074	0.142	0.603

In both models, the only variable that has a p-value lower than 0.10 is TF1. So, from the 2 linear regression models, it is possible to conclude that Shared Code is positively correlated with monitoring costs and it is also positively correlated with risks of opportunism. As a result, hypothesis 1 is validated, in other words, there is a positive correlation between the use of data collection technologies to monitor supply chain links and reduction of transaction costs. However, not all data collection technologies will have a significant effect on reducing transaction costs.

4.4.HYPOTHESIS 2

One more time, exploratory factor analysis would be used to reduce data size about the number of tiers. However, when running the Bartlett test for sphericity, the sample was rejected for data reduction methods as the p-value of 0.10 found is higher than the significance level chosen for the test (0.05). Then, the linear regression models (Table 9 and Table 10) will be run straight way using the 3 different depth items as independent variables.

Table 9 - Summary of the linear regression model of monitoring costs and items related to the number of tiers

	Coefficient	Standard Error	p-value
(Intercept)	2.290	0.292	0.000
N1	0.103	0.169	0.547
N2	0.392	0.341	0.255
N3	-0.235	0.248	0.349

Table 10 - Summary of the linear regression model of risks of opportunism and items related to the number of tiers

	Coefficient	Standard Error	p-value
(Intercept)	2.177	0.266	0.000
N1	0.084	0.154	0.588
N2	0.216	0.310	0.490
N3	-0.023	0.226	0.920

None of the p-values found in both models are lower than 0.10 so the hypothesis H2 (“there is a positive correlation between the number of tiers involved in data collection and reduction of transaction costs”) should be rejected.

4.5.HYPOTHESIS 3

As cited above, Hypothesis 3 has only one underlying item. So, 2 simple linear regression models are run (Table 11 and Table 12).

Table 11 - Summary of the linear regression model of monitoring costs and frequency

	Coefficient	Standard Error	p-value
(Intercept)	2.159	0.322	0.000
F1	0.180	0.066	0.009

Table 12 - Summary of the linear regression model of risks of opportunism and frequency

	Coefficient	Standard Error	p-value
(Intercept)	2.138	0.295	0.000
F1	0.154	0.061	0.014

From the models above, it’s possible to conclude that frequency is significantly and positively correlated with both monitoring costs and risks of opportunism. Hence, Hypothesis 3 (“there is a positive correlation between the frequency of data collection in a supply chain and reduction of transaction costs”) is validated.

4.6.HYPOTHESIS 4

To test Hypothesis 4, once again 2 linear regression models will be run – one for each factor of reduction of transaction costs (Table 13 and Table 14).

Table 13 - Summary of the linear regression model of monitoring costs and the human resources items

	Coefficient	Standard Error	p-value
(Intercept)	2.417	0.286	0.000
HR1	-0.275	0.220	0.216
HR2	0.434	0.189	0.026

Table 14 - Summary of the linear regression model of risks of opportunism and the human resources items

	Coefficient	Standard Error	p-value
(Intercept)	2.324	0.262	0.000
HR1	-0.153	0.202	0.454
HR2	0.317	0.174	0.074

As it can observed in the tables above, only HR2 (percentage of the managers of supply chain/logistics/operations/purchasing department and of the top management that have analytical skills to do a qualified evaluation of the information to make decisions) is significantly correlated with monitoring costs and risks of opportunism. Therefore, Hypothesis 4 (there is a positive correlation between the availability of human resources skills to collect data about supply chain links and reduction of transaction costs) is validated but it is only the skills of managers to make decisions from the data collected about the supply chain that is associated with the reduction of transaction costs.

4.7.HYPOTHESIS 5

To analyze supply chain collaboration, a new attempt of reducing data through exploratory factor analysis was tried and then, after eliminating 2 items because of low communality and low factor loading, just 3 items were remaining, and their variances were not homogeneous enough to be reduced as a p-value of 0.08 was found in the Bartlett’s test. Therefore, simple linear regressions were employed having all 5 supply chain collaboration items as independent variables (Table 15 and Table 16).

Table 15 - Summary of the linear regression model of monitoring costs and the supply chain collaboration items

	Coefficient	Standard Error	p-value
(Intercept)	2.124	0.367	0.000
C1	0.233	0.157	0.144
C2	-0.018	0.122	0.885
C3	0.002	0.178	0.990
C4	0.037	0.243	0.879
C5	0.060	0.207	0.773

Table 16 - Summary of the linear regression model of risks of opportunism and supply chain collaboration

	Coefficient	Standard Error	p-value
(Intercept)	2.057	0.328	0.000
C1	0.266	0.140	0.064
C2	0.029	0.109	0.790
C3	-0.078	0.159	0.629
C4	-0.170	0.217	0.438
C5	0.246	0.185	0.191

From the 2 models, it can be observed that C1 (“meetings between members of different supply chain participants”) is the only one of the supply chain collaboration items which is significantly correlated with the reduction of transaction costs and this relationship can only be observed for risks of opportunism. Hence, Hypothesis 5 (supply chain collaboration is positively correlated with reduction of transaction costs) is validated as there is at least one item associated with a decrease in transaction costs.

4.8.SUMMARY OF THE FINDINGS

In this part, the items that were used to assess the reduction of transaction costs were reduced to two groups: monitoring costs and risks of opportunism. The group of technologies called Shared Code (Cookies, QR Code and Shared Systems) is positively correlated with reductions in both types of transaction costs. The same relationship applies to both frequency of supply chain data collection and managers’ skills to use data about their supply chains to make decisions. Meetings between members of different supply chain participants are only correlated with risks of opportunism. The number of tiers involved, however, is not correlated with reductions in transaction costs. Figure 4 presents the final conceptual framework. Only the arrows that represent validated hypotheses were kept. The 4 initial transaction costs constructs were transformed into the 2 remaining factors in this framework. The values over or under each arrow represent the coefficient of correlation between the two variables being connected.

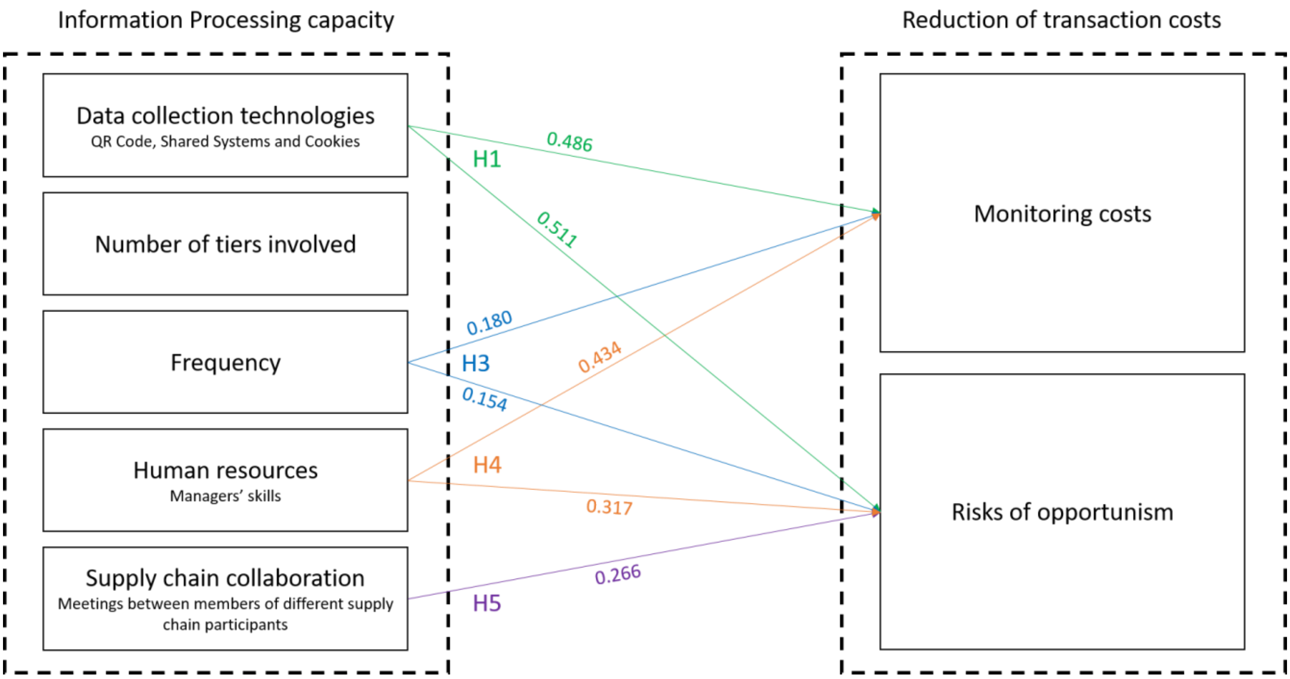


Figure 4 – Final Conceptual Framework

5. CONCLUSION

5.1.THEORETICAL IMPLICATIONS

This study introduced the relationship between the group of data collection technologies and reductions in transaction costs. Not only was the use itself of data collection technologies explored but also other aspects related to them that also contribute to increasing IP capacity such as the number of tiers involved, frequency, human resources and supply chain collaboration. The next paragraphs explain how this relationship was explored in this research.

The first hypothesis ("there is a positive correlation between the use of data collection technologies to monitor supply chain links and reduction of transaction costs") was validated since it was found a positive correlation between the use of a group of technologies composed by cookies, QR code and systems shared with suppliers and partners and the reduction of both monitoring costs and risks of opportunism.

The use of cookies in SCM has been cited as a way to track customers and improve transaction customization (Thirumalai & Sinha, 2011) through a continuous learning relationship between the parties involved in the transaction (Kaplan & Haenlein, 2006), increasing their IP capacities. However, as the use of cookies in the literatures has been focused on tracking customers, and not suppliers, the direct link between the use of cookies and the decrease in transaction costs remain unexplained. On the other hand, the use of QR Code has been associated to improves in trust through security (Yuen et al., 2019) or verification systems (Antonucci et al., 2019) used in supply chains so its effect on reducing transaction costs was expected. The relationship between the use of shared systems and the decrease in transaction costs was also expected because of the increasing trust and the reduced bounded rationality expected as more information is shared in the supply chain as an outcome of this kind of system.

Sharma, Pathak, Borah, and Adhikary (2019) associate higher numbers of supply tiers with higher coordination costs and opportunism risks for the focal firm. However, the results of the test of the second hypothesis ("there is a positive correlation between the number of tiers involved in data collection and reduction of transaction costs") contradict this as no correlation was found between these constructs.

The data supported a significant positive correlation between frequency and both monitoring costs and risks of opportunism. This implies that, although frequency of data collection is associated with an increase in IP capacity (Tushman & Nadler, 1978), it didn't translate into a decrease in transaction costs, countering the expected relationship between IP capacity and transaction costs supported in this study through the works of Massimino et al. (2018); (McIvor, 2009; Tushman & Nadler, 1978).

A more specific way to formulate the validated part of the fourth hypothesis would be the following: there is a positive correlation between the availability of managers' skills to collect data about supply chain links and reduction of monitoring costs. The literature supports the importance of managerial skills to handle technologies in SCM (Waller & Fawcett, 2013) and the importance of data collection technologies to manage transaction costs (Sanders, 2008) but the reason why the staff members' skills don't seem to have any significant impact on transaction costs remain unexplained. A possible reason to be explored is the fact that staff members' skills to collect and manipulate supply chain data may prove useless if decision makers don't know how to use it.

The last hypothesis ("supply chain collaboration is positively correlated with reduction of transaction costs") was rejected since the data didn't support significant correlation between any of the practices and the factors related to the reduction of transaction costs. This result is not supported by a previous study from Flynn et al. (2016) who conclude that supply chain collaboration is a foundation for trust. This divergence might be explained by the fact that they considered supply chain collaboration as a broader construct than it could be captured by the survey used in this study.

5.2.MANAGERIAL IMPLICATIONS

As a result of the assessment of these hypotheses, some managerial implications can be driven. Managers willing to implement projects aiming to reduce transaction costs can use the results of this research to focus their attention and resources to a specific set of technologies: Cookies, QR Code and Systems shared with suppliers and partners and the reduction of both monitoring costs and risks of opportunism. Managers should also find ways to use these technologies to collect data about their supply chains more often so that they can act faster to correct any possible problems found. As supply chain data will be used by managers to take decisions, a project of reduction of transaction costs, should also take into account the availability of managers possessing the skills to do this in the organization. To sum up, organizations should invest on technologies like Cookies, QR Code and Shared Systems, use them to increase the frequency of supply chain data collection and guarantee that their supply chain managers have the skills to make decisions based on these data. Each of these initiatives might help them to reduce their transaction costs.

5.3.LIMITATIONS OF THE STUDY

This study has some limitations. First, only Brazilian companies were surveyed. As a result, some characteristics intrinsic to the country might have influenced the results. For example, the degree of implementation of some technologies, such as blockchain, are still in the initial phase for most of the companies surveyed, which limited the capability of the study to capture a possible correlation between this data collection technology and decreases in transaction costs. Then, this study cannot guarantee the generalizability of its results to other countries.

Second, the sample size used was 55 companies. As a consequence, factor analysis couldn't be run in some cases because of the factor loading of 0.70 required for this sample size and then, items had to been analyzed individually against transaction costs items in linear regression models. So, some significant correlations that could hypothetically only be found by grouping a number of items, couldn't be identified. Moreover, R^2 values weren't used because this sample is not statically representative of the population, so this study didn't intend to find the most precise model that explains the relationships between each of the 5 constructs and transaction costs. Only the significance level and the sign of the coefficients were considered.

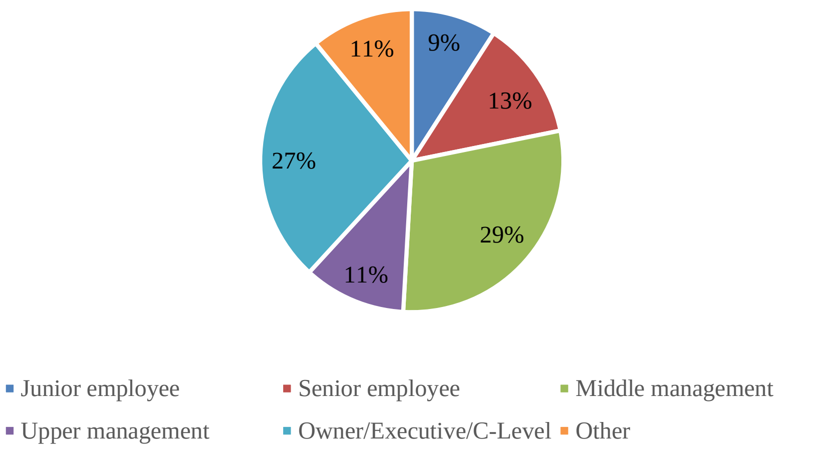
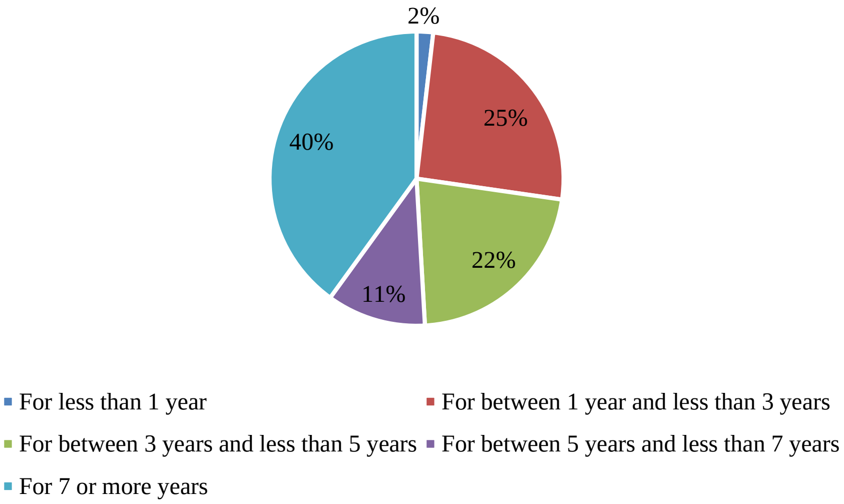
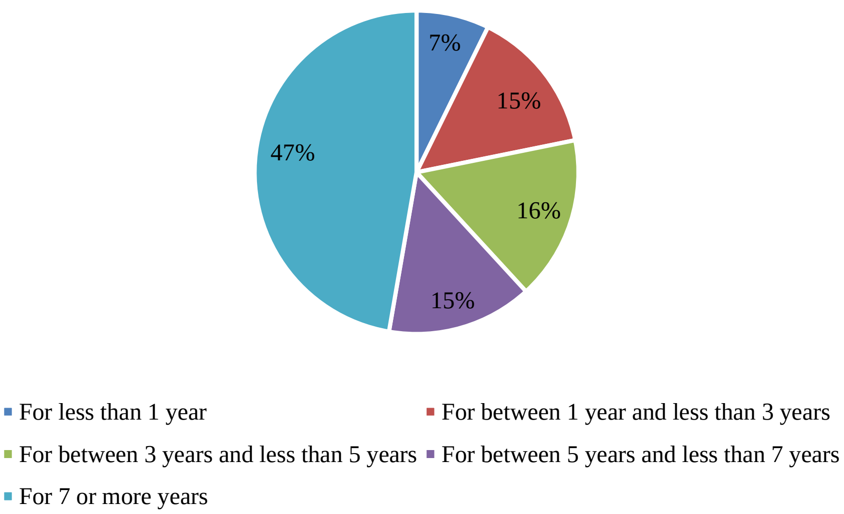
Third, the constructs were measured through questions that explored some possible items related to each of them. Even though an effort was made to find as much items to be included in each question as there are in the literature, some might still be missing, what causes a bias in the results.

Even though there were limitations, the research objective was met as the research question "Can data collection technologies that decrease IP misfits contribute to reduce transaction costs?" was answered. Five constructs related to the reduction of transaction costs through data collection technologies were found in the literature. All the related hypotheses were tested, and their results were analyzed.

5.4.FUTURE RESEARCH

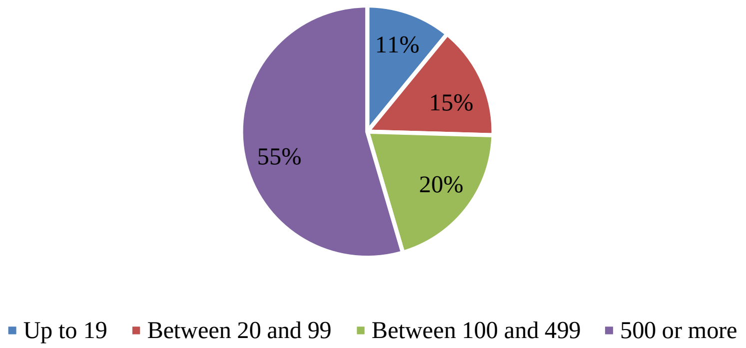
This study can be used as a support for the development of a range of different studies that could focus on areas that remain unexplored. The three elements whose significance as transaction costs reducers (data collection technologies, frequency and human resources) could also have their effectiveness compared so that managers could better understand which ones to prioritize. In this study, only correlation was assessed. So, a causality study could be the next step to understand the relationship between each of these elements and the reduction of transaction costs. Instead of using participants' opinion, a study measuring the correlation of the amounts of investments of companies in *data collection technologies, number of tiers involved, frequency, human resources and supply chain collaboration* and the reduction in transaction costs could be useful for managers to better economically evaluate their decisions around this types of investments before making them. The scale used to assess the Human resources hypothesis could also be refined so that the respondents could have a clearer understanding about the skills needed to collect data about the supply chain links. The possible role of the number of tiers involved, frequency, human resources and supply chain collaboration as mediators of the relationship between data collection technologies and the reduction in transaction costs could also be studied. Systems shared with suppliers and partners, such as VMI, need a certain level of pre-existing trust among supply chain members to succeed but the intent to implement this kind of system might create mediation effects (communication, employee commitment and top-management commitment) for building the needed trust in supply chains (Brinkhoff, Özer, & Sargut, 2015). This relationship could also be tested in new studies. Finally, this research is focused in supply chain data collection for decreasing transaction costs. Another study could identify ways of reducing transaction costs through data analysis.

APPENDIX A – Part 1

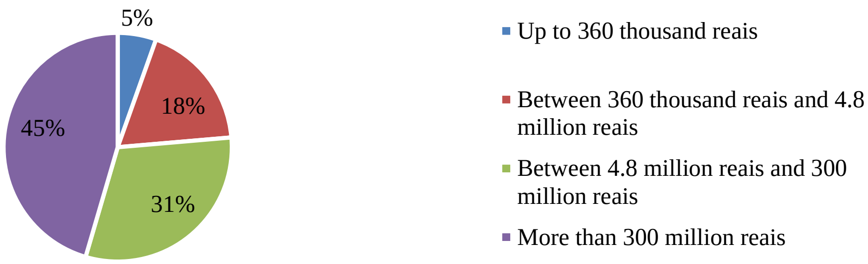
Role level

Length of time in the role

Time working for the current company


APPENDIX A – Part 2

Number of employees



Company's last year revenue



Distribution of participants among industries

Industry	**Number of participants who belong to the industry**
Education	3
Transportation and freight	9
Air transportation and aero spatial industry (including defense system)	1
Manufacturing	5
Retail and consumer durables	3
Trade and logistics	6
Financial services	1
Advertising and marketing	2
Food and beverages	2
Information technology	3
Health and pharmaceutical	3
Government	1
Public utility services	3
Machines and housing	2
Energy and mineral extraction	3
Agriculture	3
Electronics	1
Construction	2
Automotive	2

APPENDIX B – Part 1

Control Variable	Question	Answer options
Current role in the company	Which of the following best describes your current role level?	a. Junior employee b. Senior employee c. Middle management d. Upper management e. Owner/Executive/C-Level f. Other (specify)
Length of time in the role	How long have you been in this role?	a. For less than 1 year b. For between 1 year and less than 3 years c. For between 3 years and less than 5 years d. For between 5 years and less than 7 years e. For 7 or more years
Time working for the current company	How long have you worked in the company?	a. For less than 1 year b. For between 1 year and less than 3 years c. For between 3 years and less than 5 years d. For between 5 years and less than 7 years e. For 7 or more years
Number of employees	How many people work for your company nowadays?	a. Up to 19 people b. Between 20 and 99 people c. Between 100 and 499 people d. 500 or more people
Company’s last year revenue	Which of the following describes your company’s annual revenue?	a. Up to 360 thousand reais b. Between 360 thousand and 4.8 million reais c. Between 4.8 million and 300 million reais d. More than 300 million reais
Industry	Which of the following best describes the industry your company is in?	a. Agriculture b. Food and beverages c. Automotive d. Trade and logistics e. Construction f. Education g. Electronics h. Energy and mineral extraction i. Entertainment and leisure j. Government k. Manufacturing l. Machines and housing m. Unemployed n. Advertising and marketing o. Health and pharmaceutical p. Public utility services q. Financial services r. Information technology s. Air transportation and aero spatial industry (including defense system) t. Transportation and freight u. Retail and consumer durables

APPENDIX B – Part 2

Construct	Hypothesis	Question and scale items	Answer options and points
Data collection technologies (T)	H1: there is a positive correlation between the use of data collection technologies to monitor supply chain links and reduction of transaction costs	For capturing/collecting data about its supply chain, what kinds of technology does your company use? T1: Blockchain T2: Web crawlers T3: Cookies T4: RFID T5: EDI T6: Image capture T7: Bar code T8: QR code T9: Point of Sales hardware T10: Internet of Things T11: Sensors T12: Systems shared with suppliers and partners (CPFR, VMI, ECR and others)	1. We don't use it 2. We use it in a pilot project – only with 1 supplier 3. We use it with only part of our suppliers 4. We make full use of it but we're still making some adjustments 5. We use it in a complete and mature way
Number of tiers involved (N)	H2: there is a positive correlation between the number of tiers involved in data collection and reduction of transaction costs	Answer the following items relative to the suppliers over which your company has some visibility. With the use of the information that was collected/monitored/processed, your company has been able to have some visibility over its: D1: Direct supplier (first level) D2: Indirect supplier (second level) D3: Supplier beyond the second level	1. Up to 20% 2. Between 21% and 40% 3. Between 41% and 60% 4. Between 61% and 80% 5. Between 81% and 100%
Frequency (F)	H3: there is a positive correlation between the frequency of data collection in a supply chain and reduction of transaction costs	F1: How often is this data collected?	1. On demand 2. Annually 3. Semiannually 4. Monthly 5. Weekly 6. Daily 7. In real time
Human resources (HR)	H4: there is a positive correlation between the availability of human resources skills to collect data about supply chain links and reduction of transaction costs	Answer the following questions related to the analytical skills of the team. HR1: What percentage of the supply chain/logistics/operations/purchasing department are knowledgeable about analytical tools to manipulate this information (knowledge about technologies, problem solving, tools and analytical/statistics models etc)? HR2: What percentage of the managers of supply chain/logistics/operations/purchasing department and of the top management have analytical skills to do a qualified evaluation of the information to make decisions?	1. Up to 20% 2. Between 21% and 40% 3. Between 41% and 60% 4. Between 61% and 80% 5. Between 81% and 100%

Supply chain collaboration (C)	H5: supply chain collaboration is positively correlated with reduction of transaction costs	Which of these collaboration practices does your company use together with other companies in the supply chain to collect data? C1: Meetings between members of different supply chain participants C2: External audits C3: Joint planning C4: Defining and acquiring together technologies of common use to collect data C5: Defining and acquiring together technologies of common use to analyze data	1. We don't use it 2. We use it in a pilot project – only with 1 supplier 3. We use it with only part of our suppliers 4. We make full use of it but we're still making some adjustments 5. We use it in a complete and mature way

APPENDIX C

Construct	Question and scale items
Effort (TC1)	After implementing the processes of data gathering and analysis, regarding the relationship with this supplier: TC1a: The effort to gather the necessary information to define the relationship with this supplier was reduced. TC1b: It became easier to deal with the main issues and details needed to develop this relationship. TC1c: It became easier to solve details that hadn't been specified at the beginning of the relationship and had to be solved throughout the relationship with this supplier. TC1d: The effort to define the individual roles of the company and the supplier was reduced.
Monitor (TC2)	In monitoring the performance of this supplier: TC2a: It became easier to determine if we are receiving a fair treatment from him. TC2b: The effort to detect if this supplier meets the specifications and quality standards was reduced. TC2c: The effort to evaluate if this supplier serves us in a far way was reduced. TC2d: Evaluating this supplier in a precise way became easier. TC2e: Our worries about the possibility of this supplier taking advantage of this relationship were reduced. TC2f: Monitoring the performance of this supplier became less time-consuming. TC2g: Monitoring the performance of this supplier became cheaper.
Problem (TC3)	In solving problems that arised in the relationship with this supplier after the implementation of the project of data collection and analysis: TC3a: It became clearer how to solve problems in the relationship with this supplier. TC3b: We started to adopt standards of solutions or strategies for problems that might occur with this supplier. TC3c: Solving problems became less challenging due to the nature of what we buy from this supplier. TC3d: We started to be able to find solutions for the problems and they are usually less customized now.
Advantage (TC4)	Regarding the probability of this supplier taking advantage of the relationship with our company after implementing the project of data collection and analysis: TC4a: There is no more interest from this supplier to act in its self-interest in a manner that is detrimental to ours. TC4b: It is no longer easier for this supplier to alter facts and thus get what it wants. TC4c: There is no longer a strong temptation from this supplier to hold or distort information in his own benefit. TC4d: This supplier rarely fails to deliver what had been promised. TC4e: There is no longer significant motivation from this supplier to take advantage of unspecified contract terms or those that cannot cause legal duties.

REFERENCES

Alt, R. (2017). Electronic markets on transaction costs. In: Springer.

Antonucci, F., Figorilli, S., Costa, C., Pallottino, F., Raso, L., & Menesatti, P. (2019). A review on blockchain applications in the agri-food sector. *Journal of the Science of Food and Agriculture, 99*(14), 6129-6138. doi:10.1002/jsfa.9912

Aron, R., Bandyopadhyay, S., Jayanty, S., & Pathak, P. (2008). Monitoring process quality in off-shore outsourcing: A model and findings from multi-country survey. *Journal of Operations Management, 26*(2), 303-321. doi:10.1016/j.jom.2007.02.014

Barratt, M., & Oke, A. (2007). Antecedents of supply chain visibility in retail supply chains: A resource-based theory perspective. *Journal of Operations Management, 25*(6), 1217-1233. doi:10.1016/j.jom.2007.01.003

Brinkhoff, A., Özer, Ö., & Sargut, G. (2015). All You Need Is Trust? An Examination of Inter-organizational Supply Chain Projects. *Production and Operations Management, 24*(2), 181-200. doi:10.1111/poms.12234

Busse, C., Meinlschmidt, J., & Foerstl, K. (2017). Managing information processing needs in global supply chains: A prerequisite to sustainable supply chain management. *Journal of Supply Chain Management, 53*(1), 87-113.

Clemons, E. K., Reddi, S. P., & Row, M. C. (1993). The impact of information technology on the organization of economic activity: The “move to the middle” hypothesis. *Journal of management information systems, 10*(2), 9-35.

Coase, R. H. (1937). The nature of the firm. *economica, 4*(16), 386-405.

Deichmann, U., Goyal, A., & Mishra, D. (2016). Will digital technologies transform agriculture in developing countries? *Agricultural Economics, 47*(S1), 21-33. doi:10.1111/agec.12300

Dyer, J. H., & Chu, W. (2003). The role of trustworthiness in reducing transaction costs and improving performance: Empirical evidence from the United States, Japan, and Korea. *Organization science, 14*(1), 57-68.

Fawcett, S. E., Wallin, C., Allred, C., Fawcett, A. M., & Magnan, G. M. (2011). INFORMATION TECHNOLOGY AS AN ENABLER OF SUPPLY CHAIN COLLABORATION: A DYNAMIC-CAPABILITIES PERSPECTIVE. *Journal of Supply Chain Management, 47*(1), 38-59. doi:10.1111/j.1745-493X.2010.03213.x

Flynn, B. B., Koufteros, X., & Lu, G. (2016). On Theory in Supply Chain Uncertainty and its Implications for Supply Chain Integration. *Journal of Supply Chain Management, 52*(3), 3-27. doi:10.1111/jscm.12106

Foerstl, K., Meinlschmidt, J., & Busse, C. (2018). It's a match! Choosing information processing mechanisms to address sustainability-related uncertainty in sustainable supply management. *Journal of Purchasing and Supply Management*, *24*(3), 204-217.

Foerstl, K., Schleper, M. C., & Henke, M. (2017). Purchasing and supply management: From efficiency to effectiveness in an integrated supply chain. *Journal of Purchasing and Supply Management*, *23*(4), 223-228.

Galbraith, J. R. (1974). Organization design: An information processing view. *Interfaces, 4*(3), 28-36.

Grover, V., & Malhotra, M. K. (2003). Transaction cost framework in operations and supply chain management research: theory and measurement. *Journal of Operations Management*, *21*(4), 457-473. doi:10.1016/s0272-6963(03)00040-8

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis (Vol. 6)*: Upper Saddle River, NJ: Pearson Prentice Hall.

Hair Jr, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*: Sage publications.

Hallikas, J., Virolainen, V.-M., & Tuominen, M. (2002). Risk analysis and assessment in network environments: A dyadic case study. *International Journal of Production Economics, 78*(1), 45-55. doi:https://doi.org/10.1016/S0925-5273(01)00098-6

Handley, S. M., & Benton, W. C. (2013). The influence of task- and location-specific complexity on the control and coordination costs in global outsourcing relationships. *Journal of Operations Management, 31*(3), 109-128. doi:10.1016/j.jom.2012.12.003

Johnson, P. F., Klassen, R. D., Leenders, M. R., & Awaysheh, A. (2007). Utilizing e-business technologies in supply chains: The impact of firm characteristics and teams. *Journal of Operations Management, 25*(6), 1255-1274. doi:10.1016/j.jom.2007.01.005

Kaplan, A. M., & Haenlein, M. (2006). Toward a Parsimonious Definition of Traditional and Electronic Mass Customization. *Journal of Product Innovation Management, 23*(2), 168-182. doi:10.1111/j.1540-5885.2006.00190.x

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. doi:10.1111/1540-5915.02289

Klein, B., Crawford, R. G., & Alchian, A. A. (1978). Vertical integration, appropriable rents, and the competitive contracting process. *The journal of Law and Economics, 21*(2), 297-326.

Kraljic, P. (1983). Purchasing must become supply management. *Harvard business review, 61*(5), 109-117.

Kumar, A., Liu, R., & Shan, Z. (2019). Is Blockchain a Silver Bullet for Supply Chain Management? Technical Challenges and Research Opportunities. *Decision sciences*. doi:10.1111/deci.12396

Lee, I., & Lee, B.-C. (2010). An investment evaluation of supply chain RFID technologies: A normative modeling approach. *International Journal of Production Economics, 125*(2), 313-323. doi:https://doi.org/10.1016/j.ijpe.2010.02.006

Maltz, A. (1994). Outsourcing the warehousing function: economic and strategic considerations. *Logistics and Transportation Review, 30*(3), 245.

Manuj, I., & Mentzer, J. T. (2008). GLOBAL SUPPLY CHAIN RISK MANAGEMENT. *Journal of Business Logistics, 29*(1), 133-155. doi:10.1002/j.2158-1592.2008.tb00072.x

Massimino, B., Gray, J. V., & Lan, Y. (2018). On the Inattention to Digital Confidentiality in Operations and Supply Chain Research. *Production and Operations Management, 27*(8), 1492-1515.

McIvor, R. (2009). How the transaction cost and resource-based theories of the firm inform outsourcing evaluation. *Journal of Operations Management, 27*(1), 45-63.

Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). DEFINING SUPPLY CHAIN MANAGEMENT. *Journal of Business Logistics, 22*(2), 1-25. doi:10.1002/j.2158-1592.2001.tb00001.x

Min, S., Zacharia, Z. G., & Smith, C. D. (2019). Defining Supply Chain Management: In the Past, Present, and Future. *Journal of Business Logistics, 40*(1), 44-55. doi:10.1111/jbl.12201

Monczka, R. M., Petersen, K. J., Handfield, R. B., & Ragatz, G. L. (1998). Success factors in strategic supplier alliances: the buying company perspective. *Decision sciences, 29*(3), 553-577.

Ngai, E. W., Cheung, B. K., Lam, S., & Ng, C. (2014). RFID value in aircraft parts supply chains: A case study. *International Journal of Production Economics, 147*, 330-339.

Pilling, B. K., Crosby, L. A., & Jackson, D. W. (1994). Relational bonds in industrial exchange: An experimental test of the transaction cost economic framework. *Journal of Business Research, 30*(3), 237-251. doi:https://doi.org/10.1016/0148-2963(94)90054-X

Qrunfleh, S., & Tarafdar, M. (2014). Supply chain information systems strategy: Impacts on supply chain performance and firm performance. *International Journal of Production Economics, 147*, 340-350.

Rosemberg, L. M., Leonardo; Lee, NS. (2019). The transformative potential of blockchain applications for supply chain management: A combined transaction cost and value perspective. *International Journal of Information Management*.

Sanders, N. R. (2007). An empirical study of the impact of e-business technologies on organizational collaboration and performance. *Journal of Operations Management, 25*(6), 1332-1347. doi:10.1016/j.jom.2007.01.008

Sanders, N. R. (2008). Pattern of information technology use: The impact on buyer–suppler coordination and performance. *Journal of Operations Management, 26*(3), 349-367. doi:10.1016/j.jom.2007.07.003

Sharma, A., Pathak, S., Borah, S. B., & Adhikary, A. (2019). Is it too complex? The curious case of supply network complexity and focal firm innovation. *Journal of Operations Management*. doi:10.1002/joom.1067

Srinivasan, R., & Swink, M. (2018). An Investigation of Visibility and Flexibility as Complements to Supply Chain Analytics: An Organizational Information Processing Theory Perspective. *Production and Operations Management, 27*(10), 1849-1867. doi:10.1111/poms.12746

Strassmann, P. A. (1997). *The squandered computer: evaluating the business alignment of information technologies*: Strassmann, Inc.

Swink, M., & Schoenherr, T. (2015). The Effects of Cross-Functional Integration on Profitability, Process Efficiency, and Asset Productivity. *Journal of Business Logistics, 36*(1), 69-87. doi:10.1111/jbl.12070

Tan, K. C. (2001). A framework of supply chain management literature. *European Journal of Purchasing & Supply Management, 7*(1), 39-48.

Thirumalai, S., & Sinha, K. K. (2011). Customization of the online purchase process in electronic retailing and customer satisfaction: An online field study. *Journal of Operations Management, 29*(5), 477-487. doi:10.1016/j.jom.2010.11.009

Tokar, T., & Swink, M. (2019). Public Policy and Supply Chain Management: Using Shared Foundational Principles to Improve Formulation, Implementation, and Evaluation. *Journal of Supply Chain Management, 55*(2), 68-79. doi:10.1111/jscm.12190

Tushman, M. L., & Nadler, D. A. (1978). Information processing as an integrating concept in organizational design. *Academy of management review, 3*(3), 613-624.

V. Roth, A. (1996). *Achieving Strategic Agility Through Economies of Knowledge* (Vol. 24).

Waller, M. A., & Fawcett, S. E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics, 34*(2), 77-84. doi:10.1111/jbl.12010

Wamba, S. F., Gunasekaran, A., Dubey, R., & Ngai, E. W. (2018). Big data analytics in operations and supply chain management. *Annals of Operations Research, 270*(1-2), 1-4.

Wang, Z., Hu, H., & Zhou, W. (2017). RFID Enabled Knowledge-Based Precast Construction Supply Chain. *Computer-Aided Civil and Infrastructure Engineering, 32*(6), 499-514. doi:10.1111/mice.12254

Wever, M., Wognum, P. M., Trienekens, J. H., & Omta, S. W. F. (2012). Supply chain-wide consequences of transaction risks and their contractual solutions: Towards an extended transaction cost economics framework. *Journal of Supply Chain Management*, *48*(1), 73-91.

Williamson, O. E. (1983). Credible commitments: Using hostages to support exchange. *The American Economic Review, 73*(4), 519-540.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism*. New York: The Free Press.

Williamson, O. E. (2008). Outsourcing: transaction cost economics and supply chain management. *Journal of Supply Chain Management*, *44*(2), 5-16.

Wowak, K. D., Craighead, C. W., & Ketchen, D. J. (2016). Tracing Bad Products in Supply Chains: The Roles of Temporality, Supply Chain Permeation, and Product Information Ambiguity. *Journal of Business Logistics, 37*(2), 132-151. doi:10.1111/jbl.12125

Yuen, K. F., Wang, X., Ma, F., & Wong, Y. D. (2019). The determinants of customers' intention to use smart lockers for last-mile deliveries. *Journal of Retailing and Consumer Services, 49*, 316-326. doi:https://doi.org/10.1016/j.jretconser.2019.03.022

Zhou, H., & Benton Jr, W. (2007). Supply chain practice and information sharing. *Journal of Operations Management, 25*(6), 1348-1365.

Zsidisin, G. A. (2003). Managerial Perceptions of Supply Risk. *Journal of Supply Chain Management, 39*(4), 14-26. doi:10.1111/j.1745-493X.2003.tb00146.x