LUND UNIVERSITY

SCHOOL OF ECONOMICS AND MANAGEMENT

The effects of RPA in International Business

Analysing the relationship between industry and performance in firms that have undertaken Robotic Process Automation

by

Melker Eriksson

Edward Renton

June 2021

Bachelor's Programme in International Business

Supervisor: Niklas Lars Hallberg

Abstract

This study examines the effect industry has on performance in firms that have undergone Robotic Process Automation. To provide relevant information, previous research on the main concepts were analysed. A gap in existing literature was discovered, which led the authors to investigate whether or not industry has a significant effect on performance in firms that have automated one or more of their business processes. The theoretical framework of this study is based on the Resource-Based View theory as it argues that firm specific competitive advantage is derived from certain exploitable resources.

Following, the authors performed several statistical tests on empirical data containing results of over 100 RPA projects, post automation. Several carefully selected variables were derived from the data set to further analyse these results. By performing four different MANOVA tests the authors found that industry does have a significant effect on performance in firms that undergo Robotic Process Automation. Additionally, the authors also concluded what industry classification captures the benefits of Robotic Process Automation the most. This was measured in hours saved due to automation of business processes. The authors then continued to discuss the limitations of this study, as well as suggestions for future research.

Keywords: Artificial Intelligence, Robotic Process Automation, Business Processes, Firm Performance, Resource-Based View

Acknowledgements

As we are in the final stage of completing this paper, we would like to reflect on the process of writing this bachelor's thesis. The last ten weeks have been both challenging and educational, and it has been a great opportunity to amalgamate the courses that we have been taught at Lund University during the past three years. This project has confirmed once again that hard work and determination enables us to overcome challenges and succeed with whatever we put our minds to.

We are grateful for the guidance and support of everyone who contributed to this thesis. We would like to especially thank our supervisor Dr. Niklas Lars Hallberg for supervising and guiding us through this thesis under these extraordinary circumstances. Additionally, we would like to express our immense gratitude to Paul E. Marshall for his time and invaluable insights within the RPA industry. Also, we would like to thank Dr. Martin Jönsson for his advice on statistical testing methods, as well as the teachers from the statistics department at LUSEM.

Lastly, we would like to thank Dr. Martin Blom for his dedication to this International Business programme. The past three years have given us priceless insights in the field of International Business, as well as valuable knowledge and experiences.

Table of Contents

Abbreviations

AI	Artificial Intelligence
FTE	Full-Time Equivalent
HITL	Human in The Loop
HRB	Hours Returned to Business
ICC	Intraclass Correlation
ICT	Information and Communication Technology
IT	Intelligent Technologies
KPIs	Key Performance Indicators
MANOVA	Multivariate Analysis of Variances
ML	Machine Learning
PB	Payback Period
RBV	Resource-Based View
RPA	Robotic Process Automation

1 Introduction

1.1 Background and Problematization

Companies are currently embracing new technologies to optimise their performance and increase their competitive advantage (Trunk et al. 2020). Traditional constraints on business are steadily being eliminated and this is changing the rules of competition in almost all industries (Iansiti & Lakhani, 2020). Artificial Intelligence technologies no longer only characterise tech-companies since it is one of the largest, if not the largest, disruptive technology currently available to businesses. In the last decade an enormous volume of data in various formats has been generated, faster than ever before. This has demanded further development of newer technologies, accelerated technological progress, and increased computational processing capacity. As a result of this, Robotic Process Automation, which is the automation and optimisation of repetitive tasks and processes, was created (Borges et al. 2021). This raises the research question: how does industry affect performance in firms that undergo Robotic Process Automation?

Following, the three main concepts in this study will be introduced.

Industry

When analysing the implementation of technologies into firms, such as Robotic Process Automation, one must consider the industry in which the firm is operating. Increased complexity in the firm's environment can result in increased organizational risk (Wefald et al. 2020). The industry environment in which a firm is operating can also, according to Porter and McGahan (1997), affect the performance of said firm. This study will look to further analyse this relationship between industry and performance in firms that have undergone RPA automation.

When considering industry, in the study by Wefald et al. (2020), it is stated that industries have become constraining forces within which firms either adapt or perish. Meaning, industries have different attributes which can affect the performance of a given firm. Some industries may therefore be more successful in implementing and adopting certain technologies, such as RPA automation, than others. This study will further analyse which industries have the greatest and smallest effect on firm performance when RPA is implemented into their business processes.

Since this study is focused on the relationship between industry and performance, industry classification can therefore be considered relevant for this study. Industry classification categorises firms into groups, making it easier to analyse their common movements (Chan et al. 2007). This sets the study up well for analysis as similar firms are already presented in their corresponding industry in the data set of which the empirical study of this paper is based on.

Robotic Process Automation

Robotic Process Automation is derived from Artificial Intelligence and Machine Learning and is an automation of business processes. In other words, RPA is software with AI and ML capabilities, and is used to deal with repetitive tasks previously performed by humans, such as 'reading' and 'interpreting' documents to assign or organise them into their appropriate locations (Ortiz & Costa, 2020). According to a study conducted by Craig et al. (2015), a firm can expect an average of 30 % in savings when implementing RPA automation. This is measured at the project level and not at the firm level. This is due to the fact that firms cannot fully automate at once, but rather one project at the time (Craig et al. 2015).

RPA is increasingly being used in firms because of the potential gains in productivity and the savings associated with a robot performing business activities instead of an employee. RPA has proven to be effective in being able to complete repetitive tasks with a smaller error rate than humans (Supitakwong & Jamsri, 2020). The use of AI technologies within the banking industry in the form of RPA is similarly increasing with the advancement of digitalisation (Ortiz & Costa, 2020). Performing a very simple process, such as labeling and organising files of data, can be time consuming for a human but can be done in seconds by RPA (Ortiz &

Costa, 2020). The gains associated with the implementation of RPA incentivises businesses to invest their money in and adopt RPA automation in their processes.

Just like the substitution of blue-collar workers in the manufacturing plants by physical robots, RPA is the substitution of white-collar workers by non-physical robots, or software robots (Ortiz & Costa, 2020). According to Diksha and Sandhu (2021), implementing RPA in the firm's processes reduces time spent on tasks and assignments, which leads to less manpower requirements. When RPA is implemented the role of the worker changes to a supervisor of the RPA robot, meaning even fewer employees are required. This raises concerns about a coming social paradigm shift in which machines would take over business tasks more frequently and at almost all levels (Ortiz & Costa, 2020).

There are typically a few employees that are assigned the task of managing and running the robots when implemented (Park et al. 2019). These employees are the process controllers. According to Craig et al. (2015), automation does not have to make human workers redundant. They instead argue that the most optimal way of implementing automation in a firm is when human workers are using AI technologies as a tool to improve the effectiveness and accuracy of their tasks, and doing larger volumes of it due to the increase in performance that AI technologies enables (Craig et al. 2015).

Firm Performance

Implementing AI technologies, such as RPA, can have large positive effects on business performance, especially when used to automate repetitive tasks (Reis et al. 2020). This is due to the fact that RPA enables a greater workload to be completed quicker and more accurately. Tasks such as interpreting large amounts of data can be both difficult and tedious, and AI technologies have been reported to address this efficiently (Reis et al. 2020). RPA allows for both basic and complex business processes to be accomplished to a higher quality and at a faster rate than is possible by humans, thus having a positive effect on firm performance.

The theoretical base of this study will be Resource Based View theory. This theory was selected based on the framework that it uses the relationship between firm-specific resources and competitive advantage. It was found in the article by Ortiz and Costa (2020), that after six months of collected data, RPA automation had exceeded manual processing by more than

1000 % in terms of performance. Applying the arguments of Resource-Based View theory, a resource is considered valuable if it increases the absolute level of performance of a process (Gautam et al. 2005). One could therefore argue that high levels of firm performance within an industry can be determined by the gains created by RPA optimisation.

Resource-Based View theory further argues that a valuable resource can explain the variance in performance of a business process across competing firms depending on how rare and costly it is to imitate these resources (Gautam et al. 2005). Since the implementation of RPA is a subset of digital transformation that requires a specific set of skills and associate budget, it is therefore hard to directly imitate. This makes RPA a valuable resource between competing firms, according to RBV theory (Elia et al. 2021).

Research

The field of Artificial Intelligence has been researched rigorously since the 1950's (Ortiz & Costa, 2020). The automation of business processes by implementing Robotic Process Automation, however, has only gained attention from researchers more recently. Several studies have been conducted on firms in different industries to try to explain how and why businesses can benefit when adopting AI technologies, i.e. RPA. The selected literature will be analysed to explain and compare the results of RPA implementation in different industries to determine if industry has any effect on firm performance having undergone RPA automation. The authors can conclude that there is a research gap present after analysing previously conducted research, and the effects of industry on performance in firms that have undergone RPA automation has yet to be explained. Relevant information for this study was researched and collected through Lund University's database LUBsearch, Springer, and Google Scholar, by reviewing several renowned journals and articles. The empirical data was, as mentioned, collected from Blue Prism.

The empirical research in this study will be based on 102 results of automation of a single or multiple project in different firms across different industries. The data that will be analysed was retrieved from Blue Prism's database. Blue Prism is a computer software company that implements RPA robots in businesses across numerous industries in the EMEA region and North America. The data set will be empirically analysed by performing several statistical tests. The statistical tests will consist of one independent variable 'Industry' and three carefully selected dependent variables, being ‘Full-Time Equivalent’, ‘Hours Returned to Business’, and ‘Payback period’. The dependent variables will act as proxies for firm performance. This is based on research conducted by Barney et al. (2004), who argues that using results from a few business processes to represent a firm’s overall performance can generate more accurate and reliable results compared to what the actual overall firm performance is measured to be. Also, it should be noted that the results from Blue Prism’s database are self-reported.

Results from these tests will be curated, from which information will be gathered in order to answer the research question. Although the results will not be based on every existing industry, there will be enough data available to construct a comparison across ten different industries. From this the authors will get a view on the factors that affect firm performance when implementing RPA automation across industries.

The observations stated by the authors earlier in this chapter led to the formulation of the research model that will be used in this study, which can be seen below.

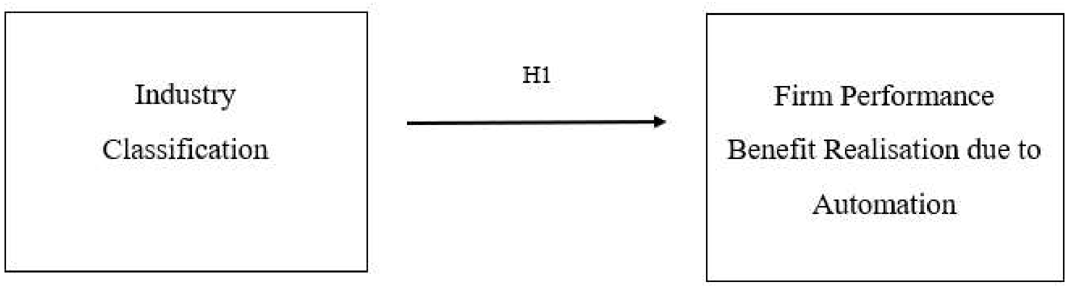

Figure 1 *Research Model*

2 Theory and Hypothesis

2.1 Industry

Industry plays an important role in how processes are employed among firms (Wefald et al. 2010). Processes are experienced differently across different industries and the managerial decisions on the deployment of key resources is linked to firm performance when considering the industry in which the firm competes (Wefald et al. 2010). Due to the fact that managerial decisions vary depending on which industry the firm competes in, different industries undergo different tasks to complete their endeavours, which can be one of the factors that explains the dissimilarity in firm performance across industries.

The industry structure is also asserted to be an important factor controlling competitive behaviours and conducts of firms (Wefald et al. 2010). As industries are structured differently, firms must conduct their processes differently depending on the industry they are operating in, in order to compete effectively. The type of industry in which a firm competes in, as well as the competitive environment of the industry, affects the firm's strategic decisions, which further explains why one can expect a variance in process performance across various industries (Wefald et al. 2010).

Classification of industry is a complex process. Firstly, Standard Industry Classification (SIC) codes are a method which uses aggregate company sales information related to the end product, or using similar production processes (Chan et al. 2007). This method classifies firms into industries based on their end product or services. Another widely used classification system is Global Industry Classification System (GICS). This is based on operational characteristics as well as investor perception of what contributes to the company's main line of business (Chan et al. 2007). Some firms may have end products in one industry, but their main operations are within another. Both methods are useful as they allow for easier analysis of related firms. It is important to classify industries as they reflect common movements in companies' underlying operating performance, which can be measured in sales growth (Chan et al. 2007).

Firms are classified into industries by their characteristics, and these diverse attributes can in turn affect the processes in which a firm undertakes. Industry type can affect organisational learning, growth, risk reduction factors, economising behaviours, firm performance, amongst others (Wefald et al. 2010). Industry type can also be used to predict firm productivity (Wefald et al. 2010). The industry type can be used to analyse how efficient a firm would be in implementing certain processes, by comparing the performance of firms' that are already competing within that industry. According to Porter and McGahan (1997), industry accounts for 19 % of the aggregate variance in a firm's profitability. This is solely based on the relationship between industry and firm performance, and not including the aspect of automation. Porter and McGahan (1997), further argues that industry accounts for larger amounts of variance in firm performance in services and trade industries, but smaller amounts in manufacturing industries. The authors will further discuss whether this is true for industries in which firms undergo RPA automation in chapter 5.

2.1.1 Industry and Information Technology

Information technology plays an essential role in supporting processes in firms (Gautam et al. 2005). The quality of Information Technology will affect the performance of processes across different industries. In light of Resource-Based View theory, firms need to access resources that are valuable, rare, and non-perfectly imitable by competitors (Elia et al. 2021). These are key factors that affect relative process performance as they create a competitive advantage (Gautam et al. 2005). High levels of competitive advantage in one industry could also be an explanation of the variance in levels of relative performance across industries since high performance fosters innovation, which in turn can increase firm performance (Ortiz & Costa, 2020).

The Resource-Based View theory can be applied to understand why the performance of processes may vary across different industries (Gautam et al. 2005). When analysing the use of AI technologies across various industries, if one industry is outperforming another this must be due to specific reasons (Elia et al. 2021). Resources rarity is a possible explanation as to this variance in the performance of a particular process across industries (Gautam et al. 2005).

If one industry possesses a certain IT resource which is rare, in terms of the RBV theory, this could be a reason for that industry to create more value from the use of the particular technology (Elia et al. 2021). Meaning, access to AI technologies which are rare and hard to imitate could be a reason for one industry to outperform another.

Organisations spend millions of dollars on IT each year to improve business performance (Gautam et al. 2005). Based on the aforementioned research, having AI technologies implemented in a firm can increase its performance, and firms which use IT technologies tend to have a competitive advantage over firms which do not use it (Gautam et al. 2005). Yet, since some IT technologies are not rare or costly to imitate, these resources by themselves are unlikely to improve relative performance of business processes (Gautam et al. 2005). For one industry to perform better than another, considering the RBV, AI technologies must be embedded in the process and be valuable by exploiting opportunities where other industries are not able to (Elia et al. 2021).

Tacit, path dependent, and socially complex IT capabilities, also known as shared knowledge, explain variations in process performance according to Gautam et al. (2005). It is not the technology itself but how it is immersed in the processes which makes it outperform competition. Shared knowledge allows for technologies to be appropriately deployed in the process (Gautam et al. 2005). Using technology correctly is more important than having more technology. Low levels of shared knowledge may even reduce performance (Craig et al. 2015). In other words, relative process performance is based on the ability to properly implement and use IT through shared knowledge, hence why one industry may have greater performance than another (Gautam et al. 2005).

2.2 Robotic Process Automation

Business processes are tasks in business that can be performed by people, systems, or more realistically, a combination of both (Ortiz & Costa, 2020). Robotic Process Automation is used to improve the efficiency of processes performed by humans, so that they are completed at a higher standard and in a shorter timespan. RPA can be defined as tools used to reduce the burden of repetitive and exhaustive tasks (Ortiz & Costa, 2020). Business processes such as the organisation of data, manually typing in customer info, and client screening, are both time-consuming and repetitive for humans, but almost instant and fully autonomous with RPA. RPA reduces time spent on tasks and it frees up workers to do work with assignments that require soft skills instead (Keding, 2020).

RPA aims to be a non-invasive technology when considering the IT infrastructure, and it is easy to configure since it does not require programming skills (Ortiz & Costa, 2020). It is considered a ‘lightweight’ IT in a way that fundamental systems are not disturbed by the software that seeks to automate it (Ortiz & Costa, 2020). RPA can be set up so that it does not affect the rest of the IT application landscape within the business. Also, it does not have invasive characteristics and can be used cohesively with other technologies. In its simplest phase, RPA is a type of simple task automation and it does not have the ability to improve itself, contrary to machine learning which can improve itself (Costa & Ortiz, 2020).

Using RPA is becoming more and more desirable as firms are becoming more digital (Craig et al. 2015). Research shows that a UiPath RPA robot takes six seconds per download and four seconds per stock, when it comes to calculation and creation of files (UiPath, 2020). As an example, it would take ten seconds per stock when assessing a portfolio by a software robot. This task would take more than a few minutes for a human to complete (Ortiz & Costa, 2020).

RPA can utilise its AI functions to learn and copy the workflow of human users, then implement them to achieve a better workflow of the system in which it is implemented (Supitakwong & Jamsri, 2020). RPA can be a self-sufficient technology in that it can learn new and improved algorithms via the use of AI, which allows it to become even more efficient over time. RPA is most productive when performing deterministic and repetitive tasks. Assignments or tasks that require cognitive thinking have proved difficult for this type of technology in isolation (Supitakwong & Jamsri, 2020).

The implementation of RPA into firms' processes is beneficial for firms as it provides better quality of work and reduces cases of re-work due to human errors (Diksha & Sandhu, 2021). Another benefit of RPA is that instead of having employees perform repetitive tasks, RPA allows them to use their skills for analysis or other important assignments that require human cognition (Diksha & Sandhu, 2021). RPA frees up employees, by substituting manpower with the use of automation technology. According to Craig et al. (2015), the typical cost saving per process is 30%, as well as improved accuracy and quality, scalability, lower error rate and increased compliance being some of the benefits of RPA implementation.

2.2.1 Artificial Intelligence

Within the banking industry AI technologies have been deemed useful in multiple business processes. Predictive applications are using AI technology to successfully handle risk and credit assessment, such as Know Your Customer (KYC) and Anti-Money Laundering (ALM) (Reis et al, 2020). AI technologies are often used in several different departments in firms, such as marketing and sales departments, market-based definitions and product and services development. This is possible since it can analyse data, such as scanning and interpreting relevant documents, as well automating repetitive tasks (Park et al. 2019). However, there are also more complex processes where AI can be utilised. It can be utilised in processes that interpret satellite imagery to determine how much oil is in vessels in transit, for example. It does so by comparing how low the oil-tanker sits in the water during its transit to allow for the calculation of oil futures (Two Sigma, 2020).

Complex tasks like these, are however, not very prevalent in research since businesses tend to automate easy, repetitive tasks (Craig et al. 2015; Park et al. 2015; Reis et al. 2020). Based on this it can be concluded that the most frequently used AI technologies in business are not the most complex ones, such as satellite imagery used for future price predictions, but rather automation of frequently manually performed tasks (i.e. those served by RPA) that takes up large amounts of employee hours over a longer period.

Research has also been made in the Healthcare industry, where digital transformation seeks to make healthcare safer, more affordable, and accessible for patients (Tyrväinen et al. 2018). It has become a rapidly growing area of research due to the prospective benefits. AI technologies can increase cost effectiveness, drive human-centric care, create new business opportunities, and diagnose patients more accurately (Tyrväinen et al. 2018). A commonly used AI function is a combination of natural language processing (NLP) used for collecting data from patient records to understand the specifics of their situation and diagnose patients with a higher level of security (Tyrväinen et al. 2018). This is possible due to the large amount of data input that can be fed into the software. This will grow more accurate over time. This is possible since the data set will grow for every patient journal it ingests, hence having a larger data set on which to base its calculations and predictions (Tyrväinen et al. 2018).

NLP is a subfield of AI which is the interaction between the human language and computers. It is used for processes where the AI needs to "read" and "understand" data from documents so it can then accurately extract information and insight as well as categorise and organise documents (Keding, 2020). NLP is also used for speech recognition which is commonly used in customer service to shorten the amount of interaction between customers and agents, hence freeing up more time for other tasks (Keding, 2020). This is also a common feature of RPA. In other words, previous research states that firm performance can increase, as well as the competitive advantage, when firms decide to automate business processes in the healthcare industry.

According to Warner and Wäger (2019) AI technologies fundamentally affect the strategy of firms when implemented. It does so by changing the way in which a firm is operating by allowing employees to focus on tasks that require cognitive thinking (Trunk et al. 2020). When used successfully, AI and humans can act in synergy to become more efficient. Since the main goal of AI is to automate processes, it allows humans to focus on activities that will allow them to add more value in innovative departments, such as R&D and HR. This gives people the time to invest in and further develop skills that AI technologies can not adequately perform, but of which are critical for the firm's performance (Trunk et al. 2020; Bharadwaj, 2000). From each of these examples, it is shown that the use of AI technologies can increase firm performance as it cuts down the amount of time required for humans to complete both repetitive and high-volume processes.

Trunk et al. (2020), states that it is important to implement the correct application to the process. Which AI application to use depends on the type, quantity, and quality of data available, which in turn results in various necessities to handle the data, such as; classification, clustering, or detection of connections (Supitakwong & Jamsri, 2020). In other words, firms use different forms of AI applications to complete various tasks. Each industry entails multiple processes, meaning that firms have to consider which AI applications are most useful in their case. The choice of AI application is influenced by various dimensions of data and what the technology is intended to be used for (Trunk et al. 2020).

2.2.2 Intelligent Process Automation

Intelligent process automation (IPA) is a preconfigured software instance that combines RPA, AI, and other technologies, to execute a combination of processes, activities, and tasks in one or more unrelated software systems (Zhang, 2019). There are some tasks that cannot be executed by RPA or AI alone. Hence the need for firms to implement IPA technologies, as it is flexible in that it combines both of these technologies.

As mentioned, RPA is used to automate repetitive tasks in business that have repeatable judgements and structured data, with a single answer or outcome (Supitakwong & Jamsri, 2020). Similar AI technologies are used with IPA in firms to aid and improve decision making. Technologies in the realm of cognitive automation, such as AI and IPA, are used to automate and augment tasks that are unstructured or structured. This means that they produce a set of likely outcomes or interpretations (Zhang, 2019). IPA spans the automation continuum from the "realm of RPA" to the "realm of cognitive automation", making IPA useful in firms as it combines the usage of both RPA and AI (Zhang, 2019). In other words, IPA technology can take unstructured data and transform it to structured data whilst using cognitive automation simultaneously, allowing for automation to take place.

Similar to RPA, if implemented well, IPA can improve efficiency and effectiveness in its own right. It can decrease time spent on repetitive tasks and free up employees to spend more time on high value areas that require professional judgement (Zhang, 2019). Therefore, IPA can enable the employee to use their time more efficiently while the robot performs repetitive tasks for them. IPA can also enhance predictive analytics, and the data collected by RPA can be sent to machine learning modules in the IPA system to help predict future outcomes, such as client behaviour (Zhang, 2019). IPA combines RPA and cognitive automation which allows firms to analyse data at a faster rate and to a more comprehensive degree than if only one of these automation methods was used. As stated in the aforementioned research, AI technologies can increase efficiency in firms which in turn can increase firm performance, yet, there is a gap in the research which compares the gains from RPA automation specifically between industries, and what effect industry has on firm performance (Zhang, 2019; Ortiz & Costa, 2020).

In this study the authors have decided to focus on RPA specifically instead of IPA and AI. One reason is that AI and IPA are used to aid firms with their decision-making processes due to its cognitive thinking abilities, which RPA alone lacks. However, decision making is difficult to analyse since it is objective and can depend on many different factors, which makes the effects it may have on firm performance difficult to quantify. Since IPA is a combination of RPA and AI and does not only focus on automation of repetitive tasks, it will not be included in this study (Zhang, 2019). RPA was chosen as the field of research of this paper as it specifically automates business processes hence making it easier to quantify as compared to other AI technologies.

2.3 Firm Performance

2.3.1 Resource Based View

It is a major challenge for firms to adapt to the ongoing technological development, and digital transformation is a necessary cost-effective means of increasing firm performance (Elia et al. 2021). Digital transformation is the implementation of digital technologies to reform firms and services (Costa & Ortiz, 2020). This is done through either substituting old technologies with current technologies, or fully transforming manual business processes into digital automated processes. Implementing digital processes can allow firms to improve their performance compared to competitors across industries. That is what the authors of this study want to confirm.

Literature on the Resource-Based View (RBV) theory states that firms can achieve this increase in performance by means of tangible and intangible resources. These resources need to be coordinated and combined through strategic capabilities in order to be effective (Elia et al. 2021).

Resources, as defined by Besanko et al. (2009), are company-specific assets, such as branding and patents, that have a direct impact on a company's ability to create a competitive advantage over others. Competitive advantage can be reached through implementing a value creating strategy that has yet to be used by any potential or current competitors (Barney, 1991). Firms can do this by utilising processes which are valuable and rare. For a firm to maintain their competitive advantage over time the company-specific value-creating strategy must not be duplicable or transferable to other companies (Newbert, 2007). This competitive advantage will in turn increase firm performance. The relationship is visualised below.

Figure 2 (Newbert, 2007)

As Barney (1991) argues, the two reasons why competitive advantages can exist are: (i) a company's resources and capabilities are imperfectly mobile, and (ii) heterogeneous. The former meaning that the resources and capabilities cannot be bought on the factor market, and the latter that they are unevenly distributed amongst companies. If these assumptions are not true then there would not be any possibility to construct a value-creating strategy that could not be duplicated by other companies within the same industry, since every player in that industry would have similar resources and capabilities.

Barney (1991) also states that for company-specific resources to lead to competitive advantages, the resource must be: (i) able to reap benefits and neutralise threats; and be (ii) scarce amongst other competitors, both existing and potential. Additionally, for a company-specific resource to be long-lasting and sustainable it must also: be (iii) difficult to duplicate, and (iv) have no equivalent substitute.

However, RBV theory has some limitations. It has been criticised for being too asymmetrical in its assumptions by not taking demand into account, and that the assumption of long-lasting competitive advantage is based on relatively stable market conditions (Foss & Hallberg, 2014; Teece et al. 1997).

2.3.2 Resource-Based View and Robotic Process Automation

Firms' performance and profits are generated based on firms' resources, to be conceived as all of the assets controlled by a firm, that enables the firm to conceive of and implement strategies that improve its efficiency and effectiveness (Elia et al. 2021). RPA is a prime example of this. As stated in the article by Trunk et al. (2020), RPA increases the amount and speed of information collected and interpreted. From a Resource-Based View perspective RPA would be classified as a valuable digital resource as it can create value for the firm in the form of increased efficiency and lower rates of error. Since these machines, or robots, can be purchased by any firm, they are not the source of sustained competitive advantage themselves. When information processing systems are deeply embedded in the decision-making process, only then can they be considered a source of sustained competitive advantage (Elia et al. 2021). Operating business structures change as soon as RPA applications are actively used, thus influencing processes and responsibilities. This brings to light RBV theory since RPA is required to be immersed in the decision-making process of a firm for it to be considered a source of sustainable competitive advantage (Newbert, 2007).

Firms' capabilities have been defined as the firm's capacity to deploy resources for a desired end result. They are information-based, tangible, or intangible processes that are firm specific and developed over time through complex interactions among the firm's resources (Elia et al. 2021). In order for RPA to be considered a capability it has to be aligned with the rest of the processes in which the firm undertakes. RPA must therefore be implemented so that it is cohesive and used as a lever for the employees to handle current operations with ease.

In order for a firm to achieve superior performance, technological resources should be combined with other firm-specific resources. That includes a human component consisting of technical skills and intangible components (Elia et al. 2021). Therefore, in light of the Resource-Based View theory, RPA can be used as an efficient digital capability when aligned with the firm's processes as it further increases the skills of the employees that operate it, ICT proficiency being one of them. ICT proficiency is the ability to use digital technologies, such as communication tools, networks, and software, to solve certain problems. In doing so it creates value for the business by increasing employee productivity and knowledge.

According to Elia et al. (2021), firms should invest in both the adoption of digital technologies and in the development of digital skills if they aim to develop a competitive advantage and increase firm performance in today's markets. Developing digital skills over time is of great importance since employees will be required to learn how to monitor and control the implemented robots. This is another reason why firms should aim to develop their digital skills in order to more efficiently manage digital technologies, such as RPA (Elia et al. 2021). As mentioned earlier, this is known as HITL. Human in The Loop is when a robot performs the task, but the human uses their insight to complete the decision making. As stated in the article by Trunk et al. (2020), RPA offers the potential of robots to augment human capabilities, while it also changes the human role to become more of a supervisor. Through the lens of the Resource-Based View theory, RPA can generate superior performance as it improves the firm's processes by developing the employee's skills and changing their role into that of a supervisor.

In the article by Elia et al. (2021), it is stated that digital resources need to be integrated into the firm's digital capabilities in order to foster a firm-level 'digital competitive advantage'. From a Resource-Based View perspective, firms need to have integrated RPA successfully in order for them to achieve a competitive advantage. However, a study made by Elia et al. (2020), shows that a deciding factor when it comes to competitive advantage is the quality of digital resources, not the quantity. In other words, the quality and completeness of the automation has a larger effect on the possible advantages that firms will experience, than the amount of automated processes in the firm. The reason that RPA is a high-quality digital resource is that it allows for tasks to be performed at a faster rate, with a lower chance of error, and offers benefits such as the amount of information analysed. In summary, Resource-Based View theory suggests that RPA can create a competitive advantage as it allows for higher quality of work completion, which in turn increases firm performance when implemented (Trunk et al. 2020).

2.4 Limitations of RPA

Looking into the use of RPA, when implemented in firms, researchers have shown that it also has its limitations. Firstly, data labelling (i.e. process development) in RPA is very laborious and time consuming since every input has to be manually explained to the system (Caner & Bhatti, 2020). Data labelling is similar to when toddlers are taught which colour is which from their parents. They have to be told what is blue and now they know that blue is blue. This cannot be explored or analysed, it is just a fact that the general environment has agreed to be true. This is also called supervised learning (Caner & Bhatti, 2020). Another limitation is the bias in AI algorithms. Since AI technologies, and RPA, are created by humans and humans are biased by default, the AI will indirectly be biased. An AI system will not be as biased as the person(s) who created it since it operates from enormous collections of data but will inherently have some amount of bias when it comes to decision making (Caner & Bhatti, 2020). This has shown to be an issue when it comes to recruitment issues, loan and legal decisions, and medical prognosis. Bias, however, can be limited by the involvement of high-level management, which means that the RPA software system is not the final decision maker, but merely assists the employee in gathering the data used for said decision, hence the need for HITL (Keding, 2020).

The main goal of implementing an RPA system is to save both time and capital, yet applying it to business processes in the first place may require a large amount of resources. There are two major reasons for this. The resources needed when automating any process is limited to whether: (i) the process can be automated simply by replacing the human worker with a robotic process, also called a Digital Worker (Blue Prism, 2020), or (ii) rearranging the business model such that it better suits the implementation of automated processing (Keding, 2020). The former strategy is faster to implement and generates improvements and results quicker since the digital worker has now almost directly substituted the human worker. The speed with which this can be done is high since the firm does not need to restructure or rearrange the way daily operations are conducted, nor their business model, for this to be realised (Trunk et al. 2020). The latter strategy is more resource and time consuming but can substantially improve business performance in the long run since the business model is now focused on the ability to have digital workers relieve as much human work as possible. This returns hours to the business that can be used to focus on other areas, such as R&D and management (Trunk et al. 2020).

Looking further into the challenges which are faced when using RPA, not every problem needs to be solved with technology. Even if machines are able to determine the most optimal decision, they are less likely to be able to sell it to stakeholders (Trunk et al. 2020). RPA enables us to complete repetitive processes with ease, yet it is not the solution for everything. Soft skills in general have become increasingly important with the introduction of RPA in organisations, making employees shift their focus toward further development of these skills in order to successfully adapt the usage (Trunk et al. 2020).

Another drawback to RPA is that it lacks explainability for its results. There is an absence of rational thinking which we have as humans, which can limit the usage of AI technologies (Caner & Bhatti, 2020). Where the firm process requires reasoning behind the results, RPA can have its constraints. Therefore, there are some business processes where AI technologies are not necessary even though it is applicable. Sometimes it can be best for the task to be done by a human due to the ability to rationally think.

2.5 Hypothesis

Based on the analysed literature presented in this chapter, the authors draw the conclusion that there is a relationship between industry and firm performance. Some industries may have higher automatability and access to rare technologies which allow them to gain a competitive advantage depending on the industry in which they operate (Gautam et al. 2005). RPA can be an example of this as it is hard to imitate since every process is unique to the firm. Also, once RPA automation is implemented into a firm it has huge potential of scalability. As argued by Krugman et al. (1997), economies of scale, which is the action of scaling production, is a key factor when firms seek to drive down costs and improve their output, hence improving firm performance. As stated by Tyrväinen et al. (2018) and Reis et al. (2020), RPA automation has been seen to show positive effects on firm performance in the Banking and Healthcare industry. RPA automation fulfils the criteria to increase firm performance as it increases overall quality by providing low risk of error and a reduction of time spent on tasks, hence it can be considered a value increasing competitive advantage (Trunk et al. 2020). In light of the

Resource-Based View theory, a process can be considered a competitive advantage if it is rare and performance increasing (Elia et al. 2021). Additionally, as argued by Gautam et al. (2005), various levels of competitive advantage within an industry can explain differences in relative performance.

Although the observations made on previous research argue that automation of business processes will result in an increase in a firm's performance, it does not fully explain the effect industry has on performance when RPA automation is implemented (Wefald et al. 2010; Chan et al. 2007; Porter et al. 1997; Elia et al. 2021; Gautam et al. 2005; Ortiz & Costa, 2020; Diksha & Sandhu, 2021; Newbert, 2007). Concluding the observations made in the selected literature, along with the research model the authors created, the following hypothesis was formulated as an answer to the research question as well as a contribution to existing research:

Hypothesis 1: *Industry has an effect on performance in firms that undergo Robotic Process Automation*

3 Methodology

3.1 Research Design

According to Bryman and Bell (2011), the research design is an extensive explanation of how the different components of the research are used to address the research problem. The research design of this study is quantitative research which will compare relevant data on the results of companies' post-automation performance, between industries. The authors will use the hypothetico-deductive (HD) method, also known as the scientific method, to explain the observations made from a dataset where the outcome is unknown. The HD method is performed by formulating a hypothesis in a way so that it can be falsified, i.e. a null hypothesis (Bryman & Bell, 2011). The hypothesis is derived from previous research and is then tested based on the empirical data that has been collected. This method is used since the authors have empirical data available. Then, the hypothesis will either be falsified or not based on statistical tests made on the observable data. The authors will in other words conduct a hypothetical deductive study on a cross sectional dataset of classified industries and firm performance measured at the project level.

As previously mentioned, the data set that will be examined is retrieved from Blue Prism's database. Blue Prism is a computer software company that specialises in providing software that enables the digitisation and optimisation of business processes across various industries. This data set consists of the results from 154 individual projects in different companies, detailing project specific firm performance pre and post-automation. Relevant information in this data set will be selected and compared systematically followed by a multivariate analysis of the variance (MANOVA) as well as several different statistical tests. What will be deemed relevant is information that makes it possible to quantify and compare results of RPA implementation in each firm, post-automation.

3.2 Sample, Variables, and Data Collection

The purpose of this study is to research if the classification of the firm's industry has an effect on performance within the firm having undertaken RPA automation. Industry will be the independent variable and the authors will use three carefully selected dependent variables as proxies for measuring firm performance. These variables are Full-Time Equivalents saved (FTEs), Hours Returned to Business (HRB), and Payback period (PB). The theoretical population are firms that have undergone RPA automation of one or more business processes with Blue Prism or one of their partners. The population is geographically limited to countries where Blue Prism operates, which is the EMEA region and North America. Also, firms that undergo RPA automation are usually mid- and large cap firms. This is due to both the costs and the complexity associated with undertaking RPA automation, hence the possibility to automate certain business processes is not very prevalent in small cap firms (Craig et al. 2015).

The data points were collected from reports constructed with a short introduction, followed by an existing problem statement that RPA automation can solve. The solution is then presented, and the results are explained. These results were then examined and extracted by the authors. It should be noted that the reported results were self-reported. Then, the sample consisting of 102 companies was selected and relevant information about the firms' performance, such as FTEs, HRB, and PB, was systematically transferred to Excel. After all the data had been analysed for relevant variables, the quantitative data was imported to R where the statistical testing was conducted, followed by interpretation of the results.

3.2.1 Sample

To obtain relevant data the authors first contacted an employee at Blue Prism. The authors stated what the main concept of the thesis was, and the employee said that he could provide data that would support the exploration of the line of reasoning that the authors put forward. When the data was received, the authors analysed the results of 154 different projects and drew a sample of companies that had a consistent set of data points. A total of 102 firms were suited for analysis out of the original 154 firms, which is a large enough sample for the sample means to be considered normally distributed, according to the Central Limit Theorem (Berenson et al. 2016). Through the perspective of international business, the firms of which the samples consist, are all from multinational companies since they conduct business in at least two different countries at the time the data was collected. The firms from which the samples are taken are operating in: Great Britain, The United Arab Emirates, Sweden, Denmark, The United States of America, and the Kingdom of Saudi Arabia, to name a few. This makes the data more diverse and is a better representation of the industries under analysis, and the degree to which industry affects firm performance, which is the goal of this study.

3.2.2 Variables

Independent variable

The population and independent variable in this study will be 'Industry' since the hypothesis is built on the notion that 'Industry' is the reason for differences in firms' performance having undergone RPA automation. The difference between independent and dependent variables is that the dependent variable is the effect, and the independent variable is the cause. In other words, the independent variable does not vary or depend on the dependent variables (Berenson et al. 2014). From the 102 data samples numerous different industries were present.

The industry classification was already stated in the project results. Since it was not possible to know the names of the firms of which the data was constructed, due to legal reasons, the only way to differentiate the projects from one another was for Blue Prism to classify them into industries. The authors used the same industry classification as was stated in the data set by Blue Prism for the analysis. The ten different industries used for analysis are:

- Telecom
- Banking/Insurance
- Consultancy/Advisory
- Healthcare
- Electricity/Gas
- Shipping/Logistics
- Hotel/Tourism,
- Manufacturing,
- IT services
- E-commerce

Dependent variables

The dependent variables that will be analysed are chosen on the basis of what is generally used as Key Performance Indicators (KPIs) when quantifying the effects of RPA. The most frequently used KPIs are: Full-Time Equivalents (FTEs), Hours Returned to Business (HRB), Average Revenue Per User (ARPU), Total Cost of Ownership (TCO), Payback period (PB), and Customer Satisfaction (CSat) (Craig et al. 2015). These KPIs are of importance because they are some of the specific reasons that firms automate their processes. The reason is that they can be improved and optimised with the help of AI technologies, such as RPA.

Total Cost of Ownership and Payback period are ways to measure the cost of implementing and operating RPA processes. Payback period is the time it takes for the investment to make its own price back in cost-savings, which is measured in months and Total Cost of Ownership is the total cost of owning or using the automation process (Ortiz & Costa, 2020).

Full-Time Equivalents is the equivalent of hours of labour that a full-time employee is executing, subsequently carried out by an automated process. Generally, that is 40 hours per week but differs depending on tasks automated and industry. This assumption will be used in this study. In other words, if a firm has saved the equivalent of 5 FTEs, they have automated a workload or workloads that previously was performed by 5 full-time employees.

Hours Returned to Business is somewhat similar but is measured in how many employee hours the organisation now can use for something other than executing the specific task that the robot now is performing. The difference is that a firm can automate processes that replace humans, which is measured in FTEs saved, or processes that focus on linear improvements, which is measured in Hours Returned to Business. (Craig et al. 2015). One FTE saved is when one human worker has been redeployed by automated processes, and one HRB is when processes are optimised by RPA to streamline the processes, such as screening time for a loan application or chat-robots, but not to specifically replace human workers (Craig et al. 2015).

Out of the aforementioned variables, Full-Time Equivalents, Payback period, and Hours Returned to Business are the most prevalent performance metrics looking at the dataset. KPIs such as CSat and ARPU are mostly used for businesses who automate processes that are directly handling customers, such as customer service departments. For that reason, CSat and ARPU will not be a part of the analysis since the variables used for comparison have to be prevalent in the data set across all industries.

Additionally, for uniformity, the variables will be converted to hours. According to Craig et al. (2015), FTE saved can be calculated differently depending on industry and task automated. In this study, one FTE saved is based on the assumption that an employee has five productive hours of work per day and is working five days per week. The number of workdays in a given year is assumed to be 260 for the analysis. This is the average in the USA for reference, though it may differ between countries (Craig et al. 2015). Payback period, which is measured in months, is transformed to hours by multiplying the number of months by 730,001 since that is the average amount of hours in any given month. Hours Returned to Business will not be transformed since it is already measured in hours.

3.3 Data Analysis

It is important when conducting a quantitative study that the data is prepared and analysed correctly for reliable results. The tool used for statistical analysis is the programming and statistical software R, which is frequently used for statistical testing and visualisation in numerous scholarly and academic papers across different fields of study. The data was first curated in Excel, to systematically organise and label the selected variables. Following, the data set was divided into industries, then to be further analysed in R.

The data set was first analysed by the authors, then imported to Excel and organised into columns of FTEs, HRB, PB, and Industry. Then, the relevant data was imported to R so that the desired statistical tests could be performed. Finally, a series of different statistical tests, including correlation and several one-way MANOVA tests, were performed to statistically test the hypothesis. The MANOVA tests were conducted with 'Industry" as the independent variable and 'FTE', 'HRB', and 'PB' as the dependent variables. Following the statistical tests, the results were summarised in tables in Excel for better representation. There will be a summarisation of the findings within each of the ten industries to visualise any dissimilarities. An extensive explanation of the results and relevant findings will be provided in chapter 4.

3.4 Validity and Reliability

The main objective when researching is to maintain high quality and use reliable sources continuously throughout the study. As mentioned, the firms used in the dataset are mid- and large-cap companies and the data was retrieved from Blue Prism's database. Blue Prism is a well credentialed and multinational company that is, according to Gartner's Magic Quadrant of RPA, one of the market leaders in the field of RPA software and implementation and was the first company to use the term Robotic Process Automation. (Gartner, 2020). However, since the data is curated by a single secondary source there is a chance of the data being biased. This will affect our statistical analysis, but it will not have any significance since the analysis will be a comparison between firms in different industries that have automated business processes, rather than comparing firms that have and have not automated these processes.

In other words, the study is not made on the relationship between automation and firm performance because of the possibility of sample selection bias, but rather on how industry classification affects firm performance of those firms who have undergone RPA automation. This mitigates the selection bias.

For the results of the statistical testing to be valid and reliable, the authors wanted to know whether or not the dependent variables chosen for this study are correlated, which would cause a biased skewness in the results. The variables were tested against each other. Further an Intraclass Correlation (ICC) test was performed to determine the correlation between the dependent variables. This was performed due to the possible occurrence of multicollinearity. Multicollinearity can arise when performing multiple variate regression analysis and is the phenomena of two or more of the variables being too correlated, which will negatively affect the statistical results. In other words, all the dependent variables were tested together, and in pairs, meaning a correlation matrix between Full-Time Equivalents and Payback period, Full-Time Equivalents and Hours Returned to Business, and Hours Returned to Business and Payback period. This was conducted to see to what degree the three variables were independent of each other, hence confirming the reliability of the broader tests.

Also, it should be noted that in this study the authors are comparing the performance of RPA automation in individual projects within firms across industries, and not between competing firms within the same industry. Therefore, the variables of imitability and rarity do not have to be considered to explain the relative performance of each industry. Other assumptions were needed for the analysis to be completed, such as the assumption that firm performance is measured by the number of hours saved due to the automation of certain processes. This is why the dependent variables act as proxies. In addition, it is assumed that the savings made in a specific project can be viewed as indicative of savings made for the company as a whole, were a larger scale of automation implementation to be undertaken. This is what the authors define as 'firm performance' throughout this study. This is based on the research conducted by Barney et al. (2004), who argues that using the effectiveness of business processes as proxies to represent the overall firm performance may be more accurate and appropriate than using the actual overall firm performance. This is explained by the fact that process level advantages are not always observable at the firm performance level (Barney et al. 2004).

These assumptions are present for several reasons. For instance, it is difficult to quantify the financial effects of automation of an individual project on the financial performance of the company as a whole. Another reason why the stated assumptions are needed is that it is impossible to fully automate all aspects of a firm's operations in a single project. Typically, company-wide digital transformation programs unfold over multiple years and are the sum of multiple projects. This is based on a study performed by Craig et al. (2015), where it is argued that automation of a company is a fluid and continuous process that has to be modified and managed due to the fact that the performance of technology is doubled every 18 months. Based on this, the analysis of the data can therefore explain the effect of 'Industry' on a firm's performance at a specific point in time, given that the firm has undergone RPA automation.

4 Results

As mentioned, for uniformity and easier statistical analysis the dependent variables 'Full-Time Equivalents', 'Hours Returned to Business', and 'Payback period', will be presented in hours. These results are presented in the appendix (Table 1, Appendix A). Following, the minimum and maximum points, mean, and standard deviation of the dependent variables, were calculated (Table 2, Appendix B). The calculated correlation coefficients between the variables can be seen below.

	FTE	*HRB*	*PB*
FTE	1		
HRB	0.575667011	1	
PB	0.446749311	0.3846521	1

Table 4.1 Correlation between the Dependent Variables (FTE; HRB; PB)

A correlation coefficient of 1 indicates a perfect positive correlation, which means that for every increase in variable X, variable Y will increase equally as much. According to Berenson et al (2014) a correlation coefficient between 0.00-0.30 is negligible; 0.30-0.50 is low; 0.50-0.70 is moderate; 0.70-0.90 is high; and 0.90-1.00 is very high. Based on this, we can conclude that the correlation between variables (FTE; PB) and (HRB; PB) makes them suitable for further statistical analysis, due to 'low' correlation.

Correlation between variables (FTE; HRB) is 'moderate', despite this the authors decided to include this variable in the analysis due to the Intraclass Correlation results. Between all the dependent variables the ICC coefficient was 0.74. The closer the coefficient is to one (1), the higher the reliability of the results, and ICC coefficient values between 0.50-0.75 indicate 'moderate' reliability (Berenson, 2014). The independent variable 'Industry' is divided into ten populations, and the frequency of RPA projects (samples, in the dataset) within each industry is summarised. There can be several reasons why the frequency of samples is different across industries and this will be further discussed in chapter 5.

Having proved that the dependent variables were suited for further statistical analysis, four different Multivariate Analysis of Variances (MANOVA) were performed. This will test if the population (i.e. the Industry) explains the variances in the dependent variables (FTE; HRB; PB). The reason for conducting several different MANOVA tests is that the results from a single MANOVA test might not be sufficient (Olson, 1974). Another reason why MANOVA tests were chosen for this research is that there are multiple dependent variables present. Also, a significance level of alpha = 0.05 is assumed for all testing. This means that if the statistically calculated p-value is less than 0.05 the hypothesis can be rejected, and that there is less than a 5 % risk of making a Type I error. A Type I error is a false positive. The significance level of 0.05 is commonly used in research as the threshold for hypothesis rejection, hence why the authors decided to apply it in this study.

The hypothesis that is going to be tested, as mentioned in 2.6, is whether or not 'Industry' has an effect on firm performance. To statistically test this, a null hypothesis was constructed. This was performed since it is not possible to prove if an experiment is true, but on the contrary one can prove if an experiment is untrue (Berenson, 2014). This hypothesis can be seen below.

H0: *Industry does not have a significant effect on performance in firms that undertake Robotic Process Automation*

When performing null hypothesis testing there has to be an alternative hypothesis, which must be true if the null hypothesis is proven to be untrue (Berenson, 2014). The alternative hypothesis will therefore be the original hypothesis:

Ha: *Industry does have a significant effect on performance in firms that undertake Robotic Process Automation*

4.1 MANOVA Tests

The main statistical test that was performed was Wilks' Lambda. Wilks' Lambda is a multivariate version of the F-test statistic. It can be used for either one-way ANOVA or MANOVA calculations and examines the differences in variance between multiple populations, or groups, against a single variable. That variable in this case is 'Industry'. Wilks' Lambda is a measurement of the percentage variance in dependent variables (FTE; HRB; PB) that are not explained by the differences in the independent variable (Industry). The ideal value is zero, which means the closer the statistic is to zero, the more the dependent variables contribute to the statistical model. The results from this test can be seen below.

Wilks' Lambda	
	Industry
Lambda	0.091
F (Observed values)	9.767
DF1	33
DF2	257
F (Critical value)	1.482
p-value	< 0.0001

Table 4.1 Wilks' Lambda

The Lambda value is close to zero which means that all dependent variables contribute to the model. From the results one can conclude that the null hypothesis should be rejected as the computed p-value is lower than the significance level alpha = 0.05, and the alternative hypothesis H1 should be accepted. The risk of a Type II error is less than 0.01% which makes the results (i.e. rejection) reliable. Also, since calculated F-value > critical F-value it once again supports the notion that the null hypothesis should be rejected.

Following, three additional statistical tests were performed to support the initial results. The results are summarised in table 4.2.

Hotelling-Lawley's Test

Hotelling-Lawley's test, or Hotelling's trace, is a positive-value statistic, similar to Wilk's Lambda. It is the sum of eigenvalues of the test matrix used when performing multivariate tests. An eigenvalue is a value or number explaining the amount of variance there is in the data in a specific direction (Olson, 1974). Contrary to Wilks' test, the desired value of the statistic (Lambda) is anything > 0, meaning the further we stray from zero the larger the difference between the multivariate variables, ergo results being more reliable since the multivariate variables are less correlated.

Pillai's Trace

Pillai's trace is a t-statistic, or test statistic, that is used to evidence if an independent variable has a statistically significant effect on the dependent variables. Its values range from zero to one, where one being strong evidence of significant effect on the dependent variables. Pillai's test is used when the assumption that 'observations are independent' is violated (Johnstone & Nadler, 2017). This is true for this research since the data is collected at one source, hence the inclusion of this statistical test.

Roy's Largest Root

The last statistical test that will be performed is Roy's Largest Root. The difference between Roy's test and the other performed tests is that the focus lies on the effect of extreme eigenvalues in variables individually. In other words, Roy's test is used for statistical tests trying to explain the effect that extreme points have on the dependent variables. Its results are usually compared with Hotelling's Trace to see if the dependent variables are correlated and if that effect contributes to the statistical model (Johnstone & Nadler, 2017). Also, positive Lambda values different from zero indicate that the null hypothesis should be rejected.

Summary

The aforementioned tests generated results that were in line with the results from Wilks' Lambda. The Lambda from Hotelling's test is positively different from zero which reassures the authors that the dependent variables' correlation has low effect on the statistical model, meaning the results are 'reliable' (Johnstone & Nadler, 2017). Comparing the Lambda of Roy's test (1.508) to the Lambda of Hotelling's Trace (3.700), it shows the authors that there is once again a low correlation between the dependent variables, and the negative effect of this on the statistical model is negligible. This is based on the observations made by Johnstone and Nadler (2017) who explain if (Lambda Roy's = Lambda Hotelling's) then there is a strong correlation between the dependent variables, and this has a strong negative contribution to the model. Even though the data was collected from a single source, the results of Pillai's Trace shows that the assumption of independent observations will have a low negative effect on the MANOVA results. This is true because Pillai's Lambda of this test is close to one (1) (Olson, 1974).

Based on these results, the authors are reassured that the independent variable 'Industry' has a significant effect on the dependent variables 'FTE', 'HRB', and 'PB'. The author can therefore falsify the null hypothesis with 99.9 % confidence in all cases, since the computed p-values (0.0001) < alpha (0.05). The research question has now been answered and it has statistically been proven that 'Industry' does have a significant effect on firm performance variables FTE, HRB, and PB, which means that firms will experience different magnitudes of increased performance based on which industry they operate in as they undertake RPA automation. The results are summarised below.

	Wilks' Lambda	Hotelling-Lawley's Test	Pillai's Trace	Roy's Largest Root
Lambda	0.091	3.700	1.642	1.508
F (Calculated value)	9.767	9.604	9.785	12.204
DF1	33	33	33	11
DF2	257	257	267	89
F (Critical value)	1.482	1.442	1.480	1.898
p-value	< 0.0001	< 0.0001	< 0.0001	< 0.0001

Table 4.2 Statistical Test Summary

Industry results

As the results from the MANOVA tests showed that there is a significant difference between firm performance within different industries, having undergone RPA automation, the authors wanted to know which industry had the best and worst performance post RPA automation. As shown in table 4.3 one can see that firms in the Banking/Insurance industry tend to benefit the most from RPA automation, and E-commerce benefited the least, based on the samples and variables used in this study. The means of the dependent variables can be seen in Table 2 (Appendix B). Further analysis and discussion will be made on these results in chapter 5.

Rank	Categories (n=10)	Observations (n=102)	Geometric Mean of total (hrs) saved
1	Banking/Insurance	24	**115,169.3**
2	Consultancy/Advisory	9	**97,423.2**
3	Electricity/Gas	4	**84,558.8**
4	Shipping/Logistics	8	**59,940.4**
5	Telecom	12	**59,036.6**
6	Healthcare	12	**55,488.8**
7	Manufacturing	8	**40,904.3**
8	IT services	10	**37,186.2**
9	Hotel/tourism	7	**28,267.6**
10	E-commerce	8	**17,841.3**

Table 4.3 Industry Performance Summary

5 Discussion and Conclusions

The authors have analysed and statistically tested the effect of ‘Industry’ on firm performance on a sample of 102 Robotic Process Automation projects. Out of the original sample of 154 projects, 102 was deemed fit for further analysis based on the completeness of the variables in each sample. As explained in chapter two, Resource-Based View was used as the theoretical basis for the relationship between using and developing digital capabilities and firm performance. From this theoretical perspective, combined with several studies on the relationship between industry, RPA, and firm performance, a hypothesis was formulated. Following, the results from the statistical analysis will be discussed along with the implications they might have.

The results in this study support the hypothesis that industry has a significant effect on performance within firms that have undergone RPA automation. Meaning, when a firm decides to undertake RPA automation, the industry in which they operate within will have a significant impact on the firm’s performance. In other words, they will experience different outcomes in regard to firm performance. Furthermore, the results also show that firms within the Banking/Insurance and Consultancy/Advisory industry tend to capitalise and benefit the most from RPA automation, whilst industries like E-commerce and IT services had relatively low benefits of RPA automation.

Since RPA is the optimisation of digital systems, the results of this study should be in line with the level of digital adoption, or digital potential, of an industry (Park et al. 2019; Gautam et al. 2005). Firms must have an existing landscape of digital systems in place otherwise there is nothing for RPA to automate. Furthermore, the process that humans conduct on those systems must be repetitive in their nature, and relatively high in value. It is not cost effective to automate low volume processes, and non-repetitive processes. It is easy to see why Banking/Insurance is a high performing industry for automation. Banks have, over the years, engaged in large scale digital transformation (Ortiz & Costa, 2020). This has been necessary to deliver the type of banking services companies and consumers expect, especially where global banking is considered.

As mentioned earlier in the paper, research conducted by Porter and McGahan (1997) claims that industry has the largest effect on firm performance in service and trade industries, and smallest effect on manufacturing industries. Comparing these observations to the results of this study it can be concluded that the results are similar. Manufacturing scored relatively low compared to Banking/Insurance and Consultancy/Advisory. However, since this study was based on a population of firms that has undergone RPA automation, the resemblance between the results of Porter and McGahan's study and the authors could be argued to be accidental.

Additionally, the nature of banking transactions is that they are high in volume and repetitive in their nature, across all aspects of the business. Areas such as performing KYC, Anti-Money Laundering, sanction screening, production origination (issuing credit cards, home loans, auto loans, insurance products) and more complex areas, such as trade finance, are all repeated hundreds of thousands of times per year in banks across the world (Ortiz & Costa, 2020; Park et al. 2018; Reis et al. 2020). These processes often take place with human operators performing steps in multiple different digital systems that have been commissioned over many years of technology adoption. There is no lack of automation potential, and typically, employee salaries are relatively high, meaning automation pays higher dividends in such firms.

Digital natives by comparison, such as firms in industries like E-commerce and IT services, are often operating on integrated or package systems that do not allow for AI technologies like RPA to automate to the same extent as the aforementioned industries. This does not mean that there is no automation potential, but rather that the landscape is different in that there are less digital systems linked together by human operated processes, and that the firm is built on being as efficient and streamlined as possible. This can be one explanation as to why E-commerce and IT services had relatively low volumes of hours saved.

Another possible reason for differing performance levels across industries, considering the RBV, is that the strategy and vision of the high performing firms could also have been transformed in order to make these processes more efficient (Elia et al. 2021). Research states that performance is greater within firms where these technologies are embedded within their processes (Gautam et al. 2005). The implementation of AI technologies requires firms to reorganise their processes, so that employees change their role to a supervisor (Trunk et al, 2020). Therefore, one can draw the conclusion that the high performing industries might have been able to reorganise their structure after RPA implementation to a greater extent than the low performing industries. Another reason behind the differences in performance between industries could be due to the fact that the processes which are being replaced by automation are increasing the absolute level of performance of the processes in that firm (Gautam et al. 2005). From this one can draw the conclusion that the industries which have the most gain, could be industries in which RPA has the most potential to improve the processes in which it is being implemented. Additionally, according to Bharadwaj (2000) the propensity to adopt technology and automatability are large deciding factors when it comes to RPA project success which also confirm the results of this study.

As stated by Iansiti and Lakhami (2020), the difference between traditional operating models and digital operating models, is that AI technologies enable scalability that has not been possible before with industrial means. This can be important to managers in industries that have a low propensity to automate as automation can provide improvements on their operating model due to scalability, hence positively affecting the relative firm performance. As often mentioned in economic theory, economies of scale is what companies strive towards to minimise the average cost per unit produced by increasing output (Krugman et al. 1997). This cost advantage arises from the fact that costs are now spread over a larger amount of goods. The difference between traditional and digital operating models is that traditional operating models stagnate much sooner due to physical limitations. This also supports the author's findings in that the industry with the highest propensity to automate has a greater relative firm performance compared to the others. A visualisation of this can be seen in the figure below.

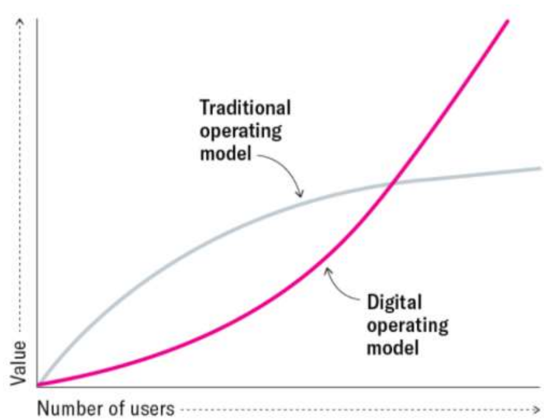

Figure 3 *Competing in the Age of AI* (Iansiti & Lakhani, 2020)

5.1 Limitations of the study

There are several limitations present in this study. The first limitation is that the samples of analysis were from ten different types of industry. This can be argued to be an unrealistic representation of industries as a whole. The reason why some industries are not included in the study is due to the low amount of data available. The researchers could therefore be missing an industry which could be a large outlier, which would affect the results and conclusion of the study. The sample size itself was large in that it was 102 case studies, yet again, this is a selected sample and one could argue that the results generated might not be true to the whole population of the firms who undertake RPA automation of their processes. Another limitation is that the sample selection was not completely random, but rather, samples were collected based on the amount of data available of each firm from Blue Prism's data set.

Since the 102 cases were retrieved from Blue Prism's database, the results might be biased in that they came from a single secondary source. Blue Prism may only want to provide results which promote the positive effects of RPA automation and therefore might not issue cases that show any negative effects of its application. This could be the case since any negative results of RPA implementation would affect their reputation and therefore the operation of their business. It could be argued that results may be positively biased due to this factor. It should also be noted that the results of the 102 RPA projects might lack negative aspects since they are self-reported.

Another limitation to this study is that the industries analysed are of different sizes, meaning the frequency of RPA projects differ in each of the ten selected industries. This means that nuance and scale of the companies within any given industry were not taken into consideration.

Lastly, the fact that there is not an established method available to measure the effects of automation could be argued to be a limitation in this study. According to Craig et al. (2015), this is due to research of RPA and its effects on firm performance are still in its infancy. One cannot calculate the performance or effects of a RPA automation itself on the financial performance of a company as a whole as not every process within a firm can be quantified (Bharadwaj, 2000). Although, the amount of time saved from the implementation of RPA automation on a project can be calculated, and therefore one can draw a conclusion from that information to provide an explanation of the differences of firm performance between industries.

5.2 Future Research

As this article focuses on the relationship between industry and performance of firms that have undergone RPA automation, further research could compare industries more nuanced. A suggestion is to consider the size and number of firms, as well as level of automatability of the firms, used for comparison. This would allow for even deeper analysis since the aforementioned factors also could have a significant effect on firms' performance. Also, since RPA automation is a relatively new subject in the business world there are no generally used blueprints on how measurements of RPA should be reported or which variables to use. This could be researched further to establish consistency and transparency.

As the variables used in this study were focused on hours saved due to automation of business processes, researchers could include metrics, such as avoided costs and fines due to lack of error and increased data reliability, staff retention- and churn rates, and up-sell and cross-sell. Another suggestion is to perform longitudinal studies on RPA automation projects instead of analysing the effects at a specific point in time.

To bring light to more detail than our study was able to provide, further research could analyse how the specific qualities and characteristics of industries give rise to the different potential of automatability. This would not only tell us which industry is the best for implementing automation processes, but the specific fundamental qualities of these industries that allow automation of business processes to be beneficial. This would also allow firms to further improve their understanding of the importance of automation as it can significantly improve their performance.

Technologies that replace human workers by automating business processes may also bring more attention to Corporate Social Responsibility (CSR) as employees could, as a result of automation, be made redundant. This would aim the analysis in the ethical direction and could provide important insight in the issues that arise when implementing automation technologies. This effect could be researched on either a project, firm, or societal level.

References

Barney, J. B. (1991). Firm resources and sustained competitive advantage. Journal of management, 17(1). pp. 99-120, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 8 May 2021]

Barney, J. B., Gautam, R., & Waleed, A. M. (2004). Capabilities, Business Processes, and Competitive Advantage: Choosing the Dependent Variable in Empirical Tests of the Resource-Based View. Strategic Management Journal, 25(1). pp. 23-37, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 3 May 2021]

Berenson, M. L., Levine, D., & Szabat, K. A. (2014). Basic Business Statistics, 13, Pearson Education

Besanko, D., Dranove, D., Shanley, M., & Scheafer, S. (2009). Economics of Strategy. 6(1). pp. 87, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 10 May 2021]

Bharadwaj, A. S. (2000). A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation. Management Information Systems Quarterly. 24(1). pp. 169-196, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 30 April 2021]

Borges, A.F.S., Laurindo, F.J.B., Spinola, M.M., Gonçalves, R.F., & Mattos, C.A., (2021). The strategic use of artificial intelligence in the digital era: Systematic literature review and future research suggestions: Systematic literature review and future research directions. International Journal of Information Management, 57, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 13 April 2021]

Bryman, A. & Bell, E. (2011). Business Research Methods, 3, New York: Oxford University Press

Caner, S. & Bhatti, F. (2020). A Conceptual Framework on Defining Business Strategy for Artificial Intelligence. Contemporary Management Research, 16(3), pp.175-205, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 11 April 2021]

Chan, C. K. L., Lakonishok, J., & Swaminathan, B. (2007). Industry Classifications and Return Comovement. Financial Analysts Journal. Available through: LUSEM Library Website https://lusem.lu.se/library [Accessed 22 May 2021]

Craig, A., Lacity, M., & Willcocks, L. (2015). Robotic Process Automation at Xchanging: A Case Study. London School of Economics. Available online: https://www.blueprism.com/uploads/resources/white-papers/LSE-Case-Study-Xchanging OUWP032015.pdf [Accessed 13 April 2021]

Diksha, (N/A)., Sandhu, K.J., (2021). Robotic Process Automation for Prioritizing the Refund in Online Travel Agency. Available through: LUSEM Library Website https://lusem.lu.se/library [Accessed 6 May 2021]

Elia, S., Giuffrida, M., Mariani, M.M., & Bresciani, S. (2021). Resource and digital export: An RBV Perspective on the role of digital technologies and capabilities in cross-border e-commerce. Available through: LUSEM Library Website https://lusem.lu.se/library [Accessed 4 May 2021]

Foss, N. J. & Hallberg, N. L. (2014). How Symmetrical Assumptions Advance Strategic Management Research. Strategic Management Journal, (35). pp. 903-913. Available through: LUSEM Library Website http://www.lusem.lu.se/library [Accessed 10 May 2021]

Gartner. (2020). Gartner's Magic Quadrant of RPA. Available Online: https://www.gartner.com/en/documents/3988021 [Accessed 19 April 2021]

Gautam, R., Waleed, M. A., & Jay, B. B. (2005). Information Technology and the Performance of the Customer Service Process: A Resource-Based Analysis. Available through: LUSEM Library Website http://www.lusem.lu.se/library [Accessed 15 May 2021]

Hyde, K.F. (2000). Recognising deductive processes in qualitative research. Qualitative Market Research: An International Journal, 3(2), pp. 82-90. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 10 May 2021]

Iansiti, M. & Lakhani, K. R. (2020). Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World. Harvard Business Review. Available through: LUSEM Library Website https://lusem.lu.se/library [Accessed 7 May 2021]

Johnstone, I. M. & Nadler. B. (2017). Roy's Largest Root test under Rank-One Alternatives. Oxford University Press. Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 12 May 2021]

Keding, C. (2021). Understanding the Interplay of Artificial Intelligence and Strategic Management: Four Decades of Research in Review. Management Review Quarterly. Available through: LUSEM Library Website https://lusem.lu.se/library [Accessed 7 May 2021]

Krugman, P. R. (2018). International Economics: Theory and Policy, 11, Pearson

Kumar, R. (2014). Research methodology: A step-by-step guide for beginners. Thousand Oaks: Sage Publications

Newbert, S. L. (2007). Empirical research on the resource-based view of the firm: an assessment and suggestions for future research. Strategic Management Journal, 28(2), pp. 121-146. Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 28 April 2021]

Olson, C. L. (1974). Comparative Robustness of Siv Tests in Multivariate Analysis of Variance. Journal of the American Statistical Association. 69(383). pp. 894-908. Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 14 May 2021]

Ortiz, F.C.M. & Costa, J. C., (2020). RPA in Finance: Applying a software robot in a portfolio optimization problem. Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 20 April 2021]

Park, Y. W., Hong, P., & Geon, C. S. (2019). Digital Transformation Technologies and Capabilities of Technology Utilization: Case Study of Japanese Firms. Technology Management in the World of Intelligent Systems. Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 12 May 2021]

Reis, C., Ruivo, P., Oliviera, T., & Faroleiro, P., (2020). Assessing the drivers of machine learning business value. Journal of Business Research, 117, pp. 232-243, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 11 April 2021]

Porter, M. E. & McGahan, A. M. (1997). How much does Industry matter, really? Strategic Management Journal, 18. pp. 15-30, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 23 April 2021]

Supitakwong, S. & Jamsri, P. (2020). The Effectiveness of RPA in Fine-tuning Tedious Tasks. Faculty of information technology King Mongkut's Institute of Technology Ladkrabang. Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 12 May 2021]

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. Strategic Management Journal, 18. pp. 509-533, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 3 May 2021]

Trunk, A., Birkel, H., & Hartman, E. (2020). On the current state of combining human and artificial intelligence for strategic organizational decision thinking. Business Research, 13(3), pp. 875-919, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 11 April 2021]

Two Sigma. (2020). Our Work and Projects, Available Online: https://dataclinic.twosigma.com/our-work/ [Accessed 5 May 2021]

Tyrväinen, P., Minna, S., Ala-Kitula, A., & Kuoremäki, R. (2018). Identifying Opportunities for AI applications in Healthcare - Renewing the National Healthcare and Social Services. Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 20 April 2021]

UiPath. (2021). AI and RPA, Available Online: https://www.uipath.com/automation/ai-and-rpa?utm_source=GoogleSearch&utm_medium=cpc&utm_term=uipathe&utm_content=511495749534&utm_campaign=AR21FWR [Accessed 9 May 2021]

Warner, K.S.R. & Wäger, M. (2019). Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. Long Range Planning Journal of Strategic Management, 53(3), pp. 326-349, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 12 April 2021]

Wefald, J. A., Katz, P. J., Downey, G. R., & Rust, G. K. (2010) Organizational Slack, Firm Performance, and the Role of Industry. Journal of Managerial Issues. Available through LUSEM Library website https://lusem.lu.se/library [Accessed 21 May 2021]

Wernerfelt, B. (1984). A Resource-Based View of the Firm. Strategic Management Journal, 5, pp. 171-180, Available through: LUSEM Library Website https://lusem.lu.se/library [Accessed 8 May 2021]

Zhang, A.C, (2019). Intelligent Process Automation in Audit. Rutgers, The State University of New Jersey Southwestern University of Finance and Economics, China, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 8 May 2021]

Appendix A

Table 1.

Industry	FTEs (# of employees)	HRB (hours)	PB (months)	FTE (hours	PB (hours	HRB (hours)
Consultancy/Advisory	70	110,000	6	455000	4380.006	110,000
Healthcare	60	220,000	12	390000	8760.012	220,000
Electricity/Gas	75	145,000	24	487500	17520.02	145,000
Shipping/Logistics	23	35,900	9	149500	6570.009	35,900
Banking/Insurance	10	15,600	6	65000	4380.006	15,600
Shipping/Logistics	37	78,000	12	240500	8760.012	78,000
Hotel/Tourism	8	12,480	12	52000	8760.012	12,480
Manufacturing	10	22,000	18	65000	13140.02	22,000
Telecom	40	72,000	9	260000	6570.009	72,000
Telecom	15	16,000	6	97500	4380.006	16,000
IT Services	75	116,000	12	487500	8760.012	116,000
IT Services	10	41,000	6	65000	4380.006	41,000
E-commerce	3	10,000	6	19500	4380.006	10,000
Healthcare	3	300,000	12	19500	8760.012	300,000
Manufacturing	20	25,000	12	130000	8760.012	25,000
Banking/Insurance	17	67,000	6	110500	4380.006	67,000
IT Services	3	7,000	3	19500	2190.003	7,000
Healthcare	2	1,600	9	13000	6570.009	1,600
Banking/Insurance	37	57,000	6	240500	4380.006	57,000
Banking/Insurance	10	16,000	12	65000	8760.012	16,000
Telecom	35	70,000	12	227500	8760.012	70,000
Banking/Insurance	72	115,000	3	468000	2190.003	115,000
Banking/Insurance	70	120,000	6	455000	4380.006	120,000
Banking/Insurance	25	40,000	12	162500	8760.012	40,000
Telecom	15	30,000	12	97500	8760.012	30,000
Hotel/Tourism	10	11,000	9	65000	6570.009	11,000
Consultancy/Advisory	25	45,000	12	162500	8760.012	45,000
Manufacturing	20	30,000	12	130000	8760.012	30,000
E-commerce	5	12,000	6	32500	4380.006	12,000
Electricity/Gas	50	100,000	12	325000	8760.012	100,000
Healthcare	20	150,000	12	130000	8760.012	150,000
Consultancy/Advisory	30	55,000	6	195000	4380.006	55,000
Shipping/Logistics	16	35,000	12	104000	8760.012	35,000
Banking/Insurance	40	70,000	6	260000	4380.006	70,000
Banking/Insurance	34	62,000	6	221000	4380.006	62,000
Telecom	10	21,000	12	65000	8760.012	21,000
E-commerce	5	16,000	8	32500	5840.008	16,000
Hotel/Tourism	12	15,000	12	78000	8760.012	15,000
IT Services	15	20,000	6	97500	4380.006	20,000
IT Services	10	12,000	8	65000	5840.008	12,000
Healthcare	14	80,000	10	91000	7300.01	80,000
Banking/Insurance	30	60,000	7	195000	5110.007	60,000
Manufacturing	10	12,000	12	65000	8760.012	12,000
Banking/Insurance	42	70,000	9	273000	6570.009	70,000
Telecom	24	45,000	9	156000	6570.009	45,000
Healthcare	22	120,000	12	143000	8760.012	120,000
Shipping/Logistics	12	27,500	15	78000	10950.02	27,500
Banking/Insurance	85	140,000	9	552500	6570.009	140,000
E-commerce	10	25,000	6	65000	4380.006	25,000
Hotel/Tourism	12	14,000	9	78000	6570.009	14,000
Consultancy/Advisory	55	85,000	9	357500	6570.009	85,000
Telecom	30	57,000	9	195000	6570.009	57,000

Healthcare		50	200,000	12	325000	8760.012	200,000
Electricity/Gas		55	10,000	24	357500	17520.02	10,000
Shipping/Logistics		26	40,000	9	169000	6570.009	40,000
Banking/Insurance		12	20,000	6	78000	4380.006	20,000
Shipping/Logistics		330	880,000	12	2145000	8760.012	880,000
Hotel/Tourism		12	16,480	12	78000	8760.012	16,480
Manufacturing		15	32,000	18	97500	13140.02	32,000
Banking/Insurance		28	62,000	6	182000	4380.006	62,000
Telecom		20	22,000	6	130000	4380.006	22,000
IT Services		60	104,000	12	390000	8760.012	104,000
Banking/Insurance		30	53,000	6	195000	4380.006	53,000
Healthcare		9	270,000	12	58500	8760.012	270,000
Manufacturing		14	21,000	12	91000	8760.012	21,000
Banking/Insurance		20	74,000	6	130000	4380.006	74,000
IT Services		6	11,000	3	39000	2190.003	11,000
Healthcare		3	2,400	9	19500	6570.009	2,400
E-commerce		5	8,000	6	32500	4380.006	8,000
Telecom		30	80,000	12	195000	8760.012	80,000
Banking/Insurance		84	130,000	3	546000	2190.003	130,000
IT Services		15	55,000	6	97500	4380.006	55,000
Banking/Insurance		22	40,000	12	143000	8760.012	40,000
Telecom		12	25,000	12	78000	8760.012	25,000
Shipping/Logistics		160	325,000	15	1040000	10950.02	325,000
Consultancy/Advisory		31	35,000	12	201500	8760.012	35,000
Manufacturing		25	40,000	12	162500	8760.012	40,000
E-commerce		5	15,000	6	32500	4380.006	15,000
Electricity/Gas		40	90,000	12	260000	8760.012	90,000
Healthcare		25	140,000	12	162500	8760.012	140,000
Consultancy/Advisory		25	65,000	6	162500	4380.006	65,000
Shipping/Logistics		140	300,000	12	910000	8760.012	300,000
Hotel/Tourism		18	15,000	9	117000	6570.009	15,000
Banking/Insurance		50	67,000	6	325000	4380.006	67,000
Telecom		15	20,000	12	97500	8760.012	20,000
E-commerce		10	20,000	8	65000	5840.008	20,000
Hotel/tourism		5	12,000	12	32500	8760.012	12,000
Banking/Insurance		77	110,000	6	500500	4380.006	110,000
IT Services		15	15,000	8	97500	5840.008	15,000
Banking/Insurance		25	45,000	7	162500	5110.007	45,000
Manufacturing		15	20,000	12	97500	8760.012	20,000
Banking/Insurance		55	58,000	9	357500	6570.009	58,000
Telecom		30	35,000	9	195000	6570.009	35,000
Healthcare		30	170,000	12	195000	8760.012	170,000
Healthcare		10	100,000	10	65000	7300.01	100,000
Banking/Insurance		90	175,000	9	585000	6570.009	175,000
E-commerce		15	23,000	6	97500	4380.006	23,000
Hotel/Tourism		10	10,000	9	65000	6570.009	10,000
Consultancy/Advisory		40	78,000	9	260000	6570.009	78,000
IT Services		20	30,000	6	130000	4380.006	30,000
Banking/Insurance		13	19,000	12	84500	8760.012	19,000
Consultancy/Advisory		60	120,000	6	390000	4380.006	120,000

Appendix B

Table 2.

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
FTE	102	0	102	13000.000	4875000.000	333752.475	747663.011
HRB	102	0	102	1600.000	1450000.000	110684.752	228445.976
PB	102	0	102	2190.003	17520.024	7025.356	2719.391

Variable	Categories (n=10)	Observations (n=102)	% (100)	FTE Mean (hrs)	HRB Mean (hrs)	PB Mean (hrs)	Geometric Mean of total (hrs) saved
Industry	Banking/Insurance	24	23.762	267,943.9	65,175.1	5,388.5	**115,169.3**
	Consultancy/advisory	9	8.824	208,021.9	78,573.1	5,674.6	**97,423.2**
	Electricity/Gas	4	3.960	227,412.9	120,415.9	7,388.5	**84,558.8**
	Shipping/logistics	8	7.921	130,683.6	40,517.9	8,619.8	**59,940.4**
	Telecom	12	11.881	133,659.7	36,259.3	7,154.8	**59,036.6**
	Healthcare	12	11.881	74,226.8	84,139.6	8,099.9	**55,488.8**
	Manufacturing	8	7.921	91,923.9	21,094.4	9,694.6	**40,904.3**
	IT services	10	9.804	82,902.6	24,016.5	4,639.4	**37,186.2**
	Hotel/tourism	7	6.863	64,121.5	12,722.3	7,958.9	**28,267.6**
	E-commerce	8	7.921	34,015.6	14,801.7	4,706.6	**17,841.3**