BI Norwegian Business School - campus Oslo

GRA 19703

Master Thesis

Thesis Master of Science

The relationship between digitalization and profitability

A cross-sectional study of firms in the Norwegian shipping industry

Navn: Åshild Turmo Lornstad, Mari Rosvoll Viken

Start: 15.01.2021 09.00

Finish: 01.07.2021 12.00

The relationship between digitalization and profitability

A cross-sectional study of firms in the Norwegian shipping industry

Master of Science in Business

Åshild Turmo Lornstad

Mari Rosvoll Viken

Abstract

In this paper, we study the relationship between profitability and the level of digitalization of firms in the Norwegian shipping industry. We found the digitalization level by using a Likert scale on a survey sent to companies within this industry. Our main goal with this study was to investigate whether there is a correlation between profitability and the level of digitalization, and to identify whether this correlation is positive or negative. The subject was chosen based on the lacking research within this field.

In addition to using cross-sectional data on digitalization, we collected data on years since the digitalization started, ROA as a measurement of profitability, the age and the size of the company, all calculated from 2019. We retrieved 39 usable responses to our survey, which were used as the sample size.

The results of our regression models were inconclusive, meaning that we could not conclude with the null hypothesis or the alternative hypothesis. Because digitalization will only affect the operating part of companies, other drivers affected by the market will not be influenced. We believe our results are a consequence of not straining out the unaffected drivers, and not because digitalization has no impact on the profitability of firms in the Norwegian shipping industry.

Acknowledgement

Firstly, we would like to express a deep gratitude to our supervisor, Siv Staubo. Her knowledge and interest in the shipping industry gave us valuable information about the topic for our master thesis.

Furthermore, we would like to thank all the participants in our survey and those who contributed to spread the survey. The responses made it possible for us to write this thesis.

Lastly, we wish to extend a big thank you to our connections in Kongsberg Maritime. Their expertise gave us much cherished insight in the Norwegian shipping industry and valuable help when constructing the survey.

Thank you!

Table of Contents

Chapter 1: Introduction

1.1 The Research Question

Several industries around the world have digitalized their business in order to become more profitable. This can be done either by decreasing their costs, increasing customer satisfaction followed by revenues, or both. However, sometimes the industry is not ripe for this change. There are numerous examples of failed digital transformation processes in many different industries. Companies such as GE, Ford and Procter & Gamble, have all failed at digital transformation (Morgan, 2019). Although one industry is ready to be digitalized does not necessarily mean that other industries will succeed with digitalization.

The shipping industry has not been particularly forward-leaning in regard to digitalization. However, in the past few years, the industry has experienced an increased interest within this field. In our Master Thesis, we want to look into the relationship between digitalization in the Norwegian shipping industry and profitability. Are there any visible trends? We want to investigate whether the industry is ready for digitalization, or whether it is wise to let the industry ripe a bit longer.

Based on data collected on the level of digitalization from companies in the Norwegian shipping industry, we want to investigate whether there is a correlation between digitalization and profitability.

For this reason, we have chosen to articulate the following research question:

"Is there either a negative or positive correlation between digitalization and profitability in the Norwegian shipping industry?"

1.2 Background information

1.2.1 The shipping industry in Norway

Norway became one of the leading shipping nations in the second half of the 19th century (Grytten & Koilo, 2019), and is still one of the major maritime nations in the world (Reve, 2009). Over time, the Norwegian maritime industry has developed into one of the world's most complete maritime environments, consisting of businesses in all parts of the value chain and with strong positions in specialized segments (Regjeringen, 2020). The industry is among the largest and most important in Norway, with a total value creation near 175 billion and more than 110 thousand employees (Innovasjon Norge, 2020).

According to Torger Reve (2009) the unique combination between ocean, technology and knowledge are the main factors that have led the Norwegian shipping industry into its leading position. In addition, Norway has one of the highest cost-levels in the world (Reve, 2009). Compared to other nations within the industry, the Norwegian shipowners have to continuously develop smarter and more cost-efficient solutions to stay competitive. Because of this, we found it even more interesting to narrow our research down to the Norwegian shipping industry.

1.2.2 Digitalization in the shipping industry

The global shipping industry as a whole, appears to be quite conservative, in terms of digitalization compared to industries such as automotive and aviation. The slow innovative processes within the industry can be explained by the network of players connected both horizontally and vertically, with a great use of standards (Arduino, et al., 2013).

Moreover, since linear ships are characterized by network problems, a solid solution for the whole network is required when investing in new technology (Bavassano, Ferrari, & Alessio, 2020). As a consequence, the cost of investing increases exponentially. The combination of high investment costs and lack of evidence supporting payoff from digital investments, can explain why shipowners are resilient to digitalize (Jotun Marine, 2018). Another barrier is cyber security, as some believe that more technology increases the risk of hacking.

Nevertheless, there is an ongoing interest to digitalize the shipping industry, which is linked to the possibility of reducing costs and making transport services more efficient. Several of the new digital technologies that recently have been introduced within the industry are technologies such as Internet of Things (IoT), Artificial Intelligence (A.I.), data analytics and blockchain (Bavassano, Ferrari, & Alessio, 2020). The new digital technologies include satellites, telematics and management systems, which provides the opportunity for the shore-based staff to influence the ship in the daily operations. In addition, digitalization helps to gather, process and make information available, which is necessary for better decision making (Splash 24/7, 2017).

Shipping Analyst Martin Stopford believe that shipping companies could reap great benefits from implementing new digital technology. However, Stopford argues that this implementation will take time, because the business model is not ready for new technology yet and must therefore be changed first. Changing the business model include implementing smart ships, introducing smart fleets with integrated management systems, and implementing smart global logistics (Jotun Marine, 2018). We find Stopford’s statement interesting, but because of the lacking research, we will not use this as a base in our study.

1.3 Key drivers

There are several key drivers in the shipping industry which will affect the cash flow and the financial results of a company. We will now consider some of the most important key drivers within the shipping industry.

1.3.1 Demand and supply

Many companies within the shipping industry either sell or buy ships, or both, at a regular basis. Since it is a competitive market, the price in this business depends on demand and supply.

The demand is influenced by shipping freight rates, second-hand prices, market expectations and sentiment, and liquidity and credit availability, while the supply is influenced by available shipbuilding berths, shipyard unit costs, exchange rates and production subsidies (Stopford, 2009, p. 631).

There are several factors that affects the demand and supply, and therefore the price in this market. A high market price when buying ships would lead to higher costs, meanwhile a low market price when selling ships leads to lowered income. The effect of the market price on the financial results may therefore be decisive.

1.3.2 Exchange rates

The shipping industry is a highly international industry. As earlier discussed, exchange rates will affect the supply of ships, and could in addition affect a company's financial results directly.

Consider a company who are building ships in Norway and selling to customers worldwide. The majority of the expenses will therefore be in domestic currency, while the main income will be in a foreign currency. Although the exchange rate does not affect the expenses, it can be crucial for the size of the income and the present value of that specific sale transaction.

1.3.3 Freight rates

Freight rates is a highly important driver for shipowners since high freight rates will directly lead to increased earnings. Moreover, freight rates will increase the demand for new ships, which is easily explained by the fact that shipowners want to expand the size of their fleets while the business is more profitable (Stopford, 2009, p. 631)

1.3.4 The cost of running ships

While freight rates are important for the income, the cost of running ships is important for the expenses. These costs can be divided into five categories, which is operating costs, periodic maintenance, voyage costs, cargo-handling costs and capital structure (Stopford, 2009, p. 225). While one single company's digitalization will not affect the categories above, it will perhaps affect the cost of running ships, and therefore have an effect on the total profitability. Based on this assumption, we identify this driver as the most important in light of our research question.

1.4 Our motivation and contribution

In the early stages of this thesis, we contacted Kongsberg Maritime and their office at Grilstad, Trondheim in Norway. They told us about their current process of converting manual operations into digitalized procedures and the large costs connected to such a transformation. This made us think about the profitability attached to these investments. Knowing that several industries have failed with investments in digitalization, we wanted to investigate whether the shipping industry in Norway profits from such investments or not.

According to Ferreira et. al (2019), the impact of digitalization on firm performance is a largely unexplored topic. In addition, there is no research looking at digitalization and profitability in the Norwegian shipping industry. Our contribution is therefore a unique study which we know, after talking with Kongsberg Maritime, the industry is interested in.

Chapter 2: Theoretical Methodology

2.1 Choice of Subject

Due to the gap in research within the shipping industry, we found it interesting to study the impact digitalization has on profitability in this specific industry. We also saw that Norway have been forced to be more innovative than the other nations within the industry, and therefore we narrowed our research down to the Norwegian shipping industry.

2.2 Preconceptions

As aforementioned, the shipping industry is a conservative industry. In addition, it is said to be highly volatile, which makes it difficult to predict the results from our findings. Since this is an unexplored topic, we have no preconceptions towards this analysis.

2.3 Methodology

Digitalization is not numerical and public information; therefore, we will collect these data using a quantitative method. A survey sent to several companies within the shipping industry will give us the information needed to calculate the digitalization level of each company.

Data collected from both the survey and the data source proff.no, will be the input to our research. From proff.no will we find data from each of the companies that have answered our survey. The collected data will be inserted in Excel to get a good view of the numbers and calculations. Then, we will use MATLAB to perform both regression analysis and hypothesis testing.

Chapter 3: Literature Review and Theory

3.1 Previous Research

A previous study on digitalization within the banking sector had results indicating the probability not being affected by the sheer level of the digitalization in a bank. The data used in the study was collected through a survey answered by 102 banks in Germany, Switzerland and Luxembourg. The level of digitalization was measured in their survey by asking the respondents about the degree of digitalization within the company. The other questions were based on a five-point Likert scale. As they mention, their measured level of digitalization suffers from the respondent's self-interpretation to which level a business model is digitalized. For further studies, they recommend including the types of digitalization tools used and the different departments' use of those (Niemand et. al., 2020).

However, a rapport by Bank of Finland state that digitalization of banks in the long run expects to increase profitability. They argue that measures to improve profitability in the long run, such as investments in digitalization, will in the short run decrease the profitability. This is caused by investments connected to digital solutions demands a lot of resources and capital (Koskinen & Manninen, 2019).

Another relevant study examines the impact digitalization has on profitability in small- and medium-size private healthcare companies in Finland. The data was obtained from a national survey in Finland where 680 private healthcare companies responded. Studies prior to this indicated that digitalization impacted profitability both positively and negatively. This study used two different performance measurements, EBIT and ROA. The results from EBIT as measurement, indicated that digitalization negatively affected the company's profitability, especially for small companies. Meanwhile, when performance was measured with ROA, the digitalization in innovation processes had a positive effect on profitability. The study concluded with the company size and business area being decisive for how digitalization impacts profitability. The study also found that the relation was impacted by the chosen profitability indicator (Holopainen, Niskanen, & Rissanen, 2019).

3.2 Literature Review

There are many types of literature reviews. As examples, we have narrative, systematic, meta-analysis and meta-synthesis (O'Gorman & MacIntosh, 2015, p. 31). In our study, we have used both narrative literature review and systematic literature review.

3.2.1 Narrative Literature Review

The purpose of this type of literature review is to analyze and summarize a body of literature. That is *"achieved by presenting a comprehensive background of the literature within the interested topic to highlight new research steams, identify gaps or recognize inconsistencies."* (O'Gorman & MacIntosh, 2015, p. 31). After we decided about the topic we wanted for our thesis, we saw that there was no existing research on digitalization and firm profitability in the shipping industry in Norway. Here, we identified a gap in previous research, and therefore decided to take a deeper look at the subject.

3.2.2 Systematic Literature Review

A systematic literature review can be defined as the attempt *"to identify, appraise and synthesize all the empirical evidence that meets pre-specified criteria to answer a given research question"* (Piper, 2013, p. 2). Since we found a gap in the literature we wanted to fill, we had to look at similar research. We found many studies which looked at the relationship between profitability and an independent variable. This was highly useful for us, both in looking at how other studies had defined profitability, but also which control variables we should include in our models.

3.4 Theories

We chose to research the relationship between digitalization and profitability within the shipping industry in Norway. The industry is chosen due to various reasons mentioned earlier. Results from previous studies in other sectors indicated both positive and negative impact from digitalization on profitability, although this relation has not been studied within the shipping industry.

Chapter 4: Research Methodology

4.1 Data Selection

4.1.1 Cross-sectional study

Before starting to collect data, we chose which data to include in our research. The first thing we took into consideration, was the time aspect. With this in mind, there are several different study designs to choose.

Times series is such a study design. This type of research requires that every individual is observed at several different occasions (Lantz, 2014). For our research, it would be useful to look into the relationship between the different variables at years before 2019, but this would either require a lot of time-consuming follow ups, or a highly advanced survey.

The opposite of a time series study is a cross-sectional study, where each individual is only observed once (Lantz, 2014). This type of study has a one-time point of view and provides a ‘snapshot’ of the situation (Levin, 2006). When using a cross-sectional study, we would not lose any data due to missing follow-up, and the survey would not be very time-consuming for the participants. Based on this, we believed that a cross-sectional study was the best fit for our research.

When using a cross-sectional study design, there follows some disadvantages. The main disadvantage we must take into consideration when discussing our results, is the possibility of a different result if we were to choose another timeframe (Levin, 2006).

4.1.2 Industry and Geographical Delimitations

Since we found it interesting to only look at the Norwegian shipping industry, we delaminated our search to Norwegian shipping companies. Within Norway, we had no other geographical delimitations.

In order to obtain an overview of the different shipping companies in Norway, we chose to focus on the companies within the industry “Shipping and sea transport” at proff.no. In addition, we added some companies who had “rederi” within their name.

4.1.3 Research on digitalization level

By looking into how previous studies measure the level of digitalization, we could not find a common way to do this. There were few previous studies about relating subjects, and none of them calculated the digitalization level as we intend to do. However, common for most of the studies is the use of the Likert scale from 1 to 5, (1: “does not fit at all”, 5: “fits perfectly”), in their surveys when collecting information about how digitalized firms are. One of the studies assessed the level of digitalization as a singular construct developed for that study. In the survey, the respondents were asked about the concrete level of digitalization in percentage, in addition, a Likert scale was used for the rest of the questions (Niemand et. al., 2020).

4.1.4 Profitability

Firstly, we wanted our measure of profitability to be comparable between the different companies in our study. We therefore decided that a ratio of profitability would be best suited. Thereafter, we had to choose one specific ratio which we would include in our model.

There are several different ratios for measuring profitability. By looking at relevant previous research we found that Hamid, Abdullah, & Kamaruzzaman

(2015) used return on equity (ROE) in their research on profitability and capital structure in family and non-family firms. Dogan (2013) chose to use return on assets (ROA) when studying firm size and profitability, and so did Malik (2011) and Andersson & Minnema (2018) in their studies on profitability.

Since ROA is a ratio that measures the performance by comparing the net income to the capital that is invested in assets, it also measures how productive and efficient the management is in their use of economic resources (Corporate Finance Institute, 2020). On the basis of this, we feel that Return on Assets is the most suitable ratio of profitability in our study.

4.1.5 Number of observations

We want to collect enough data such that the random sample from the target population, which in this case is all of the shipping firms in Norway, would be generalized. Since it can be difficult to know how many observations needed to conclude with reliability, we looked at other research and their number of observations.

Niemand et. al. (2020, p. 5) received 102 responses to their survey, which constituted their final sample. Dong and Su (2010) used a sample based on 130 firms, Salvi et. al. (2021)'s sample consisted of 114 companies, while Ibem et. al. (2018) used a sample consisting of 75 firms.

Based on previous research, we hoped to receive 100 responses to our survey, which will form our sample in the regression model.

4.1.6 Number of Dispatch

Niemand et. al. (2020, p. 5) sent their survey to 300 addresses, retrieving a response rate of 34%. Ibem et al. (2018) sent their survey to 129 firms and received usable answers from 75 of the questionnaires. This represents a response rate of 58%.

Yehuda Baruch (1999) has done a study about the reasonable response rate in academic studies. In 1995 the average response rate was 48.4%, and the author found that the response rate was declining through the years (from 55.6% average from 1975-1995) (Baruch, 1999).

Based on this, we hope to retrieve a response rate at 35%. Given our wishes for 100 responses, we would send the survey to at least 285 firms in the Norwegian shipping industry. This is equal to approximately 11% of the total number of businesses registered within shipping in Norway, as the total number is 2478 firms.

4.2 Data Collection

4.2.1 Data on digitalization

A big challenge attached to our study is how to collect data on digitalization properly. Different approaches were considered, such as looking at the company's investments in IT solutions, interviews and surveys. In order to collect data efficiently and as accurate as we saw possible, we decided to make a survey asking about their digitalization in different areas of their operations.

4.2.2 Survey

Our survey is developed in collaboration with Kongsberg Maritime to ensure the quality. Previous studies have asked the respondents to specify their digitalization level, as mentioned in these studies this is a question of interpretation. We want to make the responses more valid and remove some of the self-interpretation. By following the recommendation from the study by Niemand et. al (2019) we therefore included both digitalization tools and areas of operations in our survey. An overview of the structure of our survey is presented in the table below.

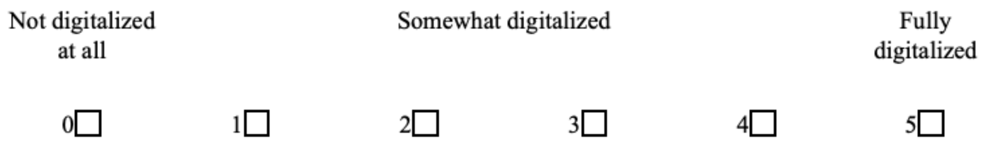
Table 1: The structure of our survey

Table 1 presents an overview of the structure of our survey. The areas of operations were divided into three main areas with three subareas each. The main area is presented in the first column, the subarea is presented in the second column.

Main area	**Subarea**
Personnel	
	Crew Management
	Crew Training & Welfare
	Document Handling
Fuel & Performance Management	
	Fuel Performance
	Fleet Tracking
	Voyage & Navigation
Maintenance & Operations	
	Maintenance
	Port Operations
	Logistics & Procurement

Under each subarea, the participants of the survey were asked to answer a scale from 1 to 5, with the following explanation of the numbers:

Figure 1: *The five-point Likert scale used in the survey. The respondents were asked to rate their level of digitalization from 0-5 under each subarea.*



To reduce the self-interpretation even more, we chose to write some examples under each subarea. Explanations for the subarea Voyage & Navigation can be observed below:

Figure 2: *A specific example of the five-point Likert scale. This example is taken from the main area Fuel & Performance Management and the subarea Voyage & Navigation.*

c. Voyage & Navigation: 0☐ 1☐ 2☐ 3☐ 4☐ 5☐

0: Minimal navigation
2: Navigation solutions without digital route planning
3: Navigation solutions with digital route planning
5: E-navigation solutions with; weather routing, emergency navigation, passage planning, automatically routing from port-to-port

At the end of the survey, the following question was asked in order to map when digitalization in the different companies began:

"If you have started digitalizing certain processes in your business, in what year did this digitalization start?"

4.2.3 Net Income, Total Assets, Company Age and Company Size

For our further calculations, we collected data for net income, total assets, company age and company size from the companies who answered our survey. We retrieved this data from the accounting figures published at proff.no.

4.3 Data Processing

4.3.1 Data Handling

To ensure that the collected data do not violate the Privacy Act and GDPR, we used Qualtrics, which has a collaboration with BI Norwegian Business School. In addition, our survey did not ask about the respondent's name or positions in the company. In accordance with NSD, our survey did not collect IP addresses, and therefore we do not handle personal information.

4.3.2 Calculating the Level of Digitalization

The level of digitalization will have a value between 0 and 5, depending on the answers from the survey. The three different main areas in our survey are weighted equally. Meanwhile, the subareas are weighted equally within the main area. Therefore, all of the main areas received a weight equal to $\frac{1}{3}$, giving each subarea a weight equal to $\frac{1}{9}$.

Some of the areas are not relevant for all of the companies. If they do not handle business in one or more areas, they do not have to respond to that category. The unanswered areas will not be included in the calculation of the total level of digitalization.

4.4 Variables

4.4.1 Return on Assets (ROA)

$$ROA = \frac{Net\ Income}{Total\ Assets} \cdot 100$$

4.4.2 Digitalization

The level of digitalization will have a value between 0 and 5, depending on the answers from the survey. Since each subarea is equally weighted, the level of digitalization will be the mean value of the answered areas.

4.5 Model

4.5.1 Multiple Linear Regression Model

A multiple regression examines the correlation between several independent variables and one dependent variable. The regression looks at the degree to which each independent variable predicts the dependent variable (Ross & Wilson, 2017, p. 49).

In order for the estimators to be unbiased for the population parameters, a set of assumption for the multiple linear regression model must be met (Wooldridge, 2015, p. 73). The assumptions are used to obtain exact statistical inference and to conclude that the OLS estimators have the smallest variance among all unbiased estimators (Wooldridge, 2015, p. xiv).

Wooldridge (2015) have called these assumption MLR.1 to MLR.6, where MLR.1 and MLR.2 says that the model must be linear in parameters and that the sample must be random with n observations, respectively. The third assumption state no perfect collinearity and the fourth require that the error term has an expected value of zero and a zero population mean. MLR.5 says that the model cannot experience heteroskedasticity, while the last assumption requires the error term to be normally distributed (Wooldridge, 2015).

4.5.2 Dependent variable

In our study, we want to examine how the level of digitalization affect the profitability. Therefore, profitability will be the dependent variable in our multiple regression model. As mentioned, we will measure profitability as return on assets.

4.5.3 Independent Variables

The main independent variable in this study is the level of digitalization, as it is the variable directly linked to our research question. We also want to add some control variables. These will be the years since digitalization began, the age of the company and the size of the company.

In our survey, we asked in which year they started to digitalize. We believe this is a highly relevant variable to include in our model, since previous research have shown that the investment can have a negative correlation with the profitability in the first years, and thereafter a positive correlation (Koskinen & Manninen, 2019).

Based on previous research, we believe that the age of the company is an important control variable. A study conducted by Loderer and Waelchli (2010, pp. 32-33) found a highly significant negative relation between company age and profitability, while Haykir and Çelik (2018, p. 135) found a convex relation between these two measures. In our study, the age will be measured as the number of years since the company was founded.

The same applies for the size of the company: studies show inconsistent conclusions. Hall and Weiss (1967) found that size tends to result in high profit rates. Niresh and Velnampy (2014, p. 63) found a weak positive relation between profitability and company size, and so did Babalola (2013, p. 92). In our study, the size of a company will be measured as the number of employees in the company.

To be certain that we do not add any irrelevant control variables, we will include five different models further in this research where we eliminate one or more of the independent variables which are not directly linked to our research question.

4.5.4 Models

Model 1:

$$ROA_{i,2019} = \beta_0 + \beta_1 LoD_i + \beta_2 Y_{i,2019} + \beta_3 A_{i,2019} + \beta_4 S_{2019} + \varepsilon_i$$

Model 2:

$$ROA_{i,2019} = \beta_0 + \beta_1 LoD_i + \beta_2 Y_{i,2019} + \beta_3 S_{2019} + \varepsilon_i$$

Model 3:

$$ROA_{i,2019} = \beta_0 + \beta_1 LoD_i + \beta_2 Y_{i,2019} + \beta_3 A_{i,2019} + \varepsilon_i$$

Model 4:

$$ROA_{i,2019} = \beta_0 + \beta_1 LoD_i + \beta_2 Y_{i,2019} + \varepsilon_i$$

Model 5:

$$ROA_{i,2019} = \beta_0 + \beta_1 LoD_i + \varepsilon_i$$

Where:

$ROA_{2019} = Return\ on\ Assets\ in\ 2019$

$LoD = Level\ of\ Digitalization$

$Y_{2019} = years\ since\ the\ digitalization\ started\ in\ 2019$

$A_{2019} = Company\ Age\ in\ 2019$

$S_{2019} = Size\ of\ the\ company\ in\ 2019$

Here, we can observe that our models meet assumption MLR.1. Our analysis has 39 random observations retrieved from the survey, which meets assumption MLR.2. As aforementioned, we wanted to retrieve over 100 observations based on calculations from previous similar research, which we were unable to collect.

If a sample size is not very large, then the t distribution can be a poor approximation to the distribution of the t-statistics when the error term is not normally distributed. Although, how large a sample size must be for the approximation to be good enough is not defined. Some econometricians state that a sample with 30 random observation is satisfactory, depending on the distribution of u (Wooldridge, 2015, p. 157).

4.5.5 Simple Linear Regression Model

Above, we can see that Model 5 only looks into the relationship between two variables, namely the profitability and the level of digitalization. This is called a simple linear regression model (Wooldridge, 2015, p. 20). A simple linear regression model needs to meet the same assumptions as the multiple linear regression model, except from the third assumption. SLR.3 states that there needs to be a sample variation in the explanatory variable (Wooldridge, 2015, p. 42).

4.5.6 Robust Linear Regression

Another linear regression model one can use, is something called a robust linear regression. This is an approach which can be useful when looking at violations of the assumptions above, such as outliers and nonnormal distributions of errors (Freund, Wilson, & Sa, 2006, p. 156).

The method of robust estimation is called the iteratively reweighted least square (IWLS) procedure. This procedure will attempt to reduce the influence of observations with large residuals (Freund, Wilson, & Sa, 2006, p. 157).

If we discover that one or more of the assumptions above is violated, we will try to use a robust linear regression to see if that gives us a better fit to the data.

4.6 Hypotheses & Hypothesis Testing

4.6.1 Hypothesis Testing

Since we would like to test a single parameter in the multiple regression model, we have to perform a t-test. The t-test indicates whether the level of digitalization is statistically associated with the profitability (ROA) or not. Under the test of significance approach, the null hypothesis will not be rejected if the test statistic lies within the non-rejection area (Brooks, 2014, p. 106).

The non-rejection area, or the critical values, will be calculated in MatLab using $T = 39$ for all five models, $k = 5$ for Model 1, $k = 4$ for Model 2 and Model 3, $k = 3$ for Model 4 and $k = 2$ for Model 5, with a suitable level of significance.

4.6.2 The Hypotheses

In the hypothesis testing framework, there are always two hypotheses, specifically the null hypothesis and the alternative hypothesis, where the null hypothesis is the one actually being tested while the alternative hypothesis represents the remaining outcomes (Brooks, 2014, p. 99).

For our research question, we have the following hypotheses, equal for all of the five models:

The Null Hypothesis H_0: There is no correlation between ROA and the level of digitalization. $H_0: \beta_1 = 0$

The Alternative Hypothesis H_A: There is either a negative or a positive correlation between ROA and the level of digitalization. $H_A: \beta_1 \neq 0$

Since there is no previous research regarding the relationship between digitalization and profitability in the Norwegian shipping industry, a two-sided test is the most suited test for our study.

4.6.2 Significance Level

We have to consider the suitable level of significance thoroughly, since this could be of high impact for the conclusion of our study. It is also important to be impartial and objective, especially since this choice have shown to be arbitrary and depending on the desire of an investigator to reject or accept a hypothesis (Keuzenkamp & Magnus, 1995, p. 20).

It is also wise to consider both Type I and Type II errors. Type I error refers to the situation where you reject the null hypothesis when it is true, while Type II error refer to the situation where you do not reject the null hypothesis when it is false (Brooks, 2014, p. 110).

If the Norwegian shipping industry finds this study interesting, the consequence for a Type I error is that digitalization seems more profitable than it really is. On the other hand, the consequence of a Type II error is that digitalization appears to be undesirable when it actually has a positive effect on the profitability. These consequences are also important to take into consideration when choosing a suitable level of significance. We will consider Type I and Type II errors to be equally important, and the expected losses from the two errors to be the same.

For our test, we want the power to be as high as possible. Studies show that when the expected loss of Type I and Type II errors are identical, one can achieve a balance between the probabilities of committing the two types of errors. Balancing those, one can get a high power of the test while taking the sample size into consideration (Kim, 2015, p. 3). Jae Kim (2015) presents the following table:

Table 2: Level of significance

Table 2 shows the level of significance when taking the sample size into consideration. The first column presents the sample size, the second and third column presents, respectively, the probability for Type I error (α) and the probability for Type II error (β). Meanwhile, the fourth and last column presents the power of the test. The table is prepared by Jae Kim (2015 p. 11).

Sample size	α	β	Power of the test
10	0.35	0.35	0.65
50	0.19	0.19	0.81
100	0.11	0.11	0.89
200	0.04	0.04	0.96

In comparison, if the sample size is 50 and the level of significance (α) is set to be 5%, the probability of a Type II error is 45% resulting in a power of the test of only 55% (Kim, 2015, p. 11).

Given our sample size of 39 observations, we are closest to 50 following the table above. We therefore set our significance level to 19%, giving our test a power equal to 81%.

4.7 Multicollinearity

In any practical context, the correlation between the independent variables will be non-zero and there will be a small degree of association between the independent variables. When this association, or correlation, becomes very high, we call it multicollinearity (Brooks, 2014, p. 217). Multicollinearity is a problem because it undermines the statistical significance of the explanatory variables (Allen, 1997, p. 176).

One way to see if it exists is to look at the matrix of correlations between the individual independent variables. Multicollinearity would be visible as a high correlation between two of the explanatory variables (Brooks, 2014, p. 218). Given our models, we would need to look at the correlation matrix of the level of digitalization, company age, company size and the number of years since the digitalization started.

In addition, we can detect multicollinearity by using Variance Inflation Factors, a so-called VIF-test. As an example, the VIF for slope coefficient 1 is

$$VIF_1 = \frac{1}{1 - R_1^2}$$

, which is the term in $Var(\hat{\beta}_1)$ that is determined by correlation between the level of digitalization (LoD) and the other independent variables. As a guideline, the value 10 is often chosen as a limit (Wooldridge, 2015, p. 86).

No multicollinearity will mean no perfect collinearity, which will make our sample meet assumption MLR.3. On the other hand, discovering multicollinearity does not mean perfect collinearity. If this is the case, we need to perform a test to decide perfect collinearity or multicollinearity. Since the only difference between our five models is that we exclude one or more control variables, we only need to do this once – where all of the variables are included.

The solution to the problem will vary, depending on the impact multicollinearity has on the outcome of the model. This can either be to ignore it, dropping one of the collinear variables or transforming the highly correlated variables into a ratio (Brooks, 2014, p. 219). If multicollinearity occurs in either size or age of the company, the simplest solution would be to eliminate the model consisting both of the variables.

4.8 Omission of an important variable

A problem that easily can occur, is omission of an important variable. In our original model we have chosen to include the level of digitalization, years since digitalization began, company age and company size. There are several more variables which can explain the differences in the profitability. The consequence of excluding an important independent variable would be that the estimated coefficients on all the other variables will be biased and inconsistent unless the excluded variable is uncorrelated with all of the other explanatory variables (Brooks, 2014, p. 225). If this were to be the case, the constant term would be biased. Dealing with this issue, it can be useful to look at the adjusted R-squared.

Since our dependent variable is profitability, more specific ROA, it is likely that we exclude important variables. This is based on the fact that there are a lot of different factors which make an impact on the profitability. We therefore expect the R-squared and the adjusted R-squared to be low.

In our study, we want to examine the effect digitalization has on profitability. As long as the excluded variables do not correlate with the other variables in our models, omission of an important variable would only affect the constant term, and therefore not affect the conclusion of the research question.

4.9 Zero conditional mean and normality

Assumption MLR.4 and MLR.6 says that the error terms must have zero population mean and be normally distributed. The way this can be tested is by plotting the error terms in a histogram. If the mean is centered around zero, and the histogram does not experience any skewness, these assumptions are met (Andersson & Minnema, 2018, p. 33).

4.10 Homoskedasticity

Since we have collected data from firms that differ in size, there is a possibility that the larger firms have more factors affecting the error terms, than the smaller ones. Therefore, the error terms will be larger (Newbold, Carlson, & Thorne, 2013, p. 578), and heteroskedasticity can occur. A consequence of heteroskedasticity being present in our data is that our analysis may not be valid. Therefore, it is important to check if our model meet assumption MLR.5.

One way to do this is to plot the residuals versus the independent variable and the predicted values from the regression. If the plot does not show any systematic relationship between the errors and the independent variable, there is no evidence of nonuniform variance (Newbold, Carlson, & Thorne, 2013, p. 578).

We can also check for autoregressive conditional heteroskedasticity (ARCH) disturbance in our model introduced by Engle. To test for this disturbance, it is recommended to use the Lagrange multiplier test (Lee, 1991), which is the test we will conduct in addition to the residual plot.

4.11 Sample Variation in the Explanatory Variable

Assumption SLR.3 for Model 5 can easily be tested for by calculating the sample standard deviation of the level of digitalization. As long as it is not zero, assumption SLR.3 holds.

Chapter 5: Empirical Results & Analysis

5.1 Descriptive Statistics

Table 3: Descriptive Statistics

Table 3 shows the descriptive statistics of our data collected from proff.no and the survey sent to companies within the Norwegian shipping industry. It contains data from 39 observations, including ROA, the level of digitalization, years since the digitalization started, the age and size of the company, all variables calculated from 2019. The descriptive statistics shown in this table is the mean value, the standard deviation, the minimum value and the maximum value.

Variable	Mean	Std. Dev.	Min	Max
ROA_{2019}	0.010858	0.26164	-0.94943	0.60185
LoD	2.7435	0.72142	1.2222	4.25
Y	6.1282	7.0977	0	33
A	21.179	20.36	1	107
S	59.949	78.812	2	328

From table 3, we observe an average ROA for the companies in our data equal to 0.01858, or approximately 1.086%. There are a lot of fluctuations between the different companies, which appears in the standard deviation, the minimum value and the maximum value. Earlier, we explained that the shipping industry is a volatile industry, but looking at this table, it looks like the profitability depends on much more than the market. These companies are doing business in the same market, and the numbers are retrieved from the same year (2019). Here, we can also observe that assumption SLR.3 for Model 5 holds.

As aforementioned in chapter 4.10, it is important to test our data for heteroskedasticity. Since a wider range between the firms in a dataset could mean a larger possibility for heteroskedasticity, it is especially important in our study. From the range in company age and size in the table, we can observe huge differences across our data.

Previously we discussed that the shipping industry is a conservative industry. The companies in our study hold an average value of 2.74 looking at the level of digitalization. Since we do not have the same data on the level of digitalization from firms in other industries, we cannot compare this number with anything outside of this paper.

5.2 Test for Multicollinearity

As explained in chapter 4.7, we can test for multicollinearity by looking at the matrix of correlations between the individual independent variables, in addition to conducting a VIF-test.

Table 4: Correlation Matrix

Table 4 shows the correlation between the level of digitalization, years since the digitalization started, the age of the company and the size of the company. The correlation matrix is based on 39 observations, which were the respondents to our survey. The table shows the correlation between the variable on the left-hand side and the corresponding variable on the top row.

	LoD	**Y**	**A**	**S**
LoD	1			
Y	0.2360	1		
A	-0.0265	0.1940	1	
S	-0.1201	0.1896	0.0561	1

The table above shows the output for the correlation matrix after calculating this in MatLab. Here, we can observe the highest correlation between the level of digitalization and years since the digitalization started, which is 0.2360. This is not considered a high correlation, indicating that multicollinearity is not present in our data.

Table 5: VIF-test

Table 5 shows the Variance Inflation Factors of the level of digitalization, years since the digitalization started, the age of the company and the size of the company. The VIF-test is performed to detect multicollinearity in the data and the limit value used is 10. The first row presents the variable, and the second row presents the Variance Inflation Factor.

Variable	**LoD**	**Y**	**A**	**S**
VIF	1.0975	1.1622	1.0452	1.0693

From table 5, we observe that the variance inflation factors all have values under our chosen limit, and by a good margin.

Both of the methods for detecting multicollinearity indicate multicollinearity not to be an issue in our model, and we are therefore confident when stating that multicollinearity is not present in our data. This also indicates that our sample meet assumption MLR.3.

5.3 Residuals

For assumption MLR.4 and MLR.6 to be met, the error terms must have a zero population mean and be normally distributed. Firstly, we calculated the mean value of the residuals to be the following:

Table 6: Residuals population mean

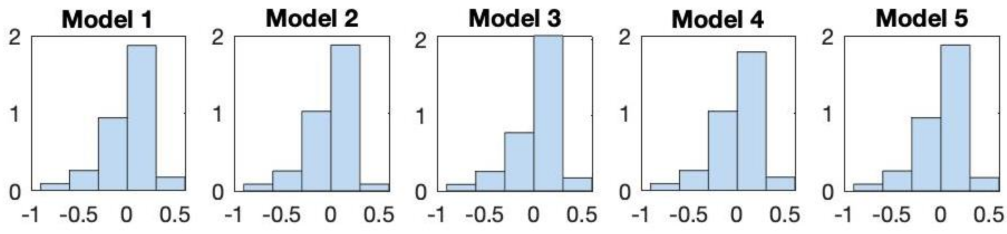
This table shows the residuals population mean for model 1 to 5. The first column presents the model, and the second column presents the residuals population mean. Model 1 are presented in row 2, going down to Model 5 presented in row 6.

	Residuals population mean
Model 1	$-2.9179e^{-17}$
Model 2	$-4.3413e^{-17}$
Model 3	$-5.2664e^{-17}$
Model 4	$2.1250e^{-17}$
Model 5	$8.9672e^{-17}$

From table 6, we can observe that the population mean of the residuals of all five models are significantly close to zero, meaning that assumption MLR.4 is met for all of our models.

To check for normality, we plotted the residuals in five histograms.

Figure 3: *Histogram of the residuals. From the left: Model 1, Model 2, Model 3, Model 4 and Model 5.*



Here, we can observe that neither of the histograms looks like its normally distributed. Therefore, we decided to conduct an additional test to check for this. We conducted a one-sample Kolmogorov-Smirnov Test. The Kolmogorov-Smirnov test, which is a goodness-of-fit test, test whether the residuals could reasonably have come from a normal distribution (IBM, 2021).

The Kolmogorov-Smirnov test rejected the null hypothesis for all of the models, which claims that the data comes from a standard normal distribution. In other words, our data does not meet assumption MLR.6. Therefore, we will conduct a robust linear regression to see if this is a better fit.

5.4 Test for heteroskedasticity

As mentioned in chapter 4.10, we have to check whether heteroskedasticity is present in our models or not. Firstly, we plotted the residuals vs. the fitted values in our models.

Figure 4: *Plot of residuals vs. the fitted values in our models. From the left: Model 1, Model 2 and Model 3.*

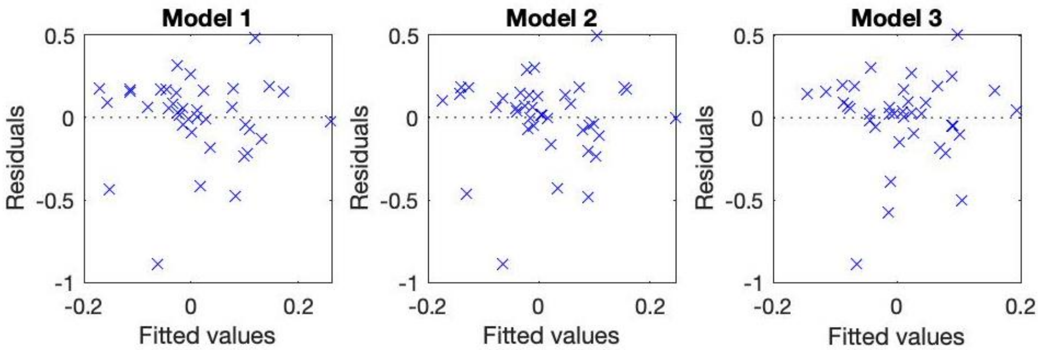
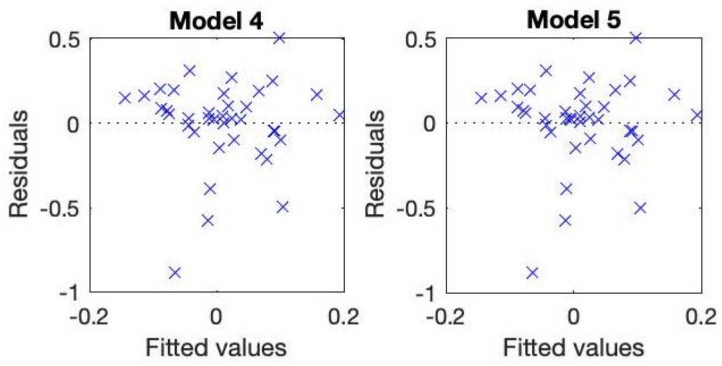


Figure 5: *Plot of residuals vs. the fitted values in our models. From the left: Model 4 and Model 5.*



Here, there is no obvious systematic relationship, which implies no heteroskedasticity. As an additional test, we performed Lagrange multiplier tests to check for ARCH disturbance.

The result of our tests indicate that we should not reject the null hypothesis of no conditional heteroskedasticity, and we therefore conclude that our sample meets assumption MLR.5.

5.5 Regression Models

In the tables below we can observe the results for our multiple linear regressions from data containing 39 observations. The coefficient of each independent variable addresses the percentage change in ROA for an increase of 1 unit of the explanatory variable, given that the other independent variables are set.

5.5.1 Regression Model 1

Table 7: Model 1

Table 7 shows the results of the multiple linear regression from Model 1, containing 39 observations. This table shows the coefficient, the standard error, the t-statistics and the p-value of the intercept, the level of digitalization, years since the digitalization started, the age of the company and the size of the company. In addition, the R-squared and the adjusted R-squared for Model 1 are presented.

	Coefficient	**Std. Err.**	***t*-stat**	***p*-value**
Intercept	−0.1467	0.1795	−0.8172	0.4195
LoD	0.0734	0.0608	1.2085	0.2352
Y	0.0047	0.0064	0.7379	0.4657
A	−0.0013	0.0021	−0.6119	0.5447
S	−0.0008	0.0005	−1.3801	0.1766
R^2	0.13			
Adjusted R^2	0.0279			

Model 1 suggests that both the age and size of a company in the Norwegian shipping industry will have a negative impact on profitability. On the other hand, the level of digitalization and years since digitalization started will have a positive impact on profitability, looking at the coefficients and not taking the p-value into account.

Using this model, one can say that the level of digitalization has a positive relationship with profitability if the p-value was as high as 23.5%. Since this is above our chosen significance level, this regression finds no significant relationship between profitability and the level of digitalization. Given a significance level of 19%, the only relationship this regression reveals are that size has a minor negative impact on profitability.

In addition, we can observe that the R-squared is 0.13 and the adjusted R-squared is 0.028. As expected, these are quite low. If we had included more control variables of relevance, this number could have gotten higher, but with our research question in mind, we decided to not include too many control variables.

5.5.2 Regression Model 2

Table 8: Model 2

Table 8 shows the results of the multiple linear regression from Model 2, containing 39 observations. This table shows the coefficient, the standard error, the t-statistics and the p-value of the intercept, the level of digitalization, years since the digitalization started and the size of the company. In addition, the R-squared and the adjusted R-squared for Model 2 are presented.

	Coefficient	Std. Err.	t-stat	p-value
Intercept	-0.1765	0.1712	-1.0311	0.3096
$\boldsymbol{LoD}$	0.0762	0.0601	1.2682	0.2131
$\boldsymbol{Y}$	0.0039	0.0062	0.6341	0.5301
$\boldsymbol{S}$	-0.0008	0.0005	-1.3971	0.1712

$\boldsymbol{R^2}$	0.121			
Adjusted $\boldsymbol{R^2}$	0.0453			

The results for Model 2 shows that the p-value for LoD has decreased slightly compared to Model 1. The significance level has to be 21.3% for the relationship between profitability and the level of digitalization to be statistically significant. Similar to Model 1, this is above our chosen level. In other words, this regression finds no significant relationship between the level of digitalization and profitability. As for Model 1, the only significant variable at our level is the size of the company.

For Model 2 where age of the company is excluded, the R-squared are somewhat lower than for Model 1. Since models including more variables often have a higher R-squared, this is as expected. An interesting thing with Model 2, is that the adjusted R-squared are higher than for Model 1, which can be an implication that this model is more suited for our data.

5.5.3 Regression Model 3

Table 9: Model 3

Table 9 shows the results of the multiple linear regression from Model 3, containing 39 observations. This table shows the coefficient, the standard error, the t-statistics and the p-value of the intercept, the level of digitalization, years since the digitalization started and the age of the company. In addition, the R-squared and the adjusted R-squared for Model 3 are presented.

	Coefficient	**Std. Err.**	t**-stat**	p**-value**
Intercept	−0.2194	0.1738	−1.2623	0.2152
$\boldsymbol{LoD}$	0.0878	0.0606	1.4488	0.1563
$\boldsymbol{Y}$	0.0028	0.0063	0.4395	0.6630
$\boldsymbol{A}$	−0.0013	0.0021	−0.6139	0.5433
$\boldsymbol{R^2}$	0.0815			
Adjusted $\boldsymbol{R^2}$	0.0028			

Looking at Model 3, the p-value of the level of digitalization is now below our chosen significance level of 19%, as the regression shows a p-value equal to 15.6%. In other words, Model 3 finds a significant positive relationship between the level of digitalization and profitability.

According to this regression, ROA increase by 8.8% for each increased level of digitalization defined in our survey. With size being excluded from the model, this is the only significant relationship discovered.

When we exclude size, we observe that the adjusted R-squared has dropped significantly from both Model 1 and Model 2, which is now only at 0.0028.

5.5.4 Regression Model 4

Table 10: Model 4

Table 10 shows the results of the multiple linear regression from Model 4, containing 39 observations. This table shows the coefficient, the standard error, the t-statistics and the p-value of the intercept, the level of digitalization and years since the digitalization started. In addition, the R-squared and the adjusted R-squared for Model 4 are presented.

	Coefficient	**Std. Err.**	t**-stat**	p**-value**
Intercept	−0.2499	0.1651	−1.514	0.1388
LoD	0.0907	0.0599	1.5129	0.1391
Y	0.002	0.0061	0.3226	0.7488
R^2	0.0716			
Adjusted R^2	0.0201			

In this regression, both size and age of the company are excluded. Similar to Model 3, the relationship between the level of digitalization and profitability is here significant given a significance level equal to 19%. The p-value of LoD has now decreased further, and in this regression, it is equal to 13.9%. The relationship between profitability and years since the digitalization started is here less significant than for all of the previous models and has a p-value as high as 74.9%.

The adjusted R-squared is here higher than for Model 3, but lower than both Model 1 and Model 2, and is equal to 0.02.

5.5.5 Regression Model 5

Table 11: Model 5

Table 11 shows the results of the simple linear regression from Model 5, containing 39 observations. This table shows the coefficient, the standard error, the t-statistics and the p-value of the intercept and the level of digitalization. In addition, the R-squared and the adjusted R-squared for Model 5 are presented.

	Coefficient	**Std. Err.**	t**-stat**	p**-value**
Intercept	−0.2504	0.1631	−1.5357	0.1331
LoD	0.0952	0.0575	1.6554	0.1063
R^2	0.069			
Adjusted R^2	0.0438			

Here we have excluded all other variables but the one directly linked to our research question, namely the level of digitalization, and are now left with a simple linear regression. Now, the p-value of the level of digitalization has decreased even further and is now equal to 10.6%.

Under our chosen significance level of 19%, this relationship is now significant. This model suggests that if the level of digitalization increase by 1 unit, ROA will increase by 9.5%.

In addition, the adjusted R-squared has now increased and are up to 0.044. This is the second highest adjusted R-squared, right below Model 2, which have an adjusted R-squared equal to 0.0453.

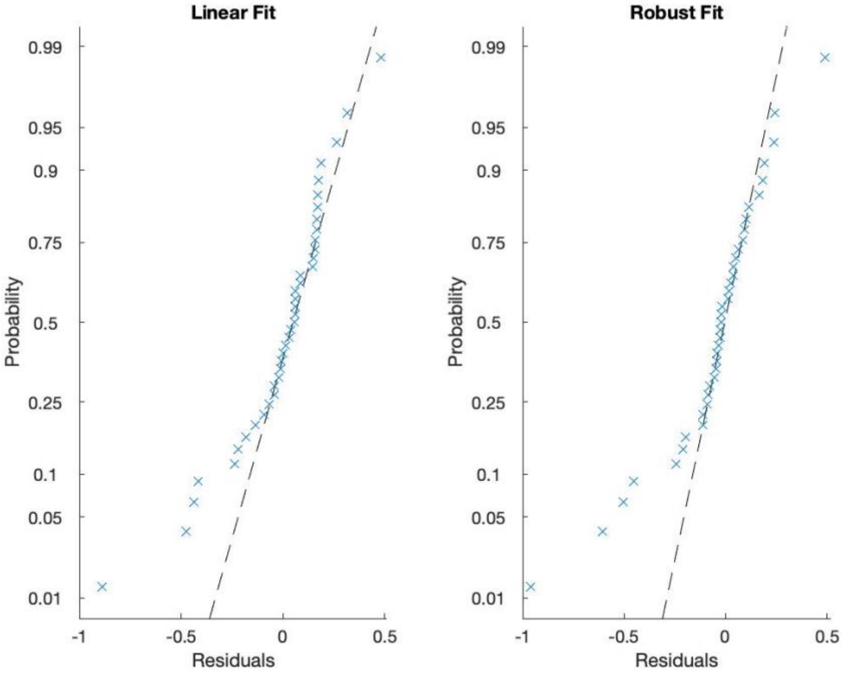
5.6 Robust Linear Regression Model

Using the Robust linear regression name-value pair argument in MatLab, we fitted a robust regression model based on Model 1.

5.6.1 Model fit

We wanted to see if the robust model had a better fit to our data than the linear fit from Model 1. Therefore, we plotted both of the models and looked at the residuals, as you can see below.

Figure 6: *Plot of the model fit vs. the residuals. The left-hand sided figure shows the linear fit of Model 1 while the right-handed side shows the robust fit.*



Here, we can see that the robust fit seems to have a slightly better fit to the residuals, since it is not affected by the observations with large residuals as much as the linear fit.

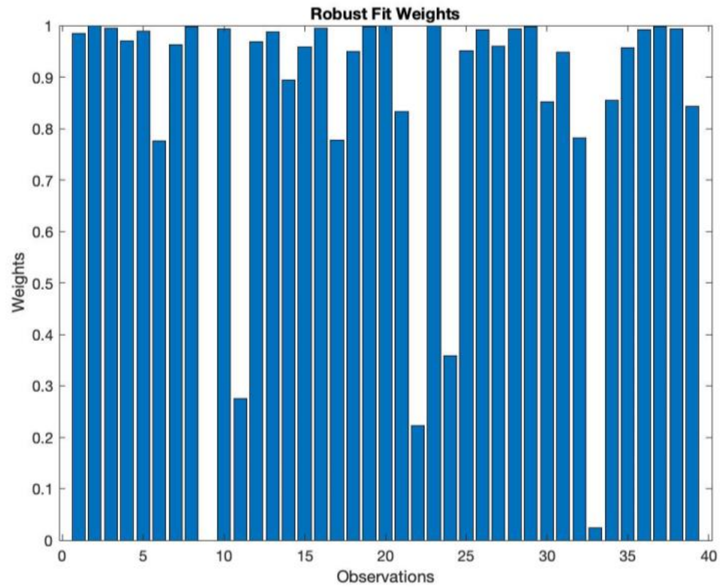
5.6.2 Outliers

In MatLab, we can find which observations who are defined as outliers.

$$Outliers\ = 9, 11, 22, 24, 33$$

The robust regression reduces the weight of the outlier observations in the regression. This is shown below.

Figure 7: *The weights of each observation in our robust regression. The weight of the outliers, which is observation 9,11,22,24 and 33, is lowered compared to the non-outliers.*



Here, we can observe that the weight of the outliers is lowered compared to the other observations in our data.

5.6.3 Robust Regression Model

Table 12: Robust Model

Table 12 shows the results of the robust regression based on Model 1, containing 39 observations. This table shows the coefficient, the standard error, the t-statistics and the p-value of the intercept, the level of digitalization, years since the digitalization started, the age of the company and the size of the company. The weight of the outliers is lowered compared to the non-outliers.

	Coefficient	**Std. Err.**	t**-stat**	p**-value**
Intercept	−0.0839	0.1462	−0.5739	0.5698
LoD	0.0509	0.0495	1.0273	0.3115
Y	0.0031	0.0052	0.5968	0.5546
A	−0.0001	0.0017	−0.0844	0.9332
S	−0.0002	0.0004	−0.5253	0.6028

Here, we receive the highest p-value for the level of digitalization compared to all of the models above. In addition, this regression finds no significant relationship between ROA and none of the variables included in this model.

When using a robust regression, R-squared and adjusted R-squared would not be meaningful (Street, Carroll, & Ruppert, 1988, pp. 152-154), and we will therefore not present these values under this regression.

5.7 Hypothesis testing

From chapter 4.6, we know that the null hypothesis will be kept if the test statistics lies within the non-rejection area. Using a significance level equal to 19%, we receive the following decisions

Table 13: t-test

Table 13 shows the name of the model in the first column, the critical value attached to that model in the second column, the absolute value of the t-statistics in the third column and the decision of the hypothesis test in the fourth column. The table shows these values for Model 1, Model 2, Model 3, Model 4, Model 5 and the Robust Model.

	Critical value	**Absolute value of t stat.**	**Decision**
Model 1	1.3373	1.2085	Do not reject H_0
Model 2	1.3365	1.2682	Do not reject H_0
Model 3	1.3365	1.4488	Reject H_0
Model 4	1.3358	1.5129	Reject H_0
Model 5	1.3351	1.6554	Reject H_0
Robust Model	1.3373	1.0273	Do not reject H_0

Chapter 6: Discussion

6.1 The results

As observed in chapter 5.7, the results of the hypotheses tests are conflicted and therefore inconclusive. There could be several reasons why we did not receive any results, and we will discuss some possible explanations below.

6.2 Sample size

Our sample size consisted of 39 companies. As discussed earlier in this paper, this is lower than we primarily had hoped for. Given a population of 2478 presented in chapter 4.1.6, our sample size is just 1.5% of the total population. Too small sample size can be consequential as it may prevent the findings from being extrapolated (Faber & Fonseaca, 2014). Keeping this in mind, a larger sample size could have helped us with receiving results to our research.

6.3 Key drivers

We discussed several important key drivers in chapter 1.3, and this is now important to revisit after receiving our inconclusive results.

6.3.1 Digitalization and market price on ships

As we saw in chapter 1.3.1, demand and supply will impact the market price of new and second-hand ships. If one company invests large sums in digitalization, it could decrease its maintenance costs or receive a higher price in the second-hand market. Despite of this, the overall market demand and supply would not be affected by a single company's investments in digitalization.

We have earlier discussed that the shipping industry is a volatile industry, and this is reflected on the market price of ships as well. When companies either experience losses or profits while selling and buying ships, this could have a significant impact on the profitability which can undermine the effect of digitalization.

6.3.2 Digitalization and exchange rates

In addition to have an influence on the demand and supply, exchange rates could also have a direct impact on a company's profitability. Companies within the shipping industry, Norwegian export companies as an example, will have a large amount of income in foreign currency, while most of their expenses are in domestic currency. Therefore, a weakened Norwegian krone would increase only the income, while not increasing the expenses.

Because the exchange rates are a product of a very large market, a firms digitalization level will not influence these rates. As for the market price of ships, there could be large sums in either losses or profits connected to the exchange rate, which could reduce the effect of digitalization on the profitability.

6.3.3 Digitalization and freight rates

Freight rates can be affected by digitalization, if a big enough number of companies manage to reduce their costs, and therefore be more competitive in the market. The reduced costs would then again reduce the income, and one could expect the profitability to be balanced out and stay the same after a while.

If a single company managed to reduce their costs due to digitalization, this would probably not influence the freight rates, and therefore not the income. This leads us to digitalization and the costs of running ships.

6.3.4 Digitalization and the cost of running ships

Digitalization could have a higher influence on a company's profitability while talking about the costs of running ships, than the other categories above. This is because a company's investment decisions can affect this directly, either by gaining larger costs due to the high investment expenses, or reduced costs due to efficient solutions.

Because this is the only driver directly connected to digitalization, our missing results may be an indicator that this category is being overshadowed by the other drivers. Despite of this, it does not mean that digitalization is not important for a company in the shipping industry. It could show that we need to measure the effect using other methods, where we strain out the impact of the other strong drivers.

Chapter 7: Conclusion

7.1 Conclusion

The results of the hypotheses tests were inconclusive, meaning that we cannot conclude with neither the null hypothesis nor the alternative hypothesis. From a researcher's perspective, not receiving a result is a result itself. In chapter 6, we discussed several possible reasons for not obtaining a result for our regression models. This can be related to a small sample size and not a suitable method.

As aforementioned, there are a lot of factors influencing the profitability of a firm. To include all of these factors in a model is therefore impossible. In the shipping industry, we often talk about large sums connected to projects and investments. Mechanisms in the market, such as exchange rates and the market price of ships, could therefore affect the profitability of a firm substantially.

The market value of ships, which is a product of demand and supply, will have a great impact on profitability, and especially since we have used ROA as a measurement. As discussed, a firm's digitalization level will not affect the market price. The same applies for the exchange rate and freight rates. For export companies, these two rates will have a large impact on the income, and a single firm's digitalization level will not affect these two rates.

On the other hand, digitalization could make an impact on the operating part of companies within the shipping industry. It is likely to believe that this impact will be positive in the long run. However, because the shipping industry is experiencing volatile market mechanisms that affect the profitability, we believe that the effect of digitalization is being overshadowed.

Therefore, we believe that our results are a consequence of not straining out the unaffected key drivers, and not because digitalization has no impact on the profitability of firms in the Norwegian shipping industry.

7.2 Contribution

We know that the impact of digitalization on firm performance is a largely unexplored topic, as we discussed in chapter 1.4, and that this is a topic of interest. Multiple firms that responded to our survey pointed out how interesting this topic is, and that they were curious about our results.

Even though we cannot present any direct conclusion, we hope to have shed a light on this unexplored and interesting matter. Hopefully, we have presented some intriguing take-aways and inspired others to look deeper into this field of research.

7.3 Further Research

Should it be of interest to investigate this topic further, we have some recommendations we would like to present. Firstly, a bigger sample size could be valuable for the research. Our sample consisted of 39 companies, which is only 1.5% of the population.

Secondly, we would also recommend finding a suitable method to strain out some of the significant market mechanisms that affects the industry as a whole. If one manages to do so, it might be easier to find a relationship between an operational variable as digitalization, and profitability in the shipping industry.

Lastly, we think it would be interesting to look at the development in the relationship over time. In our study, we have used cross sectional data. Collecting panel data on digitalization will enable researchers to observe the influence of digitalization over multiple periods. This may be valuable for the research on this topic as digitalization often are attached with large investment costs affecting the short-term payoff.

7.4 Unforeseen challenges

One of our biggest challenge attached to this study, was the amount of work it took to gather data on the level of digitalization. Previously in this paper, we discussed how many observations we desired to achieve. During the data collection, we saw that achieving such a sample size was not realistic. Although our survey was distributed to over 180 companies in the Norwegian shipping industry, the total sample size ended up being 39 observations. This is equivalent to a response rate of 21.7%. The low sample size was a combination of a lower response rate than excepted, and a lower amount of dispatch due to a higher amount of work attached to each dispatch.

7.5 Social and Ethical Considerations

As mentioned earlier, the data we have collected in this study is in accordance with NSD and do not violate the Privacy Act nor GDPR. Nevertheless, we did receive information about specific companies and their level of digitalization. Since this could be considered as sensitive information, the answers of the survey will not be made public and will be deleted shortly after submission of this paper. Because of this, we believe that we have handled the ethical considerations well.

Given that the result of our research is inconclusive, we do not indirectly recommend managers to invest or not invest in digitalization processes. Therefore, we do not violate any social considerations.

References

Allen, M. P. (1997). *Understanding Regression Analysis.* Boston, MA: Springer.

Andersson, A., & Minnema, J. (2018). The relationship between leverage and profitability: A quantitative study of consulting firms in Sweden. *Umeå School of Business Economics and Statistics*.

Arduino, G., Aronietis, R., Crozet, Y., Frouwds, K., Ferrari, C., Guihéry, L., . . . Van de Voorde, E. (2013, June). How to turn an innovative concept into a success? An application to seaport-related innovation. *Research in Transportation Economics*, pp. 97-107.

Babalola, Y. A. (2013). *The Effect of Firm Size on Firms Profitability in Nigeria.* Accounting and Auditing Department. Ukraine: Journam of Economics and Sustainable Development.

Baruch, Y. (1999). Response Rate in Academic Studies - A Coparative Analysis. *Human Relations*.

Bavassano, G., Ferrari, C., & Alessio, T. (2020, March). Blockchain: How shipping industry is dealing with the ultimate technological leap. *Research in Transportation Business & Management*.

Borenstein et. al. (2010, November 21). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, pp. 97-111.

Brewer, S. (2018, September 3). *How will digitalisation impact shipping?* . Retrieved from Linkedin : https://www.linkedin.com/pulse/how-digitalisation-impact-shipping-stuart-brewer/

Brooks, C. (2014). *Introductory Econometrics for Finance.* New York, USA: Bambridge University Press.

Corporate Finance Institute. (2020, January). *ROA Formula / Return on Assets Calculation*. Retrieved from CFI: https://corporatefinanceinstitute.com/resources/knowledge/finance/return-on-assets-roa-formula/

Dogan, M. (2013). Does Firm Size Affect The Firm Profitability? Evidence from Turkey . *Research Journal of Finance and Accounting* .

Dong, H. P., & Su, J.-T. (2010). *The Relationship between Working Capital Management and Profitability: A Vietnam Case.* International Research Journal of Finance and Economics.

Ellingsen, O., & Aasland, K. E. (2019, June 7). Digitalizing the maritime industry: A case study of technology acquisition and enabling advanced manufacturing technology. *Journal of Engineering and Technology Management*, p. 27.

Ellingsen, O., & Aasland, K. E. (2019, September 30). Digitalizing the maritime industry: A case study of technology aquisition and enabling advanced manufacturing technology. *Journal of Engineering and Technology Management*, pp. 12-27.

Faber, J., & Fonseaca, L. M. (2014). How sample size influences research outcomes. *Dental Press Journal of Orthodontics*.

Ferreira et. al. (2019, August). To be or not to be digital, that is the question: Firm innovation and performance. *Jorunal of Business Research*, pp. 583-590.

Freund, R. J., Wilson, W. J., & Sa, P. (2006). *Regression Analysis : Statistical Modeling of a Response Variable.* Elsevier Science & Technology.

Grytten, O. H., & Koilo, V. (2019). *Norsk industri*. Retrieved from Utfordringer for norske maritime næringer: krise og omstrukturering: https://www.norskindustri.no/bransjer/maritim/sjokart-for-maritim-bransje/industriutvikling/utfordringer-for-norske-maritime-naringer/

Hall, M., & Weiss, L. (1967). *Firm Size and Profitability*. The MIT Press.
Hamid, M. A., Abdullah, A., & Kamaruzzaman, N. A. (2015). Capital Structure and Profitability in Family and Non-Family Firms: Malaysian evidence. *Procedia Economics and Finance*.
Haykir, Ö., & Çelik, M. S. (2018). *The Effect of Age on Firm's Performance: Evidence from Family-Owned Companies*. Ömer Halisdemir Üniversitesi.
Holopainen, R., Niskanen, M., & Rissanen, S. (2019, June 20). The impact of internet and innovation on the profitability of private healthcare companies. *Journal of Small Business & Entrepreneurship*.
Ibem et al. (2018). Survey data on digitalization of buildingprocurement process by architecturalfirms inAbuja, Nigeria. *Data in Brief*.
IBM. (2021, May). *One-Sample Kolmogorov-Smirnov Test*. Retrieved from IBM: https://www.ibm.com/docs/en/spss-statistics/SaaS?topic=tests-one-sample-kolmogorov-smirnov-test
Innovasjon Norge. (2020). *Maritime*. Retrieved from innovasjonnorge.no: https://www.innovasjonnorge.no/en/start-page/invest-in-norway/industries/maritime1/
Jotun Marine. (2018, September 3). *How will digitalisation impact shipping?* Retrieved from Jotun Marine: https://www.jotunmarine.com/ww/en/b2b/How-will-digitalisation-impact-shipping.aspx
Keuzenkamp, A. H., & Magnus, R. J. (1995). On tests and significance in econometrics. *Journal of Econometrics*(67).
Kim, J. (2015). How to Choose the Level of Significance: A Pedagogical Note. *MPRA*.
Koskinen, K., & Manninen, O. (2019, May 29). The impact of digitalisation on bank profitability. *Bank of Finland*.
Lantz, B. (2014). *Den statistiska undersökningen: grundläggande metodik och typiska problem*. Lund: Studentlitteratur.
Lee, J. H. (1991). A Lagrange multiplier test for GARCH models . *Economics Letters* .
Levin, K. A. (2006). Study design III: Cross-sectional studies. *Evidence-Based Dentistry*.
Loderer, C., & Waelchli, U. (2010). *Firm age and performance.*
Malik, H. (2011). Determinants of Insurance Companies Profitability: An Analysis of Insurance Sector of Pakistan. *Academic Research International*.
Morgan, B. (2019, September 23). *Forbes*. Retrieved from Editors Pick: https://www.forbes.com/sites/blakemorgan/2019/09/30/companies-that-failed-at-digital-transformation-and-what-we-can-learn-from-them/?sh=1bdd721b603c
Newbold, P., Carlson, W. L., & Thorne, B. M. (2013). *Statistics for Business and Economics* (Vol. Eighth Edition). Essex, England: Pearson Education Limited.
Niemand et. al. (2020, April 25). Digitalization in the financial industry: A contingency approach of entrepreneurial orientation and strategic vision on digitalization. *European Management Journal*.
Niresh, J. A., & Velnampy, T. (2014). *Firm Size and Profitability: A Study of Listed Manufacturing Firms in Sri Lanka.* University of Jaffna, Sri Lanka, Department of Accounting. International Journal of Business and Management.

Norwegian Shipowners Association. (2018). *Think Ocean - Maritime Outlook Report.* Oslo: Norwegian Shipowners' Association.

O'Gorman, K., & MacIntosh, R. (2015). *Research Methods for Business & Management: A Guide to Writing Your Dissertation.* Oxford: Goodfellow Publishers Ltd.

Piper, R. J. (2013). *How to write a systematic literature review: a guide for medical students.* University of Edinburgh.

Plane, D. R., & Gordon, K. R. (1982, February 1). A Simple Proof of the Nonapplicability of the Central Limit Theorem to Finite Populations. *The American Statistician Volume 36*, pp. 175-176.

PwC. (2017, September). *The Digital Tranformation of Shipping.* Retrieved from PwC: https://www.pwc.no/no/publikasjoner/shipping/The-Digital-Transformation-of-Shipping_HE-NO.pdf

Rederi. (2018, March 18). *Rederi.* Retrieved from Norge er en av verdens største shippingnasjoner: https://rederi.no/aktuelt/2018/norge-er-en-av-verdens-storste-shippingnasjoner/

Rederi. (2020, januar). *Rederi*. Retrieved from Statistics and key figures: https://rederi.no/om-oss/statistikknokkeltall/

Regjeringen. (2020, December 4). *Regjeringen.* Retrieved from Meld. St. 10 (2020–2021) Grønnere og smartere – morgendagens maritime næring: https://www.regjeringen.no/no/dokumenter/meld.-st.-10-20202021/id2788786/

Reve, T. (2009). Norway – a global maritime knowledge hub.

Ross, A., & Wilson, V. L. (2017). *Basic and Advanced Statistical Tests: Writing Results Sections and Creating Tables and Figures.* Rotterdam, The Netherlands: Sense Publishers.

Salvi et al. (2021). *Online information on digitalisation processes and its impact on firm value.* Journal of Business Research.

Splash 24/7. (2017, February 10). *Splash 24/7*. Retrieved from Dr Martin Stopford on the future of shipping: https://splash247.com/dr-martin-stopford-future-shipping/

Stopford, M. (2009). *Maritime Economics* (Vol. 3rd edition). New York: Routledge - Taylor & Francis Group.

Street, J. O., Carroll, R. J., & Ruppert, D. (1988). A Note on Computing Robust Regression Estimates Via Iteratively Reweighted Least Squares. *The American Statistician*, pp. 152-154.

Studenmund, A. H. (2011). *Using Econometrics: A Practical Guide* (Vol. XI). Pearson.

Tenold, S. (2019). *Norwegian Shipping in the 20th Century.* Cham: Springer.

Tseng, P.-H., & Liao, C.-H. (2015, May 11). Supply chain integration, information technology, market orientation and firm performance in container shipping firms. *The International Journal of Logistics Management Volume 26 Issue 1*.

Wooldridge, J. M. (2015). *Introductory Econometrics - A Modern Approach* (Vol. Sixth Edition). Boston, Massachusetts, USA: Cengage Learning.

Yang, C.-S. (2019, November). Maritime shipping digitalization: Blockchain-used technology applications, future imporvements, and intention to use. *Transportation Research Part E: Logitics and Transportation Review*, pp. 108-117.