

UNIVERSITY OF TARTU  
Faculty of Social Sciences  
School of Economics and Business Administration

Rauno Juntson

Predicting New Continent Entrance of Estonian Manufacturing Firms

Master's thesis

Supervisor: Oliver Lukason, PhD

Tartu 2023

Forwarded to defense .....

(supervisor's signature)

I have composed this master's thesis independently. All materials, viewpoints from literature and other sources used to write this thesis have been referenced.

.....

(author's signature)

**Abstract**

The goal of this thesis is to create a set of prediction models that can forecast the entry to a market in a new continent for a firm. Continents are categorized according to geography as either American, Asian and Oceanian or African. 13 different variables from three domains, export revenue in a continent, the proportion of export revenue from each continent to the total export revenue and the number of markets on a continent, are used in the prediction models with 10 variables in each model. Three different prediction methods, binary logistic regression, artificial neural network and decision tree are used to create models for predicting new market entries in continents a firm had no presence in before. The used dataset was the whole population of Estonian exporting manufacturers between 2009-2019 with 25,291 observations in total. The composed models predicted the new continent entries with at least 95% accuracy. The best predictor for whether an exporting firm can be categorized as an entrant to a new continent was the total export revenue.

Keywords: new continent entrance prediction, export revenue, export share from sales, number of markets,

CERCS: S180, S186, S190

**Table of contents**

Introduction.....5

1. Theoretical background ..... 7

    1.1. Literature on internationalization process..... 7

    1.2. Literature review on studies of market entry ..... 10

2. Study design..... 12

    2.1. Data and methodology ..... 12

    2.2. Variables and descriptive statistics..... 13

3. Results and discussion ..... 18

    3.1. Predictive models ..... 18

    3.2. Importance of independent variables..... 19

        3.2.1. Entry to Africa ..... 22

        3.2.2. Entry to America..... 22

        3.2.3. Entry to Asia and Oceania ..... 23

    3.3. Theoretical implications and limitations of the study ..... 24

Conclusion ..... 24

References..... 27

    Appendix 1. Examples of empirical studies focusing on export-related dependent variables ..... 34

    Appendix 2. Individually analyzed variables with the direction of the effect using binary logistic regressions..... 37

    Appendix 3. Individually analyzed variables on the entry to an African market with coefficients and statistical significance..... 38

    Appendix 4. Individually analyzed variables on the entry to an American market with coefficients and statistical significance..... 39

    Appendix 5. Individually analyzed variables on the entry to an Asian or Oceanian market with coefficients and statistical significance ..... 40

Kokkuvõte..... 41

## Introduction

In the wind of globalization since the latter half of the 20<sup>th</sup> century, exporting and internationalization have been attracting the interest of numerous researchers. This subject has been studied using both micro and macrolevel analysis through multiple lenses, including why and how firms start to export, what capabilities a firm needs to internationalize and the barriers it might encounter. The received attention is unsurprising since if a company wants to grow a product line faster than its current markets allow, it is inevitable that one moment it must start exporting to new markets in order to keep growing. The trend of internationalization has slowed during the last few years due to the pandemic and increased geopolitical tensions, however the trend is likely to continue in the long term (Brenton et al., 2022). This emphasizes the importance of studying the mechanisms behind internationalization. Furthermore, there is evidence that governmental support for promoting exports increases the propensity to export (e.g. Hussien, 2022; Onah et al., 2022; Srhoj et al., 2021), which makes the precise aiming of government aid paramount due to different grant programs often having limited resources.

There is no universally agreed upon definition of internationalization in literature<sup>1</sup>, but there are a multitude of activities a company can do that could be considered as taking part of international markets. A number of firms choose to make foreign direct investments by acquiring whole companies or shares in them, others share knowledge and make connections on the microlevel and learn from each other. Yet some participate in global supply chains by exporting services or products through a representative, subsidiary or directly. This paper will examine internationalization through the lens of firms exporting products outside of the home market.

Internationalization cannot be considered a singular moment but rather a process, which can be viewed akin to a long list of steppingstones, that each point towards a series of increased foreign involvement (Welch and Welch, 1996; Johanson and Vahlne, 1977). A small moment in that process is when a firm first starts receiving revenue outside of home nation or in other words, the first time it starts exporting. These moments are what this paper analyzes by looking at the years when an Estonian company starts selling products to a market that is outside of the home continent of Europe. Furthermore, this paper tries to predict these entries by using different methods.

---

<sup>1</sup> For an extensive list of definitions and their analysis, see Morales (2020).

The prediction of the entries to markets that are geographically distant could be advantageous for general market analysis by different firms, trying to outcompete their competition in a specific market. Also, the findings of this paper can be beneficial for different governmental organizations building better programs of giving grants or resources to firms with the goal of promoting exports. Provided a firm has displayed the characteristics of someone that is going to start exporting to the desired location, but has not done so, targeting governmental recourses to those firms might be beneficial in fulfilling the policy of promoting exports.

This paper has the goal of creating a set of prediction models that can forecast the entry to a market in a new continent for a firm. The continents are categorized according to geography as either African, American and Asian and Oceanian. Furthermore, this thesis will analyze the created models by considering the different variables from three domains: export revenue in a continent, the proportion of export revenue from each continent to the total export revenue and the number of markets on a continent, and their individual importance in predicting an entry to a given continent. Three different methods, binary logistic regression, artificial neural network and decision tree are used to create models for predicting market entries. The variety of methods is used to verify the findings and compare different methods of prediction, which could be beneficial for future research. Binary logistic regression is also used with variables individually to find out which variables have the highest predictive accuracies.

This thesis adds to the literature of export research by analyzing firm-level data of exports and their destinations. By using data of all Estonian exporting firms from the manufacturing industry, this paper offers a unique example of export behavior of companies from a small and export-oriented economy. Also, this paper has a distinctive approach by dividing potential markets for expansion into continents, which could yield better predictive accuracy for the models created. Furthermore, there are not many studies that systematically use the previous information from markets to predict new market entries.

The structure of the paper consists of three main parts, an introduction and conclusion. The first part focuses on the theoretical background and includes a short overview of the foundations of new market entries through the lens of internationalization (process) literature and a review of empirical research focused on new market entries. The second part starts the empirical portion by discussing the data and methodology of the study and the descriptive statistics of different independent variables. The third part of the thesis brings out and discusses the results of the empirical study. The thesis ends with a conclusive

part, which summarizes the thesis, proposes some practical implications and future research directions.

## **1. Theoretical background**

### **1.1. Literature on internationalization process**

Internationalization of companies has often been traditionally researched through the lens of either the slow and incremental Uppsala model, first described by Johanson and Vahlne in 1977 and further developed by them later (i.e. 2009; 2010; 2013; 2017), or the fast internationalizers of the born global type. The born global model focuses on young companies that internationalize early on in their operations, while the Uppsala model describes a company that internationalizes more gradually (Bouncken et al, 2015). According to Vahlne and Johanson (2017) what makes the Uppsala model distinctive is its focus on not isolated decisions, but rather on the process of acquiring knowledge and committing resources. Furthermore, the model combines this process with content by agreeing that the strategy of a firm emerges from a number of processes, which are not all controlled by the management of a firm (Vahlne and Johanson, 2017). An assumption under the Uppsala model is that firms in their early stages of internationalization are lacking the knowledge of international markets, which forces them to internationalize gradually (Bouncken et al, 2015). Further features of the Uppsala model include but are not limited to the heterogeneity of firms, realism, bounded rationality and the idea that the performance of companies is influenced by management and other actors (Vahlne and Johanson, 2017).

However, similarly to the Uppsala model, the born global approach focuses on knowledge and capabilities, which when viewed through the lens of a knowledge-based view, allow the management of a firm to focus its efforts on a market further away from their origin country early on after the firm's foundation (Bouncken et al, 2015; Knight and Cavusgil, 2004). In this regard these theoretical frameworks are similar as they both agree on the importance of knowledge as an influencer of firms' internationalization. Only the effect of knowledge is viewed through slightly different lenses in the way it influences the internationalization process.

There is evidence that there are hierarchies in the markets that companies choose to internationalize to (e.g Lawless, 2009). As explained in the Uppsala model as path dependence (Vahlne and Johanson, 2017), a similar idea of (extended) gravity (Morales et al, 2019) or sequentialism (Albornoz et al, 2012; Welch and Luostarinen, 1988) states that companies first go to markets that are the most similar to the home market or easiest to export to by the means of geography, language, income per capita, expected profitability or culture.

All the mentioned factors can reduce the costs of a market entry (Lawless, 2009; Vahlne and Johanson, 2017; Morales et al, 2019; Alborno et al, 2012; Welch and Luostarinen, 1988). Indeed, there is evidence that a new market for a company is more likely to be on the same continent as the company itself (da Silva et al, 2023).

Hence, the hierarchies of export markets can be dependent on not only the geography of export markets but also other factors. The logic behind this decision making can also be known as “psychic“ distance (O’Grady and Lane, 1996; Ojala, 2015). Johanson and Vahlne (2017) explain this incrementality as a way for firms to respond to risk – expanding and making decisions incrementally rather than at once allows companies to prevent the potential of substantial mistakes and the losses that could be incurred. The increased probability of exporting companies to start exporting to further markets has, in some models of market entry, been explained partly by the idea of sunk costs. The idea is that involvement in new markets is affected by preceding experience, or in other words, participation in some export markets increases the probability of firms to participate in further export markets (Roberts and Tybout, 1997; Fichera and Shevtsova, 2020).

However, research also suggests that a minority of companies implement a distinct strategy that is opposite to the strategy of sequential internationalization, labeled non-sequential internationalization, by selecting markets that are different from their home market as their first market to internationalize to (Cuervo-Cazurra, 2011). Furthermore, in contrast to the idea of path dependence some companies inhibit the behavior of constantly entering and exiting a market (Vissak and Masso, 2015; Bernini et al., 2016). In their typology of exporters Vissak and Masso (2015) defined one type of internationalizers as serial nonlinear internationalizers if they “have de- and re-internationalized at least twice“. Firms that enter and exit markets repeatedly can constitute a significant part of a nation’s exporting companies (eg. Blum et al, 2013).

There is debate in literature whether this type of non-linear internationalization constitutes a failure for the internationalization process of a firm. This is due to the ability of companies to learn from their past experiences. Welch and Welch (2009) have found evidence that some re-internationalizes retain a baggage of knowledge and networks that help them internationalize faster in the future while for other firms the failed internationalization attempt makes them less interested in re-internationalization. The former is being also supported by Vissak et al (2020) who found that attaining knowledge from previous operations in export markets means that future market entries are more systematic as opposed to making decisions just when opportunities arise. In case studies of French manufacturing

SMEs Dominguez and Mayrhofer (2017) found results similar to the evidence found by Welch and Welch (2009).

In short, firm level exports have been found to be unstable (Vissak and Masso, 2015) and there is debate whether this exiting of a market can constitute a failure for an exporting firm, hence, there could be multiple reasons for exiting a market and it might not be reasonable to expect stability among the variables. It is key to note that this paper does not address the probability of companies leaving markets after they have entered them.

Another aspect of internationalization are the spillover effects. This means there could be other companies in the vicinity of a potential exporter, that have experience in internationalizing or have already started exporting to other countries, which would influence the decision to export. Contradicting somewhat with the basic assumption under the Uppsala model in the regard that firms have little or no knowledge of international markets in the beginning of their expansion process (Bouncken et al, 2015), spillover effects from nearby exporters can reduce the uncertainties of entering new markets and thereby could increase the likelihood of exporting to a market (Segura-Cayuela and Vilarrubia, 2008).

Taking Mexican industrial plants as an example, Aitken et al. (1997) found evidence of spillover effects from multinational companies but not from the general population of exporters. In contrast to these findings Fichera and Shevtsova (2020) found on the example of Ukraine that information-sharing between geographically close exporters and ownership structure of exporters did not significantly influence the probability of entering a new export market. The export prediction model that is discussed in this paper does not separately take into account the possible spillover effects, which could be used to explain a part of the findings, but the analysis of which is out of the scope of this paper.

In summary, the process of internationalization has been researched through the lens of fast internationalizing born-globals or the slow and incremental internationalizers according to the Uppsala model. There is evidence that the markets that firms enter are not random and have a hierarchy in them. Contrary to that, some firms internationalize non-sequentially or exhibit the behavior of entering and exiting markets multiple times. This is regarded as a failure by some researchers while others find that the internationalization attempts give firms valuable knowledge for expansion in the future. Furthermore, there is a potential of spillover effects on internationalization between firms, meaning that a firm could learn from nearby firms and reduce the uncertainties of entering new markets.

## 1.2. Literature review on studies of market entry

Firm-level data of precise export locations and the value of those exports has been historically largely hard to find, hence this area has been under-researched. However, the usage of firm-level data for quantitatively analyzing entries into new export markets has not been entirely deficient from research articles.

There is a diverse range of factors that could influence the decision to internationalize. These include both financial and non-financial factors, knowledge, systems and operations (Welch and Welch, 1996). As mentioned earlier, when viewing internationalization through the lens of a born global firm, what allows a firm to expand to further markets are the firm's innovativeness, knowledge and capabilities (Knight and Cavusgil, 2004). Multiple researchers agree that indeed, different kinds of internal capabilities of a firm contribute to the ability to internationalize (eg. Tallman and Fladmoe-Lindquist, 2002; Prange and Verdier, 2011; Raymond et al, 2014; Chen et al, 2016). A simple proxy for the resources and capabilities that a company possesses could be the size of the company (Tang et al, 2021), which has been used extensively in market entry analysis (see Appendix 1).

The table in Appendix 1 will outline a selection of articles that use firm-level data to analyze market entries and internationalization. The articles were found by the author of this thesis by searching "market entry" and "internationalization prediction" from Google Scholar, Scopus and Web of Science and filtering out relevant articles. This list of articles outlined in the table does not mean to be exhaustive but rather illustrative of the gap that exists in literature of analyzing and predicting entries into markets on new continents.

All the selected articles have variables relating to exporting propensity as the analyzed dependent variables. These include, for instance, the number of export markets (Lawless, 2009; Fichera and Shevtsova, 2020), the proportion of revenue from European markets to total revenue (Lukason and Vissak, 2020), exporting directly or indirectly (Lu et al., 2017; Razzolini and Vannoni, 2011), the proportion of export revenues to total revenues (Ye et al., 2019) or others. A dependent variable most common in the analyzed articles was the exporting status of firms (Buch et al., 2010; Eaton et al., 2011; Draz et al., 2015; Micocci and Rungi, 2022). The list of countries in which the analyzed firms were from tended to be large, such as China (Ye et al., 2019), France (Eaton et al., 2011; Micocci and Rungi, 2022), Germany (Buch et al., 2010), Ukraine (Fichera and Shevtsova, 2020), Italy (Razzolini and Vannoni, 2011) and Pakistan (Draz et al., 2015), with a few being from smaller countries, Ireland and Estonia (Lawless, 2009; Lukason and Vissak, 2020). One study included data from 29 different developing countries (Lu et al., 2017). The amount of data points included

in the studies varied greatly from 132 observations in total (Draz et al., 2015) to up to 230,423 different firms being analyzed over 7 years (Eaton et al., 2011).

There is a multiplicity of independent variables analyzed. The usage of the internal capabilities, including financial capabilities and potential barriers, along with other characteristics of a firm are widespread to explain internationalization as dependent variable (see Appendix 1). Productivity and size as variables have been thoroughly researched, both of these factors yielding mostly positive results on internationalization (Lawless, 2009; Fichera and Shevtsova, 2020; Buch et al., 2010; Razzolini and Vannoni 2011). Furthermore, some researchers use variables relating to the human capital (Lukason and Vissak, 2020; Draz et al. 2015; Ye et al., 2019; Lu et al., 2017; Razzolini and Vannoni, 2011; Draz et al., 2015), ownership (Fichera and Shevtsova, 2020; Lukason and Vissak, 2020; Ye et al., 2019) or products of a company (Draz et al. 2015) or try to explain internationalization to a given market according to a common hierarchy (Lawless, 2009; Eaton et al., 2011). A set of factors that have been used to predict the moment when companies start exporting have been the financial metrics that companies inhibit right before they start to internationalize (Micocci and Rungi, 2022; Buch et al. 2010). Using financial indicators is in accordance with the view that firms need to have a good set of capabilities to expand to further markets or that there are barriers which make internationalization less likely, financial capabilities are a part of the capabilities needed to internationalize or financial constraints being barriers that bar a company from it (Buch et al. 2010).

In summary, there are multiple factors that have been used in analyzing firm-level data in the context of market entries of firms. The internal capabilities of a firm, including financial and non-financial are important factors in expanding to different markets. Variables that have been used in analyzing the entries into markets range widely and these include those relating to the size, financial metrics, productivity, human capital, market coverage, relationship with other firms and ownership of a firm. Exporting and market entry have been researched using various methods, including correlation analysis, different types of regression analysis, and neural networks and decision trees for prediction. The amount of observations included in different studies varied considerably and the data used were mostly samples and not the general population of companies. Many of the studies used data from large countries with there being fewer studies using data from a small and open economy such as Estonia. Furthermore, none of the studies used a method of analyzing entries into continents that the firm had no presence in. The usage of independent variables related to its existing export activities is limited in the studies found. However, this could be due to the reason of the focus

of the studies being on the start of exports, the propensity of exports and other variables related to exports but not the entry to a continent that is the focus of this study.

## 2. Study design

### 2.1. Data and methodology

The study was conducted using data from the Business Registry of Estonia, which compiles data from annual reports uploaded to its system. The data is gathered from an annex of a firm's annual reports that states revenues by geographic location. Companies included in the analysis are all manufacturing firms, i.e. according to the Estonian Classification of Economic Activities<sup>2</sup> the companies are classified as class C (EMTAK, 2008) with the exclusion of industrial equipment repairs (EMTAK C subcode 33). The geographic locations are compiled into four large geographic sets, roughly corresponding to continents. They are America (this includes both South and North America), Asia and Oceania, Africa, and Europe. With transcontinental nations, a nation is categorized into a continent by the location of its capital. Analyzing different variables according to regions is common in internationalization studies (Patel et al, 2018; Fernhaber et al, 2008; Townsend et al, 2009) and it has also been used in prediction models for internationalization (Lukason et al, 2021).

The dataset includes data from the years 2009-2019 which was a relatively long period of expansion in the Estonian economy with the Estonian GDP growing nominally every year by an average rate of 7,02% (Statistics Bureau of Estonia, 2023a). Before the year 2009 the data regarding exporting is not available in a processable format in the Business Registry, hence the first year of the dataset was chosen to be 2009. The last year to be included is 2019 because the coronavirus SARS-CoV-2 began spreading in the year 2020 which brought with it lockdowns and a sharp decrease in economic activity, both of which might skew the analyzed data.

The dataset has 25,291 observations, with 3325 companies in total. The data is heavily skewed towards companies that export to Europe with 18,037 observations having at least some exports to a European country. Similarly to the Estonian data, in many other nations a significant proportion of exporters have only a few markets that are geographically close to them (e.g. Eaton et al, 2007; Fichera and Shevtsova; 2020). Furthermore, this general observation about Estonian exporters is in line with the theoretical findings that companies are keener on exporting to markets that, among other factors, are closer to them in geography,

---

<sup>2</sup> EMTAK is a local variation of the European Nomenclature of Economic Activities (NACE) codes. The first four classes of EMTAK codes, which include EMTAK C, correspond with their NACE equivalents (EMTAK, 2008).

language or culture (Vahlne and Johanson, 2017); furthermore, Estonia being in the European Union makes exporting to Europe simpler than exporting to countries further away with possible trade barriers.

Three models are used to analyze and predict market entries: an entry to American markets, Asian and Oceanian markets, and African markets<sup>3</sup>. For each analysis two subsets of data are filtered out from the larger dataset. One subset is composed of when companies first started exporting to a given continent. The other subset includes companies that were exporters but did not, had not and would not export to a given continent during the years included in the analysis. Variables were taken from the year before the entry. Due to the majority subset being overestimated, the datasets are made roughly equal using the Synthetic Minority Oversampling Technique (SMOTE) method by replicating the minority, i.e. the dataset with the moments of market entry. A method of minority overrepresentation is chosen instead of majority underrepresentation, because removing observations from the majority class could also possibly remove valuable data from the models.

The study then uses three analytical methods to predict the entry to a new continent and to determine the best indicators for that. These methods are binary logistic regression, multilayer perceptron artificial neural networks, and decision tree. Binary logistic regression is used in classification or in predictive analysis by determining the probability of an event occurring with there being only two possible outcomes (IBM, 2023). The other two methods used are types of computer-learning algorithms. Artificial neural networks are computational models that take inputs, process them through nodes, each of which is a linear regression model on its own, and produce outputs that can be used in predictions by iteratively adjusting the strengths of connections between nodes (IBM, 2019; Krenker et al, 2011). The third method used, decision trees, partition the population into smaller subsets based on the inputs and creates a tree, which provides an output or a prediction based on the input data (Song and Lu, 2015). As indicated in the literature review, these methods have previously been used by researchers in categorizing companies as either an exporter or not (Lukason and Vissak, 2020; Draz et al., 2015; Micocci and Rungi, 2022).

## **2.2. Variables and descriptive statistics**

Each model of market entry prediction uses 10 different independent variables to predict market entry, with 13 indicators in total. These independent variables are outlined in

---

<sup>3</sup> European market entries are not analyzed because there are very few companies that have exports but do not export to Europe or have not had any exports to Europe during the years 2009-2019, hence the opposing group for the predictive model could be easily skewed by a few observations. The dataset includes only 16 companies that had exports but did not have any exports to a market in Europe.

the table below (Table 1). The variables can be divided into three types, each containing variables relating to a specific continent.

The first type are variables relating to the number of markets. The number of export markets has been used as a variable in internationalization research (Gripsrud et al., 2023; Lawless, 2009; Lukason and Vissak, 2020; Fichera and Shevtsova, 2020), but where this study differs from previous studies is that the number of markets is checked separately for each of the continents. This is done to not mix up different types of exporters and by that to try to increase the predictive accuracy of the model. A firm that has a few export markets in Europe is likely to be a different kind of exporter than the one that has only a few markets in America, with the former following a more traditional expansion path as described in the Uppsala model (Vahlne and Johanson, 2017) and the latter searching for more profitable markets or producing an unconventional product which has demand for it in geographically distant markets.

The second and third sets of variables are derived from the revenue of a company. These variables are the export revenue to a given continent, the proportion of revenue from a given continent to total export revenue and the total revenue from export. The usage of revenue as an indicator for size is widespread in literature (Dang et al., 2018). The revenue variables could be taken as a proxy for the size of a firm as an exporter, with total export indicating the total size of the firm and the revenue on each of the continents the size of a firm as an exporter on the given continent. This differentiation is done on the same grounds as the one for markets, i.e. to more clearly separate different types of exporters and to create more accurate predictive models with this. Predictors using revenue are not used in their net sum but instead the natural logarithm of that sum. This is done to reduce the impact of extreme values and make companies more comparable for the model.

As mentioned, in addition to the sum of total export sales and export sales on a continent, three more variables are created by dividing the revenue from a continent with total export revenue. These variables are used on different continents separately regarding the same considerations as for the other sets of variables.

**Table 1.** Variables used in the models

<b>Variable subgrouping</b>	<b>Variable</b>	<b>Variable explanation</b>	<b>Variable used in entries to African markets</b>	<b>Variable used in entries to Asian and Oceanian markets</b>	<b>Variable used in entries to American markets</b>
Variables relating to Africa	Export to Africa	Natural logarithm of revenue from markets that are in Africa.	-	+	+
	Proportion of African exports	Export revenue from African markets divided by total revenue from export.	-	+	+
	Number of African markets	Number of different markets in Africa.	-	+	+
Variables relating to Asia and Oceania	Export to Asia and Oceania	Natural logarithm of revenue from markets that are in Asia or Oceania.	+	-	+
	Proportion of Asian and Oceanian exports	Export revenue from Asian or Oceanian markets divided by total revenue from export.	+	-	+
	Number of Asian and Oceanian markets	Number of different markets in Asia or Oceania.	+	-	+
Variables relating to South and North America	Export to America	Natural logarithm of revenue from markets that are in Africa.	+	+	-
	Proportion of American exports	Export revenue from America divided by total revenue from export.	+	+	-
	Number of American markets	Number of different markets in South or North America.	+	+	-
Variables relating to Europe	Export to Europe	Natural logarithm of revenue from markets that are in Europe.	+	+	+

	Proportion of European exports	Export revenue from European markets excluding Estonia divided by total revenue from export.	+	+	+
	Number of European markets	Number of different markets in Europe excluding Estonia itself.	+	+	+
Other variables	Total export	Natural logarithm of revenue from all export markets.	+	+	+

Note: in the last three columns a plus sign indicates that the variable was used in the model for the entry into the respective continent and a minus sign indicates that it was not used

Source: compiled by the author

As described previously, this paper analyses the start of exports to a given continent. A moment of market entry is marked when a company did not have any revenues in any of the markets in a given continent the previous year and does have revenues in it the next year. There are 87 entries to Africa, 287 to America, 887 to Europe and 380 to Asia and Oceania. Averages for these moments and their opposing group, except for the moments to Europe because of the absence of the opposite group, are brought out in the table below (Table 2).

A factor to note from the general statistics is that companies in all groups tend to be more prolific in exporting to European markets with them having the largest share of their exports being to European markets, the most exports to European markets in nominal terms and having most export markets in Europe as opposed to other continents. This coincides with the general statistics of Estonian exports, where in 2022 81.7% of exports go to countries that are classified as European in this study (Statistics Bureau of Estonia, 2023b).

According to the export revenue and share of export revenue to total revenue, in the data analyzed, exporting to Europe is followed by exports to Asia and Oceania and then America. However, according to the total export statistics of Estonia, American markets are in the second place and Asian and Oceanian markets in the third place with them being 7.2% and 6% of Estonian exports, respectively (Statistics Bureau of Estonia, 2023b). When comparing the average scope of exports to different continents, an exporting firm had the least exports, fewest markets and the smallest proportion of exports coming from markets that are in Africa. This again is similar to the general trade data of Estonia where less than 2% of Estonian exports went to African markets (Statistics Bureau of Estonia, 2023b).

**Table 2.** Means of different variables across analyzed groups

Means	Export to Africa	Proportion of African exports	Number of African markets	Export to Asia and Oceania	Proportion of Asian and Oceanian exports	Number of Asian and Oceanian markets	Export to America	Proportion of American exports	Number of American markets	Export to Europe	Proportion of European exports	Number of European markets	Total export
Entry to an African market				8.68	0.12	2.30	4.36	0.06	1.07	44.89	0.67	10.29	63.02
Firms with no exports to Africa				0.65	0.02	0.15	0.37	0.02	0.08	11.48	0.79	2.49	12.91
Entry to an Asian or Oceanian Market	0.65	0.01	0.18				1.87	0.04	0.46	26.14	0.72	5.74	34.63
Firms with no exports to Asia or Oceania	0.00	0.00	0.00				0.08	0.01	0.02	8.38	0.79	1.83	8.72
Entry to an American market	1.07	0.01	0.30	3.79	0.06	0.94				29.19	0.71	6.44	39.60
Firms with no exports to America	0.02	0.00	0.00	0.23	0.01	0.05				9.09	0.79	1.97	9.62

Source: compiled by the author

Comparison of the means indicates that for entries into all new continents, the entrants tend to be multiples larger in size as measured by revenue from exports, with the entrants to Africa having a logarithmic revenue from exports that is 4.88 times larger than that of the opposing group, for entrants to Asia and Oceania it is 3.97 times larger and for entrants to America it is 4.12 times larger. Furthermore, this holds true when looking at the sizes of export revenues to other analyzed continents, with the firms that enter new continents having on average more revenue from all continents that they have started to export to, as opposed to the average of those who did not export to the analyzed continent during the analyzed period. This might hint that the firms that do not enter new continents have not yet depleted their avenues of growth from the continents which they already have a presence on.

A similar pattern to the revenues emerges when looking at the number of markets an entrant or a non-entrant has. The mean number of markets for a firm that enters a market on a new continent is larger than that of a firm that does not. For example, looking at the entries to African markets, the mean number of markets an entrant had in Europe was 10.29, in America 1.07 and in Asia and Oceania 2.3 while the mean of markets of those exporters that did not have any exports to African markets during the time period included in the study was 2.49 markets in Europe, 0.08 markets in America and 0.15 markets in Asia and Oceania. This pattern holds true for entries into Asian and Oceanian and American markets as well.

Looking at the revenue proportions of different continents to the total export revenue, the firms that have a higher proportion of their export revenue from markets outside of the home continent tend to be in the group that expands to new markets outside of Europe. The only variable where on average the entrants have a lower number from the opposing group is the proportion of export revenues that come from the home continent of Europe. However, the means for the proportion of European exports to total exports is relatively similar for both groups with the entrants to Africa having 67% of its exports from European markets as opposed to the opposing group having 78.7% of its exports from Europe. Furthermore, the numbers for entrants to Asia and Oceania and America and their opposing groups are more similar with the percentages being 72% versus 79% and 71% versus 79% respectively.

### **3. Results and discussion**

#### **3.1. Predictive models**

The independent variables from Table 1 were used to create models for predicting a firm's expansion onto a new market. Three methods were used for predictive models, and they were neural networks, decision trees and binary logistic regression using all variables together. To ensure the robustness of the results, a set of five neural networks and five

decision trees were created. The results presented here are the average of these five. The predictive properties of the variables were validated with the models correctly categorizing at least 95% of firms as either exporter or non-exporter on a given continent.

**Table 3.** Prediction accuracies of different models

	Entry to an African market	Entry to an Asian or Oceanian market	Entry to an American market
Binary logistic regression	96.1%	95.6%	94.8%
Neural networks	95.62%	97.3%	96.85%
Decision trees	94.94%	96.79%	96.14%

Source: compiled by the author

As seen from the table (Table 3) all predictive models were largely similarly precise in predicting the entry to a market. No pattern of one method being clearly superior to another emerges, with machine learning methods being better for predicting the entry to an Asian and Oceanian or American market and binary logistic regression being more accurate in predicting the entry to Africa. All methods had prediction accuracies of at least 95% and the best model for categorizing companies was found to be neural networks for entry into Asia and Oceania, having an accuracy over 97%. Neural networks also performed better than other methods for entry to an American market, and if ignoring the binary logistic regression, also to an African market.

**3.2. Importance of independent variables**

Three sets of binary logistic regressions with 10 variables each, that means 30 runs altogether, were carried out to check whether there is effect for a variable on the decision to export to a given continent, what the direction of the effect is and what are the predictive properties of each variable. There are 13 variables in total but for each of the models, the indicators for the region that is analyzed are excluded, because they would be null and thus there would be no reason in including them in the analysis. Due to the possibility of multicollinearity between the variables, each independent variable was checked independently as opposed to checking them all at once.

The results from binary logistic regressions are in Appendix 2 where “+“ means the increase in the value of independent variable increasing the likelihood of new continent entry,

while in case of „-,“ the opposite occurs. All of the predictors had a positive effect on expanding to continents outside of Europe except for one - Europe as a proportion of exports - which had a negative effect in all results. Indeed, analyzing the data we can find that it has a dense concentration of exporters that only export to a few countries in Europe with there being 13,601 observations out of 25,291 with exports to only 1-4 European markets in total and no other exports. Such a finding that a significant part of a country’s exporters export to only a few markets, and only to markets close to home country is in line with findings about other countries’ exporters (i.e. Eaton et al, 2007; Fichera and Shevtsova, 2020).

The reason for this could be a type of firm that is satisfied with its markets in the EU and thus has no need to expand to further markets on other continents. Furthermore, looking at the means of this variable for entrants and non-entrants could paint a picture of another type of firm: with both groups having a similar proportion of European exports to total exports and the non-entrant having multiple times less markets and revenue from exports on average (see Table 2), the non-entrant might not have yet depleted the demand for its products from its extant markets while it is producing its products at a level that is near the firm’s full manufacturing capacity.

As indicated in Appendix 3, Appendix 4 and Appendix 5, analyzed variables were individually statistically significant. Appendix 3, Appendix 4 and Appendix 5 also include the coefficients of different variables included in the analysis. The variables were ranked according to their individual predictive accuracies using binary logistic regressions (see Table 5) as is common in internationalization literature (e.g. Fernandez-Alles et al., 2023). With the creation of prediction models a list of most influential predictors was also created for different machine learning models by their normalized importance to the models (Table 5). Normalized importance is not the predictive accuracy of a variable using the method but rather the sensitivity of the variable as an input to the given model (Ju et al., 2019). Ranking these predictors by normalized importance to the model creates a somewhat different list of most important variables as opposed to using binary logistic regression with there still being many similarities. The reason for these differences is that some descriptive characteristics of the variables are explained by other variables in machine learning algorithms.

Table 5. Ranking of different independent variables

	Market that is entered	Export to Africa	Proportion of African exports	Number of African markets	Export to Asia and Oceania	Proportion of Asian and Oceanian exports	Number of Asian and Oceanian markets	Export to America	Proportion of American exports	Number of American markets	Export to Europe	Proportion of European exports	Number of European markets	Total export
Predictive accuracy	African market			5 / 78.1%	6 / 77.5%	4 / 78.6%	8 / 74.2%	7 / 74.2%	9 / 74.2%	3 / 80.9%	10 / 66.4%	2 / 81.6%	1 / 84.5%	
	Asia or Oceanian market	10 / 51.5%	8 / 51.5%	9 / 51.5%			7 / 61.8%	6 / 61.8%	5 / 61.9%	2 / 74.6%	4 / 65.2%	3 / 74.1%	1 / 80.3%	
	American market	10 / 53.6%	9 / 53.6%	8 / 53.6%	5 / 65%	4 / 65%	6 / 65%			2 / 74.7%	4 / 62.4%	3 / 74.7%	1 / 78%	
Normalized importance for neural networks	African market			7 / 42.04%	4 / 43.54%	8 / 35.92%	9 / 33%	3 / 52.74%	10 / 27.32%	2 / 65.98%	4 / 43.3%	6 / 43.04%	1 / 100%	
	Asia or Oceanian market	10 / 32.28%	5 / 59.38%	9 / 48.38%			8 / 48.44%	4 / 61.90%	7 / 51.20%	3 / 67.48%	2 / 84.12%	6 / 52.66%	1 / 87.84%	
	American market	9 / 36.12%	8 / 40.16%	7 / 44.14%	6 / 52.28%	4 / 59.58%	10 / 27.28%			2 / 79.14%	4 / 67.42%	5 / 52.70%	1 / 100%	
Normalized importance for decision trees	African market			6 / 46.53%	5 / 47.34%	7 / 46.48%	9 / 66%	8 / 1.26%	10 / 0.60%	4 / 74.68%	5 / 93.83%	3 / 79.02%	1 / 100%	
	Asia or Oceanian market	8 / 0.03%	9 / 0.02%	10 / 0.01%			6 / 0.07%	5 / 1.16%	7 / 0.05%	4 / 60.49%	1 / 100%	3 / 60.81%	2 / 81.09%	
	American market	10 / 0.21%	8 / 0.24%	9 / 0.23%	7 / 1.01%	5 / 1.48%	6 / 1.05%			3 / 59.25%	1 / 100%	4 / 56.25%	2 / 81.38%	

Source: compiled by the author

Note: each cell reflects rank and accuracy or normalized importance of specific independent variable (in column) to predict the continent entry (in row), by using the prediction method noted in the first column

### 3.2.1. Entry to Africa

Analyzing the 87 entries into Africa, the variable with the best predictive accuracy individually was the total revenue from exports, which had an individual predictive accuracy of 84.5% (see Table 5). This means that the more revenue a firm receives from exports overall, the likelier it is to start exporting to Africa. This finding is in line with previous findings from other authors that state that the size of a company increases its probability of entering new markets (e.g. Buch et al., 2010; Razzolini and Vannoni, 2011). Total export was also the most important variable to the machine learning methods used in predictions when measured by the normalized importance.

In the list of variables by their predictive properties, the total export was followed by the number of export markets a firm has in Europe and the amount of export revenue it has in Europe, with the accuracies being 81.6% and 80.9% respectively (Table 5). This was followed closely by the number of markets in Asia and Oceania and the revenue from that subgrouping of markets. This could indicate that as a company increases its revenues and markets in Asia, Oceania and Europe, it might reach a saturation point to its products, meaning in order to keep growing it has to enter new markets. The third variable relating to Europe, the proportion of European export revenues to total revenues from export, was however the variable with the least predictive accuracy with an accuracy of 66.4% (Table 5). For this analyzed dependent variable, all independent variables relating to America were in the latter half in the list according to their individual predictive properties, meaning they had relatively poor predictive properties (Table 5).

### 3.2.2. Entry to America

The predictive accuracies different variables had on the 287 entries to American markets were similar to the ones for entries into African markets with the variable with the best predictive accuracy of 80.3% being total revenues from exports (Table 5). In entries to America this was then followed by the number of European markets and the revenue from European markets, both of them having the same accuracy of 74.7% (Table 5). Hence, the three most precise variables for prediction are the same for entries into both American and African markets.

Similarly to the entries into African markets, in the entries to American markets, the first three variables were followed by the three variables relating to Asia and Oceania. After rounding the percentages, the aforementioned variables have the same precision of 65% in predicting the entries into America (Table 5). As mentioned before, the proportion of European revenue from total export revenue was the variable with the least predictive

accuracy for entries into Africa. This is where the models differ somewhat, with this variable being the seventh most precise in predicting the entries into American markets. The variables relating to Africa have the least predictive precision on their own in predicting entries into American markets with the binary logistic regression model being able to correctly categorize 53.6% as new continent entrants (Table 5).

### **3.2.3. Entry to Asia and Oceania**

The results of entries into Asian and Oceanian markets did not considerably differ from the predictions into Africa and America discussed earlier. As with the entries into other continents, the most important variables on 380 entries into Asian and Oceanian markets to entries individually analyzed were, in order of importance, the export to all markets, export to European markets and the number of markets a firm has in Europe, with prediction accuracies of 80.3%, 74.6% and 74.1% (Table 5). Once again, this finding indicates that the size of a firm as an exporter is a good variable in predicting entries into a market that is outside of the continents a firm already exports to, possibly indicating greater capabilities of exporting and expanding its export portfolio.

The proportion of European exports to total exports was more important in predicting entries into Asian and Oceanian markets ranking-wise than to other markets and it was 65.2% (Table 5). However, this variable was the most important in decision trees and the second most important in neural networks as measured by normalized importance (Table 5). This was then followed by different independent variables describing a firm's operations in America and then Africa. When grouped by continents, i.e. variables relating to America or Africa, the variables had noticeably similar predictive properties, with the American variables having a prediction accuracy of around 61.8% and the variables derived from a firm's operations in Africa being able to correctly predict a firm as an exporter or not around 51.5% of the time (Table 5).

A further aspect to point out in models of entries into Asian and Oceanian or American markets, is that a significant part of their decision trees are built using only four variables. These four variables are total export, export to Europe, number of European markets and the proportion of European exports to total exports. When looking at the normalized importance of different variables to the decision trees predicting entries into Asian and Oceanian and American markets, there is a steep drop-off in normalized importance after the variable of European export revenues and number of European markets, respectively (Table 5).

### **3.3. Theoretical implications and limitations of the study**

This study has multiple theoretical implications. Different internal capabilities of a firm facilitate its ability to internationalize (Raymond et al, 2014; Chen et al, 2016) and the size of a firm has been used as a proxy of its internal capabilities (Tang et al, 2021). This study used revenues from export as an indicator of the size of a firm as an exporter and this yielded positive results on further expansion to new continents. The findings that previous market coverage is a good indicator of further expansion to new export markets (Fichera and Shevtsova 2020; Roberts and Tybout, 1997) were also confirmed in this study. Export revenues from different continents have not been used to create prediction models on further expansion to new continents in previous studies, but in this study they had high predictive abilities with the best being revenues from the home continent of Europe. The proportion of export revenues from a continent to the total export revenues also had predictive properties, however these variables were not as good as other sets of variables, i.e. the number of markets or export revenue from a specific continent. The study was able to produce highly accurate models of predicting entries into markets on new continents.

The findings brought out in this paper have some limitations that need to be considered when generalizing and interpreting the results. The predictive models and their properties are only valid for the given dataset in the sense of Estonian exporters during the years 2009-2019. The findings cannot be generalized to other countries because the firms that operate in a small nation like Estonia could be somewhat different from the set of exporters that operate in another country due to the size of the home market and the human capital a manufacturer can draw upon. Furthermore, Estonia is in the EU, making the possible expansion to new markets easier by either being able to export without tariffs inside the EU or as a part of a trading block outside the EU. However, analysis of different trade barriers and similar possible obstructions to trade and exports is beyond the scope of this study. Further research on other markets is needed to validate the findings.

### **Conclusion**

This paper had the goal of creating prediction models of exporting companies entering new continents and analyzing the models, to find the best independent variables included in them. In the extant literature, entries into markets have been analyzed using a diverse set of methods, including different types of regression analysis, and neural networks and decision trees for prediction. The independent variables used for analyzing entries into new markets included both financial and non-financial information and were varied, with size, previous market coverage and efficiency positively indicating an expansion into new export markets.

The empirical study of this thesis was carried out using data from the Business Registry of Estonia during the years 2009-2019 and included 25,291 observations. The dataset was skewed towards companies that export to European markets as opposed to markets in other parts of the world. This is in line with the theoretical standpoint, findings from other studies and general Estonian export statistics. The study was carried out in three sets, each involving 10 independent variables with there being 13 variables in total included in the studies. Variables came from three domains: the export revenue in a continent, the proportion of export revenue from each continent to the total export revenue and the number of markets on a continent. Models were created to predict the entries to a market in Asia and Oceania, America or Africa using binary logistic regression, neural networks, and decision trees. Prediction models created in the empirical part of this paper were highly accurate in predicting entries with the different models having accuracy of at least 95%.

As shown in this paper, all the analyzed variables had at least some predictive characteristics to explain when companies start exporting to a given continent. The coefficients were positive for all variables, except for exports to European markets as a proportion of all exports - a positive coefficient meaning that the larger a given variable was, the more likely a firm was to enter a market on a continent it was not exporting to before. The predictors were found to be statistically significant.

Overall, the most important variables for predicting entries into markets on new continents in different methods were the size of a company as an exporter, as measured by its total exports, followed by the different variables relating to Europe, meaning its revenue from European export markets, and its revenue from European export markets as a proportion of total export revenue from all export markets and the number of markets in Europe a firm exported to.

There were similarities in the ranking and individual predictive properties of variables relating to either of the four continents. The best predictors were variables relating to a firm's activities in European export markets and the worst were variables related to the African markets. Furthermore, ranking different variables, the export variables of Asia and Oceania followed European variables in their predictive properties, after which variables related to America followed.

Despite the close similarities of variables relating to a specific continent, a pattern emerges amongst these variables. When comparing the variables relating to different continents amongst themselves, meaning, which variable relating to a specific continent achieved the best prediction accuracy, and their importance in different models, the export

revenue and the number of markets on a continent tended to be better at predicting an entry to a new continent compared to the proportion of revenue from that continent to total export.

The findings of this thesis can benefit companies that want to compare their life cycle with their competitors, and to analyze the possibility of expanding to further markets, as comparatively they might also exhibit the potential in doing so. Furthermore, these findings can be advantageous for different governmental organizations trying to aim governmental aid more precisely with the intention of promoting exports to a specific set of markets. This thesis can be further developed by expanding the scope of the analysis and taking into account more financial and non-financial metrics or adding a qualitative aspect to the analysis to better understand the reasons for expanding into markets that are further from the home continent.

### References

1. Aitken, B. Hanson, G. H. and Harrison, A. E. (1997). Spillovers, foreign investment, and export behavior. *Journal of International Economics*, Vol 43, pp. 103-132.
2. Albornoz, F., Calvo-Pardo, H. F., Corcos, G. and Ornelas, E. (2012). Sequential exporting. *Journal of International Economics*, Vol. 88, Issue 1, pp. 17-31. DOI: 10.1016/j.jinteco.2012.02.007.
3. Bernini, M., Du, J. and Love, J. H. (2016). Explaining intermittent exporting: Exit and conditional re-entry in export markets. *Journal of International Business Studies*, Vol. 47, Issue 9, pp. 1058–1151.
4. Blum, B. S., Claro, S. and Horstmann, I. J. (2013). Occasional and perennial exporters. *Journal of International Economics*, Vol. 90, Issue 1, pp. 65-74. DOI: 10.1016/j.jinteco.2012.11.002.
5. Bouncken, R., B., Schuessler, F. and Kraus, S. (2015). The Theoretical Embedding Of Born Globals: Challenging Existing Internationalization Theories. *International Business & Economics Research Journal*, Vol. 14, No. 1.
6. Brenton, P., Ferrantino, M., J. and Maliszewska, M. (2022). Reshaping Global Value Chains in Light of COVID-19: Implications for Trade and Poverty Reduction in Developing Countries. Washington, DC: World Bank. doi:10.1596/978-1-4648-1821-9.
7. Buch, C. M., Kesternich, I., Lipponer, A. ja Schnitzer, M. (2010). Exports versus FDI revisited: Does finance matter?. *Deutsche Bundesbank Discussion Paper, Series 1: Economic Studies*, No 03/2010. doi: 10.2139/ssrn.2785367.
8. Chen, J. and Sousa, C. M. P. and Xinming, H. (2016). The determinants of export performance: a review of the literature 2006-2014. *International marketing review*, Vol. 33, Issue 5, pp. 626-670.
9. Cuervo-Cazurra, A. (2011). Selecting the country in which to start internationalization: The non-sequential internationalization model. *Journal of World Business* Vol. 46, Issue 6, pp. 426–437. DOI: 10.1016/j.jwb.2010.10.003.
10. da Silva, R. V., Lopes, A. F., Carvalho, H. and Duarte, J. M. (2023). When Firms Go International: Deliberate or Random?. *Review of International*

Business and Strategy, Vol. 33, Issue 2, pp. 219-245. DOI: 10.1108/RIBS-01-2021-0020.

11. Dang, C., Li, C. and Yang, C. (2018). Measuring firm size in empirical corporate finance. *Journal of Banking & Finance*, Vol. 86, pp. 159-176. DOI: 10.1016/j.jbankfin.2017.09.006
12. Dominguez, N. and Mayrhofer, U. (2017). Internationalization stages of traditional SMEs: Increasing, decreasing and re-increasing commitment to foreign markets. *International Business Review*, Vol. 26, Issue 6, pp. 1051-1063. DOI: 10.1016/j.ibusrev.2017.03.010.
13. Draz, U., Jahanzaib, M. and Asghar, G. (2015). Identification of Determinants for Globalization of SMEs using Multi-Layer Perceptron Neural Networks. *Mehran University Research Journal of Engineering & Technology*, Vol. 35, No. 1, pp. 39-52.
14. Eaton, J., Eslava, M., Kugler, M. and Tybout, J. (2007). Export Dynamics in Colombia: Firm-Level Evidence. National Bureau of Economic Research, Working Paper 13531. DOI: 10.3386/w13531.
15. Eaton, J., Kortum, S. and Kramarz, F. (2011). An Anatomy of International Trade: Evidence from French Firms. *Econometrica*, Vol. 79, No. 5 (September 2011), pp. 1453-1498. DOI: 10.3982/ECTA8318.
16. Estonian Business Registry. (2008). Estonian Classification of Economic Activities. Accessed 10.04.2023, <https://www.rik.ee/et/e-ariregister/emtak-tegevusalad>. Reference inside the text: *EMTAK, 2008*.
17. Fernandez-Alles, M., Vissak, T. and Lukason, O. (2023). Which variables predict the internationalization type of academic spin-offs?. *Heliyon*, Vol. 9, Issue 2, e13172. DOI: 10.1016/j.heliyon.2023.e13172.
18. Fernhaber, S. A., Gilbert, B. A. and McDougall, P. P. (2008). International Entrepreneurship and Geographic Location: An Empirical Examination of New Venture Internationalization. *Journal of International Business Studies*, Vol. 39, Issue 2, pp. 267-290.
19. Fernhaber, S. A., Gilbert, B. A. and McDougall, P. P. (2008). *Journal of International Business Studies*, Vol. 39, No. 2, pp. 267-290.
20. Fichera, E. and Shevtsova, Y. (2020). Sunk exporting costs and export market coverage. *The Manchester School*, 88, pp. 599–616. DOI: 10.1111/manc.12311.

21. Gripsrud, G., Hunneman, A. and Solberg, C. A. (2023). Speed of internationalization of new ventures and survival in export markets. *International Business Review*. DOI: 10.1016/j.ibusrev.2023.102121.
22. Hussien, M. S. (2022). Government Financial Incentive and Firm Performance: Evidence from Turkey. *International Journal of Economics, Management and Accounting*, Vol. 30, No. 2, pp. 259-283.
23. IBM. (2019). IBM SPSS Neural Networks 26. Accessed 04.05.2023, [https://www.ibm.com/docs/en/SSLVMB\\_26.0.0/pdf/en/IBM\\_SPSS\\_Neural\\_Network.pdf](https://www.ibm.com/docs/en/SSLVMB_26.0.0/pdf/en/IBM_SPSS_Neural_Network.pdf).
24. IBM. (2023). What is logistic regression?. Accessed 04.05.2023, <https://www.ibm.com/topics/logistic-regression>.
25. Johanson, J. and Vahlne, J.-E. (1977). The Internationalization Process of the Firm—A Model of Knowledge Development and Increasing Foreign Market Commitments. *Journal of International Business Studies*, Vol. 8, Issue 1, pp. 23-32. DOI: 10.1057/palgrave.jibs.8490676.
26. Johanson, J. and Vahlne, J.-E. (2009). The Uppsala internationalization process model revisited: From liability of foreignness to liability of outsidership. *Journal of International Business Studies*, Vol. 40, Issue 9, pp. 1411-1431. DOI: 10.1057/jibs.2009.24.
27. Johanson, J. and Vahlne, J.-E. (2010). Markets as networks: Implications for strategy-making. *Journal of the Academy of Marketing Science*, Vol. 39, Issue 4, pp. 484-491. DOI: 10.1007/s11747-010-0235-0.
28. Johanson, J. and Vahlne, J.-E. (2013). The Uppsala model on evolution of the multinational business enterprise - from internalization to coordination of networks. *International Marketing Review*, Vol. 30, No 3, pp. 189-210. DOI: 10.1108/02651331311321963.
29. Johanson, J. and Vahlne, J.-E. (2017). From internationalization to evolution: The Uppsala model at 40 years. *Journal of International Business Studies*, Vol. 48, Issue 9, pp. 1087-1102. DOI: 10.1057/s41267-017-0107-7.
30. Ju, E., Kim, K. H., Chang, K. H., Shim, J. B., Kim, C. Y., Lee, N. K., Lee, S. and Park, C. G. (2019). Artificial Neural Network-Based Variable Importance Analysis of Prognostic Factors Related to Radiation Pneumonitis in Patients with Lung Cancer: Preliminary Study. *Journal of the Korean Physical Society*, Vol. 75, No. 4, pp. 277-282. DOI: 10.3938/jkps.75.277.

31. Knight, G. A. and Cavusgil, S. T. (2004). Innovation, Organizational Capabilities, and the Born-Global Firm. *Journal of International Business Studies*, Vol. 35, Issue 2, pp. 124-141.
32. Krenker, A., Bešter, J. and Kos, A. (2011). Introduction to the Artificial Neural Network. K. Suzuki (editor), *Artificial Neural Networks - Methodological Advances and Biomedical Applications* (pp. 3-18). Rijeka: InTech.
33. Lawless, M. (2009). Firm export dynamics and the geography of trade. *Journal of International Economics*, Volume 77, Issue 2, pp. 245-254. DOI: 10.1016/j.jinteco.2009.01.006.
34. Lu, J., Lu, Y., Sun, Y. and Tao, X. (2017). Intermediaries, Firm Heterogeneity and Exporting Behaviour. *The World Economy*, Vol. 40, Issue 7, pp. 1381-1404. DOI: 10.1111/twec.12423.
35. Lu, J., Lu, Y., Sun, Y. and Tao, Z. (2017). Intermediaries, Firm Heterogeneity and Exporting Behaviour. *The World Economy*, Vol. 40, Issue 7, pp. 1381-1404. DOI: 10.1111/twec.12423.
36. Lukason, O. and Vissak, T. (2020). Export behavior and corporate governance. *Review of International Business and Strategy*, Vol. 30, No. 1, pp. 43-76. DOI: 10.1108/RIBS-07-2019-0097.
37. Lukason, O., Vissak, T. and Segovia-Vargas, M.-J. (2021). How Does Managerial Experience Predict the Internationalization Type of a Young Firm?. *IEEE Access*, Vol. 9, pp. 18148 – 18166. DOI: 10.1109/ACCESS.2021.3054116.
38. Micocci, F. and Rungi, A. (2022). Predicting Exporters with Machine Learning. September 16, 2022 version. DOI: 10.2139/ssrn.3882811.
39. Morales, E., Sheu, G. and Zahler, A. (2019). Extended Gravity. *Review of Economic Studies* Vol. 86, Issue 6, pp. 2668–2712. DOI: 10.1093/restud/rdz007.
40. Morales, J. S. L. (2020). Internationalization: An Analysis of 26 Definitions. *Intersticios Sociales*, No. 19, pp. 87-99. DOI: <https://doi.org/10.55555/IS.19>.
41. O’Grady, S. and Lane, H. W. (1996). The Psychic Distance Paradox. *Journal of International Business Studies*, Vol. 27, Issue 2, pp. 309-333.

42. Ojala, A. (2015). Geographic, cultural, and psychic distance to foreign markets in the context of small and new ventures. *International Business Review*, Vol. 24, Issue 5, pp. 825-835. DOI: 10.1016/j.ibusrev.2015.02.007.
43. Onah, B. C., Amuka, J. I., Asogwa, F. O., Onuigbo, F. N. and Ezeudeka, C. E. (2022). Does Incentive Improve Export Performance of Firms? An Evidence. *Journal of Economics, Finance and Accounting Studies*, Vol. 4, No. 1, pp. 437-444. DOI: 10.32996/jefas.2022.4.1.27.
44. Patel, P. C., Criaco, C. and Naldi, L. (2018). Geographic Diversification and the Survival of Born-Globals. *Journal of Management* Vol. 44 No. 5, pp. 2008–2036. DOI: 10.1177/0149206316635251
45. Patel, P. C., Criaco, G. and Naldi, L. (2018). Geographic Diversification and the Survival of Born-Globals. *Journal of Management*, Vol. 44, No. 5, pp. 2008–2036. DOI: 10.1177/0149206316635251.
46. Prange, C. and Verdier, S. (2011). Dynamic capabilities, internationalization processes and performance. *Journal of World Business*, Vol. 46, Issue 1, pp. 126-133. DOI: 10.1016/j.jwb.2010.05.024.
47. Raymond, L., St-Pierre, J., Uwizeyemungu, S. and Dinh, T. L. (2014). Internationalization capabilities of SMEs: A comparative study of the manufacturing and industrial service sectors. *Journal of International Entrepreneurship*, Vol. 12, pp. 230–253. DOI: 10.1007/s10843-014-0123-7.
48. Razzolini, T. and Vannoni, D. (2011). Export premia and subcontracting discount: Passive strategies and performance in domestic and foreign markets. *World Economy*, vol. 34, no. 6, pp. 984–1013.
49. Roberts, M. J. and Tybout J. R. (1997). The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs. *The American Economic Review*, Vol. 87, No. 4 (Sep. 1997), pp. 545-564.
50. Segura-Cayuela, R. and Vilarrubia, J. M. (2008). Uncertainty and Entry into Export Markets. Banco de España, Working Paper No. 0811. DOI: 10.2139/ssrn.1144564.
51. Song, Y.-Y. and Lu, Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry*, Vol. 27, No. 2, pp. 130-135. DOI: 10.11919/j.issn.1002-0829.215044.
52. Srhoj, S., Škrinjarić, B. and Radas, S. (2021). Bidding against the odds? The impact evaluation of grants for young micro and small firms during the

- recession. *Small Business Economics*, Vol. 56, Issue 1, pp. 83-103. DOI: 10.1007/s11187-019-00200-6.
53. Statistics Bureau of Estonia. (2023). *Statistics Estonia: Foreign Trade Application*. Accessed 14.05.2023, <https://valiskaubandus.stat.ee/profile/country/ee/?locale=en>. Reference inside of text: *Statistics Bureau of Estonia, 2023b*.
54. Statistics Bureau of Estonia. (2023). *Statistika andmebaas*. Accessed 12.04.2023, <https://andmed.stat.ee/et/stat>. Reference inside of text: *Statistics Bureau of Estonia, 2023a*.
55. Tallman, S. and Fladmoe-Lindquist, K. (2002). Internationalization, Globalization, and Capability-Based Strategy. *California Management Review* Vol. 45, No.1, pp. 116-135.
56. Tang, R. W., Zhu, Y., Cai, H. and Han, J. (2021). De-internationalization: A Thematic Review and the Directions Forward. *Management International Review*, Vol. 61, Issue 3, pp. 267–312. DOI: 10.1007/s11575-021-00446-x.
57. Townsend, J. D., Yenyurt, S. and Talay, M. B. (2009). Getting to Global: An Evolutionary Perspective of Brand Expansion in International Markets. *Journal of International Business Studies*, Vol. 40, No. 4, pp. 539-558.
58. Vissak, T. and Masso, J. (2015). Export patterns: Typology development and application to Estonian data. *International Business Review*, Vol. 24, Issue 4, pp. 652-664. DOI: 10.1016/j.ibusrev.2014.11.004.
59. Vissak, T., Francioni, B. and Freema, S. (2020). Foreign market entries, exits and re-entries: The role of knowledge, network relationships and decision-making logic. *International Business Review*, Vol. 29, Issue 1, pp. 1-22. DOI: 10.1016/j.ibusrev.2019.101592.
60. Welch, C. L. and Welch, L. S. (2009). Re-internationalisation: Exploration and conceptualisation. *International Business Review*, Vol. 18, Issue 6, pp. 567-577. DOI: 10.1016/j.ibusrev.2009.07.003.
61. Welch, D. E. and Welch, L. S. (1996). The Internationalization Process and Networks: A Strategic Management Perspective. *Journal of International Marketing* , Vol. 4, No. 3, pp. 11-28.
62. Welch, L. S. and Luostarinen, R. (1988). Internationalization: Evolution of a Concept. *Journal of General Management*, Vol. 14, No.2, pp. 34-55. DOI: 10.1177/030630708801400203

63. Ye, Z., Zhang, F. and Zhang, S. (2019). Export effect and influence mechanism of foreign ownership. *International Review of Economics & Finance*, Vol. 76, pp. 258-276. DOI: 10.1016/j.iref.2021.06.006.

Appendix 1.

Examples of empirical studies focusing on export-related dependent variables

Authors	Countries analyzed	Sample size	Years	Methods used	Dependent variable	Analyzed variables	Relevant findings
Lawless (2009)	Ireland	751 firms	2000-2004	Regression analysis	Number of export markets	Firm productivity; hierarchy of markets	Firm productivity affects the dependent variable. Weak evidence of hierarchy of markets. Market coverage of firms is dynamic.
Fichera and Shevtsova (2020)	Ukraine	3272 firms	2000-2006	Regression analysis	Number of export markets; exporting status	Number of export markets; size; productivity; intangible assets; foreign ownership; capital intensity	The number of export markets increases the probability to expand to new export markets. Foreign ownership increases the probability to export but does not affect the number of markets. Other variables increase both the likelihood to export and to have more export markets
Lukason and Vissak (2020)	Estonia	9530 exporting and 73619 non-exporting firms	2015	Binary logistic regression; ordinary least squares regression	9 different variables including exporting status, number of markets and revenue from European markets to total exports.	Different variables relating to the board, its composition and ownership of the firm.	Board size and maximum tenure on the board had positive coefficients with being an exporter. Results for majority owner were not statistically significant. All other variables negatively affected being an exporter.

Authors	Countries analyzed	Sample size	Years	Methods used	Dependent variable	Analyzed variables	Relevant findings
Buch et al. (2010)	Germany	187 086 firms with no international activities, 12 252 exporters and 11 876 with foreign direct investments	2002-2006	Bivariate probit model	Exporting status; having foreign direct investments	Size; cost efficiency; cash flow; debt ratio; share of fixed assets to total assets.	Size and cash flow had positive and fixed asset share and debt ratio negative impact on internationalization.
Eaton et al. (2011)	France	230 423 firms of which 34 558 are exporters	1986-1992	Simulation algorithm and regression analysis	Exporting status; distribution of sales across markets, domestic sales conditional on foreign market entry; export intensity.	Domestic revenue; export revenue; efficiency; hierarchies in a market	Most of the variation in market entry is attributable to efficiency.
Ye et al. (2019)	China	25 700 observations	2003-2016	Correlation analysis, regression analysis	Export propensity	Foreign ownership. Size, age, book value to market value, leverage, current ratio, productivity per employee and fixed assets to total assets as control variables.	Foreign ownership promotes firm exports. Control variables gave mixed results.

Authors	Countries analyzed	Sample size	Years	Methods used	Dependent variable	Analyzed variables	Relevant findings
Lu et al. (2017)	29 developing countries	12 679 firms of which 9008 do not export	2002-2006	Correlation analysis, regression analysis	Exporting status, exporting directly or through intermediaries	Output, employment, capital, output per worker, capital per worker, total factor productivity.	Productivity increases the probability to export.
Razzolini and Vannoni (2011)	Italy	1537 firms	1998-2003	Logistic regression analysis	Exporting status, exporting directly or indirectly	Productivity (total factor productivity), labor productivity, size	Exporters are larger in size and more productive by total factor and labor productivity.
Draz et al. 2015	Pakistan	132 observations		Neural networks (multilayer perception)	Exporting status	13 different variables including size, firm membership in a cluster, education of employees and variables relating to potential product	Model achieved accuracy of 69.4%. Most important variables were being registered as an “Original Equipment Manufacturer” in Pakistan and the size of the firm.
Micocci and Rungi (2022)	France	57 016 firms	2010-2018	Binary logistic regression, decision trees	Exporting status	52 different metrics including financial accounts, ratios, proxies, and company categories	The model using decision trees was deemed to be the best having achieved accuracy up to 90%. <sup>4</sup>

Source: compiled by the author

<sup>4</sup> The project is still ongoing as of 18.05.2023.

## Appendix 2.

Individually analyzed variables with the direction of the effect using binary logistic regressions

	Start of exporting to Africa	Start of exporting to Asia and Oceania	Start of exporting to America
Export to Africa		+	+
Proportion of African exports to total export		+	+
Number of African markets		+	+
Export to Asia and Oceania	+		+
Proportion of Asian and Oceanian export to total export	+		+
Number of Asian and Oceanian markets	+		+
Export to America	+	+	
Proportion of American exports to total export	+	+	
Number of American markets	+	+	
Export to Europe	+	+	+
Proportion of European exports to total export	-	-	-
Number of European markets	+	+	+
Total export	+	+	+

Note: plus sign meaning statistically significant and positive effect, minus sign meaning statistically significant and negative effect and an empty cell meaning that this variable was not used in this analysis; this table is based Appendixes 3, 4, 5, which indicate the coefficients and significances of logistic regressions with single variables for the prediction of specific continent entry

Source: compiled by the author

## Appendix 3.

Individually analyzed variables on the entry to an African market with coefficients and statistical significance

<b>Entry to an African market</b>	<b>B</b>	<b>Sig.</b>
Proportion of American export to total export	18.219	<,001
Proportion of Asian and Oceanian export to total export	14.647	<,001
Number of American markets	2.182	<,001
Number of Asian and Oceanian markets	1.649	<,001
Proportion of European export to total export	-1.300	<,001
Export to America	0.488	<,001
Number of European markets	0.380	<,001
Export to Asia and Oceania	0.366	<,001
Total export	0.083	<,001
Export to Europe	0.082	<,001

Note: B in the table is the beta coefficient of the variable and Sig. meaning the statistical significance or P-value

Source: compiled by the author

## Appendix 4.

Individually analyzed variables on the entry to an American market with coefficients and statistical significance

<b>Entry to an American market</b>	<b>B</b>	<b>Sig.</b>
Proportion of African exports to total export	20.405	<,001
Proportion of Asian and Oceanian export to total export	10.261	<,001
Number of African markets	2.953	<,001
Number of Asian and Oceanian markets	1.755	<,001
Proportion of European exports to total export	-0.774	<,001
Export to Africa	0.694	<,001
Number of European markets	0.415	<,001
Export to Asia and Oceania	0.387	<,001
Total export	0.097	<,001
Export to Europe	0.084	<,001

Note: B in the table is the beta coefficient of the variable and Sig. meaning the statistical significance or P-value

Source: compiled by the author

## Appendix 5.

Individually analyzed variables on the entry to an Asian or Oceanian market with coefficients and statistical significance

<b>Entry to an Asian or Oceanian market</b>	<b>B</b>	<b>Sig.</b>
Proportion of African exports to total export	18.462	<,001
Proportion of American export to total export	12.399	<,001
Number of African markets	3.585	<,001
Number of American markets	2.668	<,001
Export to Africa	0.854	<,001
Proportion of European exports to total export	-0.674	<,001
Export to America	0.625	<,001
Number of European markets	0.477	<,001
Total export	0.126	<,001
Export to Europe	0.101	<,001

Note: B in the table is the beta coefficient of the variable and Sig. meaning the statistical significance or P-value

Source: compiled by the author

## Kokkuvõte

### EESTI TÖÖSTUSETTEVÕTETE UUDE MAAILMAJAKKU SISENEMISE PROGNOOSIMINE

Rauno Juntson

Käesoleva magistritöö eesmärgiks on prognoosida eksportiva ettevõtte sisenemist uues maailmajaos asuvale turule ning tuua loodud mudeleid analüüsid väljaks selleks parimad sõltumatud muutujad. Töö struktuur jaguneb kolmeks osaks, teoreetiline osa, empiirilise uurimise metodoloogia tutvustus ning empiirilise uurimuse tulemused, millele lisanduvad sissejuhatus, kokkuvõte ja lisad.

Varasemalt avaldatud uurimistöodes on rahvusvahelistumist uuritud mitmesuguseid meetodeid kasutades, sealjuures on leitud, et ettevõtte sisemised võimekused aitavad kaasa selle rahvusvahelistumisele. Varasemalt ei ole kasutatud ettevõttepõhiseid kvantitatiivseid andmeid ettevõttele uues maailmajaos asuvale turule laienemise prognoosimiseks.

Töös läbiviidud uuringu aluseks oli Äriregistrilt saadud andmestik ettevõtete majandusaasta aruannete lisa kohta, kus on välja toodud ettevõtte käibe jaotumine turgude lõikes. Käesolevas töös jaotati turud neljaks maailmajaoks, Euroopaks, Ameerikaks, Aasiaks (sisaldab Okeaniat) ning Aafrikaks. Kõik andmestiku ettevõtted on töösusettevõtted ning andmestik sisaldab 25 291 vaatlust aastatest 2009-2019. Sisenemist prognoositi kolmes maailmajaos, Ameerikas, Aasias ja Ookeanias ning Aafrikas, olevale turule kasutades 10 erinevat sõltumatut muutujat. Kokku oli kasutusel 13 sõltumatut muutujat kolmest valdkonnast, naturaallogaritmide ettevõtte müügitulu igas maailmajaos, turgude arv igas maailmajaos ning iga maailmajaos müügitulu osakaal kogu eksportitulust. Kuivõrd iga mudel prognoosis sisenemist kindlasse maailmajakku, ei olnud kasutusel selle maailmajaoga seotud muutujad. Lisaks oli kasutusel kogu ettevõtte müügitulu väljaspool Eestit.

Prognoosimudelite ülesehitamiseks kasutati kolme meetodit, binaarne logistiline regressioon, närvivõrgud ja otsustuspuud. Loodud mudelid olid täpsed uutele turgudele sisenemise prognoosimisel ning omasid täpsust, mis ületas 95%. Kõige paremad muutujad ettevõtete uutesse maailmajagudesse sisenemise prognoosimiseks olid ettevõtte kogu müügitulu väljaspool Eestit ning seejärel muutujad, mis olid seotud ettevõtte Euroopa eksport-turgudega ehk ettevõtte müügitulu Euroopa eksportturgudel, Euroopa eksportturgude arv ning Euroopa eksportturgudelt saadud müügitulu osakaal kogu eksportturgudelt saadud müügitulust.

Magistritöö järeldused saavad olla kasuks erinevate organisatsioonide eksportimist propageerivate programmide ülesehitamisel, sest programmide täpsem sihtimine nendest

tõenäoliselt kasusaavatele ettevõtetele võimaldab kasutada piiratud ressursse efektiivsemalt. Lisaks saavad ettevõtete juhid võrrelda enda ettevõtte näitajaid töö leidudega ning seejärel analüüsida võimalust kaugematele turgudele laienemiseks, kuivõrd nende juhitaval ettevõttel võib olla potentsiaalselt võimekust seda teha.

Non-exclusive licence to reproduce thesis and make thesis public

I, Rauno Juntson

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

Predicting New Continent Entrance of Estonian Manufacturing Firms,

supervised by Oliver Lukason,

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.
3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.
4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

*Rauno Juntson*  
18.05.2023