UNIVERSITY OF TARTU

Faculty of Social Sciences

School of Economics and Business Administration

Gerda Hermann

PREDICTION OF SUBSIDISED FIRM FAILURE WITH FINANCIAL RATIOS, TAX ARREARS AND ANNUAL REPORT DELAYS

Master’s thesis

Supervisors: Senior Research Fellow Oliver Lukason (PhD) and Tõnis Tänav (PhD)

Tartu 2021

Name and signature of supervisor: ....

Allowed for defence on: ....

(date)

I have written this master's thesis independently. All viewpoints of other authors, literary sources and data from elsewhere used for writing this paper have been referenced.

……………………………………..

(author's signature)

Table of contents

Abstract

Studies about the failure of firms receiving government support are infrequent, especially in the agricultural sector. This thesis aims to predict the failure of firms that have received agricultural subsidies. Logistic regression is applied to create prediction models by using variables portraying financial performance, tax arrears, subsidies, and reporting delays. Statistical tests indicate the potential of variables from all those domains to distinguish failed and non-failed firms. The most accurate prediction model can detect around a quarter of failed firms, at the same time not misclassifying non-failed firms. The results can be implemented in practice by agencies providing agricultural support.

Keywords: failure prediction, subsidised firm failure, farm viability, reporting delays, tax arrears

CERCS: S181, S190, S192

1. Introduction

Agricultural sector depends highly on subsidies. If statistical methods can be adequate in predicting subsidised firm failure, then they may be useful instrument to help the Agricultural Registers and Information Board (ARIB) estimate failure risk. “Failure” has different meanings, while this study examines firm failure as a process which starts with post-support non-eligibility for subsidy and can end with firm closure. Problems that have started after subsidy provision time like reclamation, and/or ARIB control was not passed, and/or the company were not active are considered as failed. The prediction of failure plays critical role. If financial distress in agricultural businesses can be predicted with timely warnings, then appropriate action can be taken and losses reduced (Stulpinienė, 2011). Good applicant with low failure risk would be granted a subsidy, while a high risk applicant would be denied. Subsidies are funded in return to get positive net return for society.

The data used in this study is from The Agricultural Registers and Information Board (ARIB) - an Estonian paying agency of the European Union. This thesis aims to predict the failure of firms that have received agricultural subsidies. Logistic regression is used to examine financial and non-financial indicators which may influence the likelihood of agricultural firm failure.

The overall prediction accuracy of the unweighted logistic regression prediction model with variables from all domains is above 73%. This study generates prediction models that can relatively precisely predict non-failed firms, at the same time it can eliminate about quarter of failed firms. To the best knowledge of the author, this is the first study that predicts failure of firms that have received agricultural subsidies. The novel findings to the extant literature include that previous tax arrears, delays with annual reports and variables such as subsidy ratio and also firm age are different between failed and non-failed groups. The results of this paper could be used by ARIB in subsidies decisions.

This research has the four main sections. The next section is a review of literature, where previous studies about the failure of agricultural firms is described. This is followed by the second section of an overview of used dataset, variables and methods. Empirical part finishes with results and discussion. Conclusion section summarizes the research.

2. Review of literature

Failure prediction is widely researched but there isn´t a lot of failure prediction researches about agricultural firms. Agriculture is very important part of the economy not only for providing food, while it also gives employment opportunities to a very large part of the population.

Terms ‘financial failure’, ‘business failure’, and ‘bankruptcy’ are often used with the same meaning and a common theme for them is the ability to unable to pay liabilities that have become due and which ends up with a firm not being able to continue its operations (Gestel, Baesens, & Martens, 2010); companies default or enter to the insolvency status, which is usually final status before deletion (Klepac, & Hampel, 2017). Jolly, Paulsen, Johnson, Baum, & Prescott (1985, p. 1108) describe financial stress following: “Financial stress occurs when the capacity of an individual or firm or a specific sector of the economy to adjust to the forces causing stress is exceeded”. Pannell, Malcolm and Kingwell (2000) discusses farming practices in agricultures and points out that climate, crop diseases, soil types, crop species, irrigation, marketing policies and technology are connected by forming and changing the uncertainties of alternative farming practices.

There are some earlier researches about agricultural companies’ failure prediction such as Barney, Graves and Johnson (1999) who found that accounting data are useful for predicting farm debt failure. And Miller and LaDue (1988) who investigated dairy farm borrowers using logit model to discriminate between acceptable borrowers and also borrowers who have defaulted and found that financial measures of liquidity, profitability and operating efficiency indicates borrowers quality. Frank's work (1998) confirmed the importance of increasing return on equity, which reduces the likelihood of failure. Farm´s economy researchers says that there isn’t one indicator that works good for agricultural firms that tells if farm is viable or not. If there is an appropriate prediction model, Altman says (1968) that failure can be predicted two year prior to the event and after two years the discriminant model becomes unreliable in its predictive ability. In agricultural companies failure prediction models may not work well in long-term before failure (Klepac, & Hampel, 2017) because they pass through acute failure process, where failure risk becomes high very shortly before failure (Lukason, & Laitinen, 2019) also unforeseen external causes makes firms´ failure not always observable (Lukason, & Hoffman, 2015).

Larger farms are usually borrowing larger amount of money to leverage up and manage their operations, and also government payments positively improve their financial indicators (Katchova, 2010). The larger the farm, the smaller the probability of non-viability (Argilés, 2001). Farm space and economic size do not directly combine when comparing income and agricultural firm size, as different farming categories have contrasting relationships between area/income (European Commission, 2018). Limsombunchai, Gan and Lee (2005) study that analysed credit scoring model for agricultural loans found that higher gross income to total assets ratio indicates higher probability to fail and higher value of farm assets shows higher creditworthiness.

EU policy is supporting farmers to obtain revenues from activities that differs from usual agricultural activities. Agricultural firm diversification support is good for many beneficial reasons – for the farmers, for the sector sustainability and the added value it brings to economies in rural areas (European Commission, 2016). It is found that agricultural diversification positively affects viability (Barnes, Hansson, Manevska-Tasevska, Shrestha, & Thomson, 2015) that is for example renting out of farm building and fixed equipment, agricultural tourism and contract work (Hansson, Ferguson, & Olofsson, 2010). In addition to previously mentioned agricultural diversification possibilities Barnes, Hansson, Manevska-Tasevska, Shrestha and Thomson (2015) adds farm shops or other activities outside of the usual farming to diversification opportunities. In year 2013, 5.2% of all EU farms were diversified (European Commission, 2016).

Failure prediction models has mostly been used as multivariate approach because there isn’t one single reason for a firm failure (Altman, 1968). Veganzones and Severin (2020) says that prediction methods plays a key role in corporate failure forecasting and classify prediction methods based on the overview of literature into three broad group: single statistical, artificial intelligence and ensemble methods.

An overview and results from previous studies using different prediction models are presented in table 1. Analysed articles are composed of all articles found from Google Scholar, Emerald insight and ScienceDirect by keywords: agricultural failure, viability and financial stress. As can be seen from table 1, the logistic regression has predominated in recent failure studies, and is preferred for its simplicity and because it does not assume error terms (residuals) to be normally distributed (Savitha, & Kumar K., 2016). Ordinary Least Squares (OLS), neural network (NN), multinomial logit (MNL), support vector machine (SVM), decision trees and Adaptive Boosting prediction methods are also being used in analysing agricultural firms’ failure. Sample sizes are rather small and sample period is mostly short, generally two year period. The data that are used in analysed studies in table 1 are collected mostly from developed countries. Dependent variables are typically used as default (payment default; insolvency; non-viable) and years before insolvency.

Table 1

Overview of previous studies and variables that did make a positive effect

Author	Barney, Graves, & Johnson (1999)	Savitha, & Kumar K. (2016)	D´Antoni, Mishra, & Chintawar (2009)	Argilés (2001)	Klepac, & Hampel (2017)	Barnes, Hansson, Manevska-Tasevska, Shrestha, & Thomson (2015)
Purpose	To improve on the FmHA and PW models by developing and comparing 3 statistical models.	To find the factors influencing credit repayment behaviour.	Investigate the aspects that predict financial stress of young and beginning farmers.	To test whether accounting would significantly improve the prediction of farm non-viability.	To find out if it´s possible to predict financial distress 1–3 years ahead with a solid accuracy based on agriculture firms and to test the prediction accuracy of different classification methods.	Estimate the impact of diversification on the performance of farm businesses using measures of economic viability among farms.
Country	USA	India	USA	Spain	EU	Sweden and Scotland
Data years	1990-1992	2015	2004-2006	1989-1991	2009-2013	2000-2012
Sample size	244 (68 failures)	590	19 638	82(19 viable and 63 not viable).	250(188 active and 62 defaulted).	6044(Scotland) 8712(Sweden)
Prediction methods	OLS, logit, and the neural network.	Binary logistic regression and multinomial regression analysis.	Multinomial logit (MNL) model.	Logistic regression.	Logistic regression, support vector machines method with RBF ANOVA kernel, decision trees and Adaptive Boosting.	Multinomial logistic regressions.
Dependent variable	Debt payment on date.	Non-default or default in LR. In MNL ability to pay loan: standard, sub-	Log odds ratios(combining a farm's net farm income & solvency	Is defined on the basis of financial data from three years.	Years before businesses went default or fell into insolvency proceedings in 2014.	Short term and long term non-viable.

		standard, doubtful.	position): Marginal Income/ Favorable; Marginal Solvency/ Favorable; Vulnerable/ Favorable			
Research results	The accounting data (information in the current value balance sheet and projected cash basis income statement) are useful for predicting farm debt failure.	Borrower age, years of being bank client, yield of the crop, distance to bank branch, size and tenure of the loan and farm size is related to prompt loan repayment. Leverage and efficiency ratio. Loan amount and capital ratio increases the likelihood of default.	Size and year of operation, farmer's age, and farm type are significant determinants of financial stress. Off-farm income is important to farm financial performance and its balance sheet. The older the farmers gets the likelihood of being vulnerable (financially stressed) decreases.	Found that accounting-based variables added significant information to predict farm non-viability. The greater the utilized agricultural area and the more professionalized the farm, the more probability of being viable.	Liquidity, rentability and debt ratios. Even companies which did not default in reality can be predicted – they show serious financial difficulties, such as the insolvency management or the default of payments and the differences which are not shown for active companies. Highest accuracy achieved using the methods of Adaptive Boosting and Decision Trees.	Agricultural diversification positively affects viability. Tenanted farmers are more viable than owner-occupied farmers. The less favoured area variable is indicative of spatial disadvantage, which is a defining characteristic in both countries and for Sweden this positively relates to non-viability status.

Source: compiled by the author

Table 1 consist of six previous studies where agricultural firms´ failure have been investigated. Previous studies in table 1 displays different financial and non-financial variables that have been used and did make a positive effect in prediction. Non-financial information are used, such as borrower/farmer age (Savitha, & Kumar K., 2016; D´Antoni, Mishra, & Chintawar, 2009); years of banking/operation (Savitha, & Kumar K., 2016; D´Antoni, Mishra, & Chintawar, 2009); yield of the crop (Savitha, & Kumar K., 2016); size and tenure of the loan (Savitha, & Kumar K., 2016); agricultural diversification (Barnes, Hansson, Manevska-Tasevska, Shrestha, & Thomson, 2015); farm size (Savitha, & Kumar K., 2016; D´Antoni, Mishra, & Chintawar, 2009) and less favoured area for farming (Barnes, Hansson, Manevska-Tasevska, Shrestha, & Thomson, 2015). Table 1 displays that financial information is good for predicting firms viability and variables such as debt ratio (Savitha, & Kumar K., 2016; Klepac, & Hampel, 2017); liquidity (Klepac, & Hampel, 2017); leverage and efficiency ratio (Savitha, & Kumar K., 2016) and also profitability (Klepac, & Hampel, 2017) have been good in describing differences between failed and non-failed firms. It can be concluded that the financial data are good to predict the failure of agricultural firms’. Tax arrears as independent variables have never been used. Variables such as less favoured area and borrower’s age are not used in this dataset because in Estonia the same conditions apply to all agricultural enterprises and the data from the enterprises perspective are being analysed in this research. However, this study uses the age of the company at the time of application as one of the possible variable in predicting failure.

Previous studies analysed in table 1 also provide information on the values of variables by groups, i.e. if higher or smaller value describes more failed or non-failed groups. Loans with shorter terms have lower probability of default (Savitha, & Kumar K., 2016). A longer time period of the overdue debts and liabilities agreements describes failed companies not active ones (Klepac, & Hampel, 2017). Farmer, who has been longer associated with bank or operating longer then the probability level of good loan is higher (Savitha, & Kumar K., 2016). Off-farm income and agricultural diversification is important to farm financial performance (D´Antoni, Mishra, & Chintawar, 2009; Barnes, Hansson, Manevska-Tasevska, Shrestha, & Thomson, 2015). Smaller farm increases the probability of failure (D´Antoni, Mishra, & Chintawar, 2009; Savitha, & Kumar K., 2016) and greater the agricultural area the smaller the probability of failure (Argilés, 2001). A higher value for profitability decreases the probability of failure (Klepac, & Hampel, 2017). The debt ratio as the leverage displayer is higher for the non-viable agricultural companies than for the viable firms and liquidity as well as efficiency ratios are lower for non-viable agricultural enterprises (Klepac, & Hampel, 2017).

Barney, Graves and Johnson (1999) research compared different prediction methods accuracies and found that the neural network model is better than OLS and logit. Argilés (2001) achieved accuracies with logistic regression within viable group around 42% and in failed group around 90% in structural model; in accounting model 79% viable and ca 95% failed farms were classified correctly; in third model (involved structural and accounting variables) within viable farms 79% and failed farms 95%. Klepac and Hampel (2017) study achieved highest accuracy using the methods of Adaptive Boosting (in total accuracy, one year before failure was reported in failed group around 94%) and Decision Trees (in total accuracy, one year before failure was reported in failed group over 92%), and also their study shows that the failure prediction accuracies are higher when the time period is closer to business failure. The marginal effect of significant determinants in D´Antoni, Mishra and Chintawar (2009) research were small on average, which was explained with the used data years and pointed out that more meaningful information about financial stress might have been collected by using years previously to the farm crises. Savitha and Kumar K. (2016) binary logit model predicted 81% of the cases correctly.

3. Dataset, variables and methods

3.1. Dataset

The Agricultural Registers and Information Board (ARIB) is paying agency of the European Union. ARIB is an Estonian government agency that is responsible for organising the granting of national subsidies, European Union agricultural and rural development subsidies, European Maritime and Fisheries Fund subsidies, and market management subsidies (ARIB, 2021). ARIB is also in control for preserving national registers and other databases related to agriculture. One of the foundation of the process of economic and political integration in European Union is Common Agricultural Policy (CAP), where one of the nine objectives assigned to the CAP is to ensure a viable farm income (European Commission, 2018). One of the major concerns for ARIB is to know whether farms are viable or not, therefor being able to predict farm viability through prediction model for subsidies decisions can reduce failure risk.

Financial data and reporting delays dataset used in the analysis has been gathered from the Estonian Business Register and tax arrears dataset are from the Estonian Tax and Customs Board. Analysis also use dataset from ARIB, containing firms that had applied at least one of the three measures from ARIB in 2015-2020 period. Firm may be in the dataset more than once because they can apply in different year and for multiple measures in one year. Each year will be new information about firm performance, because variables used in this dataset are gathered previously before application time – minus one (t-1) and minus two years (t-2) back. Tax payment delays have been analysed over twelve month ends, the beginning of tax arrears are usually same as the t-2 time period that have been used for calculating financial variables and reporting delays for minus two years. All the financial data based variables (including subsidies variables) from annual reports are called “financial variables” and other domains are called “non-financial variables”.

Data on the performance of the Estonian farm enterprises in ARIB measures have been used for the research and are grouped as failed, coded as 1 and non-failed, coded as 0 in the prediction model. The sample obtained from three measures are agricultural diversification measure, investment support to improve farm performance and LEADER measure to support entrepreneurship. This study considered failed as a subsidy that had problems after the subsidy application. Failed group were classified when the subsidy had at least one of these conditions: reclamation, and/or ARIB control were not acceptable, and/or the company was no longer active. Non-failed group involved all the companies that hadn´t any of those previous problems and are now in finished status. The dataset of this paper includes 1353 observations which are gathered from 1077 unique companies. Sample consist of 361 failed and 992 non-failed firms who applied for subsidy measure. The sample reduces because not all of the next three values of variables are available - minus two year annual report delays variable (ARDCOUNT2), change in balance sheet total (CHBS2) and change in business revenues (CHBR2).

Firms have up to six months to submit their annual report to Estonian Business Register when the end of the 12-month financial period is over (Kohv, & Lukason, 2021). Businesses who apply for agricultural subsidy have to have previous two annual reports to include with application or already submitted reports in Estonian Business Register. In this study t is subsidy application year and all financial variables are based on t-1 and t-2 annual reports according to the previous and over the previous reports that are available on application time. It is assumed if annual report aren´t already submitted in Estonian Business Register then company have to submit their previous year’s financial information earlier to ARIB. This also gives more relevant information about firm latest performances and prediction model are more accurate.

Information about taxes in Estonian Tax and Customs Board is accessible to all, companies in Estonia must pay taxes two times a month (on the 10^{th} and 20^{th} dates in each month) (Kohv, & Lukason, 2021). Businesses who apply for subsidy can´t have any current tax arrears. In this research tax debts are classified as total tax and deferred tax arrears and tax debts by the end of the months are considered. The delayed frequencies of two previous annual reports submitted to the Estonian Business Register are counted in variables ARDCOUNT (t-1) and ARDCOUNT2 (t-2). Non-financial variables and formulas in this study are based on Kohv and Lukason (2021) and Lukason and Andresson (2019) researches; financial variables are compiled on the basis of table 1.

3.2. Variables

There are 35 independent variables in this dataset that are included in this analysis, all these indicators are described in the table 2.

Table 2

Financial and non-financial variables' content, abbreviations and formulas for prediction model

Dependent variables: Failed and non-failed firms		
Independent variables:		
Financial variables	**Abbreviation**	**Explanation**
ARIB subsidies in total to balance sheet ratio (1)	TSUB	all previous year ARIB subsidies amount/balance sheet total
Subsidy ratio (1)	SUB	subsidy amount/balance sheet total
Profitability (2)	PBS; PBS2	net profit/balance sheet total
Profitability (2)	PBR; PBR2	net profit/business revenues
Solvency/capture structure (2)	EBS; EBS2	equity/balance sheet total
Share of long term loans (2)	LTLBS; LTLBS2	long-term loans/balance sheet total
Liquidity (2)	CASTLBS; CASTLBS2	(current assets - short-term liabilities)/ balance sheet total
Share of current assets (2)	CABS; CABS2	current assets/balance sheet total
Productivity (2)	BRBS; BRBS2	business revenues/balance sheet total
Change in balance sheet total (2)	CHBS; CHBS2	balance sheet total (t)/balance sheet (t-1)
Change in business revenues (2)	CHBR; CHBR2	business revenues (t)/business revenues (t-1)
Proportion of ancillary activities (2)	ANCA; ANCA2	(business revenues - sales revenue)/ business revenues
Size (2)	SIZEBS; SIZEBS2	ln(balance sheet total)
Size (2)	SIZEBR; SIZEBR2	ln(business revenues)
Non-financial variables	**Abbreviation**	**Explanation**
Age	AGE	age in subsidy application time

Delays with annual reports (3)	ARDCOUNT and ARDCOUNT2	report submitted-report due. Frequency in two year period.
If maximum tax arrears are >1 then natural logarithm of maximum tax arrears over twelve month ends in total (TMAX) and deferred (DMAX)* (4)	TMAX and DMAX	$= ln[max(x_1 \dots x_{12})]$
If maximum tax arrears are >1 then natural logarithm of median tax arrears over twelve month ends in total(TMED) and deferred (DMED)* (4)	TMED and DMED	$= ln[median\ (x_1 \dots x_{12})]$
Number of months ending with tax arrears of 1 euros or more over twelve month ends (4)	TCOUNT and DCOUNT	$= \sum_{k=1}^{12} TA_k$ Where $TA_k = \{1 \text{ if tax arrears } > 1; \text{else } 0\}$

Notes. * If <1, given value is 0. All variables marked with numbers in brackets are used in table 7 according to domains, except „Age“, which is only used in the combined model. Source: Author’s composition based on table 1 studies (Barney, Graves, & Johnson, 1999; Klepac, &Hampel, 2017, Savitha, & Kumar K., 2016; D´Antoni, Mishra, & Chintawar, 2009; Barnes, Hansson, Manevska-Tasevska, Shrestha, & Thomson, 2015; Argilés, 2001) and non-financial variables and formulas based on Kohv and Lukason (2021) and Lukason and Andresson (2019)

Financial distress of the farm business can be determined by inspecting long-run characteristics like profitability, liquidity, solvency, riskbearing ability (Jolly, Paulsen, Johnson, Baum, & Prescott, 1985). Several financial ratios such as profitability, solvency, share of long term loans, liquidity, share of current assets, productivity, proportion of ancillary activities and changes and sizes in balance sheet total and in business revenues are considered in models. Also firm age on subsidy application time and subsidy ratios are used to predict failure in this dataset.

3.3. Methods

SPSS software were used to analyse data. The classification in this dataset is the performance of business as NF (non-failed) or F (failed), which are used as a dependent variable in logit model. The binary variable takes the value 1 if the business failed and 0 otherwise.

Previous studies have found that neural network models are lacking comprehensibility (Zhou, Jiang, & Chen, 2003; Barakat, & Bradley, 2010; Foster, Zurada, & Barney, 2010) and that it´s not easy to expand the model into a decision support system for new cases – low transportability. Altman, Iwanicz-Drozdowska, Laitinen and Suvas (2020) says that logistic regression and neural networks are superior to other approaches. Logistic regression model have been used before in agricultural probability of default determining studies (Barney, Graves, & Johnson, 1999; Jouault, & Featherstone, 2006; Featherstone, Roessler, & Barry, 2006; Savitha, & Kumar K., 2016; Barnes, Hansson, Manevska-Tasevska, Shrestha, & Thomson, 2015).

Classical statistical (logistic regression, noted as LR) tool is used in this paper for composing the prediction models. LR is theoretically preferred to models such as the discriminant model because it is more robust in the estimation of parameters (Altman, Iwanicz-Drozdowska, Laitinen, & Suvas, 2017).

The LR is calculated using the following formula based on Altman, Iwanicz-Drozdowska, Laitinen and Suvas (2017) research:

$$p(Y = 1|X) = \frac{1}{1+e^{-L}} = \frac{1}{1+e^{-(b_0 + b_1 X_1 + \cdots + b_N X_N)}} \qquad (1)$$

where b_i (i = 0, …N) – coefficients

X_i (i = 1,…,N) – independent variables.

This study uses 35 independent variables in 5 domains to explain the relationship between dependent variable.

4. Results and discussion

Data were analysed using descriptive statistics that included financial and non-financial variables. Descriptive statistics of the financial variables can be seen in table 3.

Table 3

Descriptive statistics of independent financial variables

Variable	Non-failed				Failed				Total			
	N	Mean	Median	Std. Dev.	N	Mean	Median	Std. Dev.	N	Mean	Median	Std. Dev.
TSUB	992	0.08	0.00	0.124	361	0.08	0.01	0.141	1353	0.08	0.01	0.129
SUB	992	0.42	0.13	1.028	361	1.18	0.25	2.188	1353	0.62	0.15	1.470
PBS	992	0.11	0.07	0.212	361	0.12	0.07	0.221	1353	0.11	0.07	0.214
PBR	992	0.13	0.10	0.297	361	0.21	0.12	0.348	1353	0.15	0.10	0.314
EBS	992	0.63	0.65	0.252	361	0.62	0.62	0.278	1353	0.62	0.64	0.259
LTLBS	992	0.19	0.12	0.219	361	0.21	0.12	0.233	1353	0.20	0.12	0.223
CASTLBS	992	0.21	0.16	0.291	361	0.26	0.16	0.347	1353	0.22	0.16	0.307
CABS	992	0.39	0.33	0.274	361	0.43	0.37	0.307	1353	0.40	0.34	0.283
BRBS	992	1.17	0.68	1.446	361	1.01	0.59	1.381	1353	1.13	0.66	1.431
CHBS	992	1.43	1.10	1.188	361	1.67	1.11	1.752	1353	1.49	1.10	1.365
CHBR	992	1.49	1.13	1.375	361	1.62	1.05	1.859	1353	1.53	1.11	1.519

ANCA	992	0.18	0.04	0.240	361	0.18	0.02	0.249	1353	0.18	0.03	0.242
SIZEBS	992	11.93	12.00	1.526	361	12.05	12.21	1.635	1353	11.96	12.05	1.556
SIZEBR	992	11.59	11.54	1.467	361	11.52	11.55	1.649	1353	11.57	11.55	1.518
PBS2	992	0.11	0.07	0.248	361	0.14	0.07	0.278	1353	0.12	0.07	0.257
PBR2	992	0.14	0.10	0.307	361	0.19	0.12	0.361	1353	0.16	0.11	0.323
EBS2	992	0.61	0.64	0.290	361	0.60	0.64	0.328	1353	0.61	0.64	0.301
LTLBS2	992	0.19	0.10	0.226	361	0.18	0.04	0.239	1353	0.18	0.08	0.230
CASTLBS2	992	0.21	0.14	0.331	361	0.23	0.14	0.415	1353	0.21	0.14	0.355
CABS2	992	0.41	0.33	0.297	361	0.45	0.38	0.322	1353	0.42	0.34	0.304
BRBS2	992	1.21	0.69	1.544	361	1.13	0.65	1.579	1353	1.19	0.67	1.554
CHBS2	934	1.53	1.12	1.376	324	1.77	1.13	1.897	1258	1.59	1.12	1.530
CHBR2	934	1.71	1.12	1.923	324	2.08	1.21	2.419	1258	1.80	1.13	2.067
ANCA2	992	0.18	0.04	0.249	361	0.18	0.01	0.260	1353	0.18	0.03	0.252
SIZEBS2	992	11.71	11.80	1.624	361	11.73	11.91	1.895	1353	11.71	11.83	1.700
SIZEBR2	992	11.38	11.41	1.512	361	11.31	11.36	1.805	1353	11.36	11.40	1.595

Notes. Std. Dev. is Standard Deviation. Winsorizing has been used in case of all variables in order to avoid extreme values.

Source: own elaboration

Table 4 describes tax arrears means, medians and standard deviations in non-failed, failed and total groups.

Table 4

Descriptive statistics of independent non-financial variables

Variable	Non-failed				Failed				Total			
	N	Mean	Median	Std. Dev.	N	Mean	Median	Std. Dev.	N	Mean	Median	Std. Dev.
AGE	992	10.33	9.04	5.918	361	9.19	7.52	5.519	1353	10.03	8.82	5.834
ARDCOUNT1	992	0.24	0.00	0.427	361	0.39	0.00	0.489	1353	0.28	0.00	0.449
ARDCOUNT2	956	0.23	0.00	0.423	333	0.38	0.00	0.487	1289	0.27	0.00	0.445
TMAX	992	0.93	0.00	2.469	361	1.59	0.00	3.036	1353	1.11	0.00	2.647
TMED	992	0.25	0.00	1.326	361	0.39	0.00	1.587	1353	0.29	0.00	1.401
TCOUNT	992	0.49	0.00	1.669	361	0.93	0.00	2.321	1353	0.60	0.00	1.875
DMAX	992	0.71	0.00	2.127	361	1.39	0.00	2.829	1353	0.89	0.00	2.353
DMED	992	0.07	0.00	0.670	361	0.25	0.00	1.243	1353	0.12	0.00	0.865
DCOUNT	992	0.28	0.00	1.109	361	0.68	0.00	1.886	1353	0.39	0.00	1.371

Notes. Std. Dev. is Standard Deviation

Source: own elaboration

In addition to descriptive statistics there are Welch´**s** robust ANOVA and independent-sample median test being done to analyse the dataset. Table 5 documents the Welch´s robust ANOVA and independent-sample median test values of subsidies and financial variables.

Table 5

Welch´s robust ANOVA test and independent-sample median test values and results of independent financial variables

Variable	Welch´s robust ANOVA test Sig.	Independent-samples median test Sig.	Welch´s robust ANOVA test result (ANOVA)	Independent-samples median test result (ISMT)
TSUB	0.374	0.7		
SUB	0.000	0	+	+
PBS	0.398	0.889		
PBR	0.000	0.136	+	
EBS	0.555	0.399		
LTLBS	0.241	0.987		
CASTLBS	0.018	0.987	+	
CABS	0.025	0.317	+	
BRBS	0.060	0.02		+
CHBS	0.015	0.431	+	
CHBR	0.241	0.02		+
ANCA	0.801	0.225		
SIZEBS	0.237	0.106		
SIZEBR	0.529	0.987		
PBS2	0.098	0.79		
PBR2	0.020	0.262	+	
EBS2	0.752	0.635		
LTLBS2	0.377	0.01		+
CASTLBS2	0.347	0.725		
CABS2	0.039	0.106	+	
BRBS2	0.392	0.399		
CHBS2	0.038	0.747	+	
CHBR2	0.013	0.138	+	
ANCA2	0.725	0.068		
SIZEBS2	0.829	0.451		
SIZEBR2	0.470	0.725		

Notes. “+” defines that the p-value is less than 0.05

Source: own elaboration.

Welch´s robust ANOVA test (ANOVA) compares means (Fujikoshi, 1993) and independent-sample median test (ISMT) compares whether medians are significantly different (Sheskin, 2011). The p-values of ANOVA and ISMT being ≤0.05 in this thesis indicate that variable means/medians are significantly different through the failed and non-failed firms.

Author discusses that the probability of failure could be lower if firm have received subsidies before, on the other side that could refer that company depends highly on the amounts of the grants. ARIB subsidies in total to balance sheet ratio variable (TSUB) were not significantly different. Means and medians between two groups are similar and TSUB didn´t indicate failure in this study. Subsidy ratio variable (SUB) means and medians between failed and non-failed groups were significantly different. Descriptive statistics in table 3 shows that SUB means and medians are higher in failed group. This defines that failed firms are more likely to have bigger subsidy ratio (SUB). If the amount of subsidy is quite large in relation to agricultural firm income then this variable shows that subsidised firm is more likely to be failed. This is also logical, because if subsidy amount is multiple times bigger than firm's balance sheet in total, the firm applies for investments that actually exceed their financial capabilities.

Profitability variables PBS and PBS2 were not significantly different but the other profitability variables that were counted minus one year (PBR) and two years (PBR2) back showed that these variables means are significantly different between failed and non-failed groups. Importance of profitability is approved in agricultural firms' studies before (Klepac, & Hampel, 2017; Jolly, Paulsen, Johnson, Baum, & Prescott, 1985). Descriptive statistics in table 3 displays that PBR and PBR2 means are higher in failed group, which shows that firms may have increased some financial indicator on purpose. In this dataset solvency/capture structure variables EBS and EBS2 showed no significantly different means or medians between two groups. On the other hand Franks (1998) MNL model confirmed the importance of increasing returns to equity for decreasing probability of failure. Jouault and Featherstone (2006) binomial logit regression in agribusiness loans research found that as leverage increases, profitability decreases.

In this research, share of long term loans variable LTLBS2 was significant at the level of 0.01 in ISMT. LTLBS2 descriptive statistics displays that the medians are higher in non-failed group. This don´t follow previous studies findings, that firms with high long-term debts have higher probability of failure (Savitha, & Kumar K., 2016; Klepac, & Hampel, 2017). Author discusses that long-term loans are generally given to companies in a better financial position. It could also be explained that having long-term loans, there is less pressure to repay them quickly and the repayment is rather spread over a long period of time. Share of long term loans variable LTLBS, which are counted from minus one year annual report, shows no significantly differences between two groups.

Liquidity variable CASTLBS means are significantly different between failed and non-failed group. CASTLBS means are higher in failed group, although the value for the latter (0.26) does not exceed much the value (0.21) in the non-failed group. Previous study says that liquidity as well as efficiency ratios are lower for non-viable agricultural enterprises (Klepac, & Hampel, 2017). Previous studies support that high capital ratio increases the probability of default (Savitha, & Kumar K., 2016). Share of current assets variables (CABS and CABS2) were statistically significant in ANOVA, the means are higher in failed group. Productivity variable (BRBS) did make a positive effect in one year in ISMT but didn´t play any role in ANOVA Welch test and in minus two year (BRBS2). Lukason and Käsper (2017) statistical tests also did not indicate differences in productivity in government funded start-up firms.

A higher value of farm assets suggests a higher creditworthiness (Limsombunchai, Gan, & Lee, 2005). Changes in balance sheet total variables (CHBS, CHBS2) and changes in business revenue (CHBR, CHBR2) are significantly different between failed and non-failed group. Descriptive statistics in table 3 shows that CHBS, CHBS2 and CHBR2 means are bigger in failed group. ISMT indicates that CHBR medians were higher in non-failed group.

For trying to make difference from farm and non-farm income proportion of ancillary activities (ANCA) variable are used. ANCA means and medians didn´t show any differences in two groups. Previous studies showed that off-farm income and agricultural diversification are important to farm financial performance (D´Antoni, Mishra, & Chintawar, 2009; Barnes, Hansson, Manevska-Tasevska, Shrestha, & Thomson, 2015). One reason for explaining the ANCA variable similarity between groups could be the reason that one of the measures firm were applying for was agricultural diversification measure. Firm size can be measured in different ways, in this dataset size variables SIZEBS and SIZEBS2 that were based on balance sheet in total and also SIZEBR and SIZEBR that were based on business revenues, there were no significantly different means or medians between two groups.

When considering ANOVA and ISMT tests together, it can be said that non-failed and failed firms do not differ in respect to values of financial variables. Only subsidy ratio variable (SUB) means and medians between failed and non-failed groups were significantly different, and the p-value were 0.00. It is possible that variables that were significant only in one test (either in ANOVA or ISMT) there are different types of failed firms reflected.

Table 6 documents the Welch´**s** robust ANOVA and independent-sample median test values of independent non-financial variables.

Table 6

Welch´s robust ANOVA test and independent-sample median test values and results of independent non-financial variables

Variable	Welch´s robust ANOVA test Sig.	Independent-samples median test Sig.	Welch´s robust ANOVA test result	Independent-samples median test result
AGE	0.001	0.015	+	+
ARDCOUNT1	0.000	0	+	+
ARDCOUNT2	0.000	0	+	+
TMAX	0.000	0	+	+
TMED	0.129	0.054		
TCOUNT	0.001	0	+	+
DMAX	0.000	0	+	+
DMED	0.008	0.001	+	+
DCOUNT	0.000	0	+	+

Notes. “+” defines that the p-value is less than 0.05

Source: own elaboration

Agriculture needs significant initial investment to get started. Younger firms are usually low on capital and experience in farming. Both the Welch´s robust ANOVA test and the independent-sample median test found that firm age (AGE) are significantly different between two groups. Non-failed group has higher mean and median in AGE variable than failed group (showed in table 4). Failed groups are younger in operating years than non-failed group. This follows previous study findings that the likelihood of being financially stressed are higher for younger firms (Savitha, & Kumar K., 2016).

It can also be concluded that the medians and means of annual reports delays variables (ARDCOUNT1 and ARDCOUNT2) across the failed and non-failed groups are significantly different. Failed firms are more likely to submit annual reports in delay. This also follows the recent study findings that financially distressed SMEs are more likely to submit their annual reports later than should have, factors for that could be behaviour of managers and corporate governance characteristics (Lukason, & Camacho-Miñano, 2021).

Results displayed in table 6 shows the significance of tax arrears, TMAX, TCOUNT, DMAX, DMED and DCOUNT variables that were based on twelve month ends prior to the month of application time in. Descriptive statistics in table 4 shows that means are all higher in failed groups. It can be said that failed firms are more likely to have bigger tax debts and this could lead to failure. This also follows the findings in Lukason and Andresson (2019) study. Tax arrears are good failure prediction variables. In summary independent non-financial variables shows that means and medians are significantly different through the failed and non-failed firms, only one variable TMED showed no significant difference.

Logistic regression (LR) are used to predict subsidised agricultural firms failure. As the dataset is unbalanced, there are more non-failed firms than failed firms in the sample, a weighted LR is being used additionally to unweighted LR. The procedure of weight for logit analysing method is used previously already back in the 1988 in credit assessment model in farm borrowers (Miller, & LaDue, 1988). To make two groups in the analysis to be equal the weight applied for failed firms is calculated as 0.5/(the share of failed firms in the sample) and for non-failed firms as 0.5/(the share of non-failed firms in the sample). In table 7 domain-based and across-domain logistic regression accuracies results are displayed.

Table 7

Accuracies of all logistic regression models composed (%)

Domain	Logistic regression (LR)					
	Unweighted			Weighted		
	F	NF	All	F	NF	All
Subsides (1)	10.2	98.2	74.7	32.7	83.6	58.1
Financial variables (2)	10.8	98.5	75.9	52.2	63.5	57.8
Reporting delays (3)	0.0	100.0	74.2	52.0	65.4	58.7
Tax arrears (4)	5.8	98.3	73.6	21.3	88.8	55.1
All combined	23.9	96.9	78.2	61.5	70.7	66.1

Notes. F - failed. NF - non-failed. All the domains variables in brackets are listed in Table 2 according to the numbers.

Source: own elaboration

Surprisingly in this dataset subsides variables SUB did make an effect both in ANOVA and in ISMT. LR model correctly predicted 74.7% of subsidies domain cases. In unweighted LR there are subsidies and financial ratios domains that displays good accuracies. Unweighted LR can also identify failed group. Lukason and Andresson (2019) says that models founded on tax arrears are more accurate than models founded on financial ratios in the short term and financial ratio-based models are more valuable in the long duration. The model show a positive association between the tax arrears but was compared to the others domain little bit less accurate. It can be concluded that the bigger and frequent tax debts the higher the risk of a subsidies default.

Failed agricultural firms tend to delay with annual reports and LR model results displays that surprisingly unweighted LR model didn´t hit any failed subsidies, but if data was weighted then it did hit up to 52% failed group in reporting delays domain.

The highest accuracy achieved all combined domain with the accuracy 78.2%. The discriminant model correctly classified 66.1 % of the all combined firms in weighted LR where 61.5 % failed and 70.7 % non-failed subsidies got hit. The overall prediction accuracy through every domain is above 73% in unweighted logistic models and above 55% in weighted regression. Unweighted LR is better for practical use, because accuracy is in the majority class (in non-failed group). Unweighted LR predicted correctly over 73% in each domain. Although weighing can significantly increase the accuracy of failed companies, it would result in a very large drop in the accuracy of non-failed companies.

This study generates prediction models that can relatively precisely predict non-failed subsidies, at the same time it can eliminate about quarter of failed subsides. Since predicting the firms failure that have received government support is accordingly new to the scientific point of view, there aren´t studies to compare models accuracies. If we compare this study with previous overall bankruptcy researches (Altman, Iwanicz-Drozdowska, Laitinen & Suvas, 2020; Lukason, & Andresson, 2019) then this study didn´t achieve so much accuracies which is also logical because the dataset that were used in this study was pre-selected. For an example Lukason and Käsper (2017) bankruptcy prediction models classified correctly 63.8% for t+1 and 67.8% for t+2 and data was also pre-selected. In addition there may be firms that wasn´t failed because of financial reasons and that's why it doesn´t have to display automatically problems.

5. Conclusion

The purpose of this study was to create a business failure prediction model. The companies that received ARIB support were used in the study. Prediction models for analysing data was: subsides; financial variables; reporting delays; tax arrears and all combined model. Financial data was gathered from submitted reports in Estonian Business Register. Tax arrears information was collected from Estonian Tax and Customs Board. To the best knowledge of the author, this is the first study that predicts failure of firms that have received agricultural subsidies. This study used logistic regression to generate prediction models that can relatively precisely predict non-failed subsidised firms, at the same time it can eliminate about quarter of failed subsidised firms. Findings suggest that agricultural firms tax arrears, firm age, subsidy ratio and delays with submitting annual reports offer high predictive performance to subsidised firms´ failures. Firm´ age variable seems to fit within the previous findings of agricultural firms' failure studies and tax arrears and delays with submitting annual reports in default prediction studies.

When a subsidised firm fails, as the worst scenario ARIB could lose the subsidy fully accompanied by potential legal proceeding costs, while opportunity costs also occur because of not funding a non-failing firm. The results can be implemented in practice by agencies providing agricultural support. ARIB could use created prediction model as a tool in subsidies decisions and as a result, the losses caused by incorrectly issued subsidies would be significantly reduced. Even if the prediction model won´t be used as an automatic tool, then in assessing the grants ARIB could look in addition if applicant have: previous tax arrears, previous delays with submitting annual reports, big subsidy ratio, and young age. As these factors are describing more failed enterprises, additional conditions should be applied when deciding on a subsidies designation. This study also revealed that some financial ratios were higher in the failed group, although previous literature indicated that this should have been the opposite. There may be firms that wasn´t failed because of financial reasons and that's why it doesn´t have to display automatically problems. Although analysing the data, the changes in the annual reports of small companies stood out. If accrual-based indicators are not appropriate, companies should submit additional requirements-based reports – detailed financial data (for example bank statement), which would give a more accurate picture of firm's financial performance.

Restrictions in this study could be the pre-selection of companies and the Estonian context that may hinder the generalization of the results in the best possible way. The obtained results could be further improved by using different classifiers settings (for example the behaviour of the manager of the company). Neural network and decision trees could be better for practical use, as some independent variable indicated that different types of failed companies may be involved. Neural network and decision trees are better for considering different types of firms.

List of references

1. Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial Distress Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model. *Journal of International Financial Management & Accounting, 28*(2), 131-171. DOI: https://doi.org/10.1111/jifm.12053
2. Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2020). A Race for Long Horizon Bankruptcy Prediction. *Applied Economics, 52*(37), 4092-4111. DOI: 10.1080/00036846.2020.1730762
3. Altman. E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*(4), 589-609. DOI: https://doi.org/10.2307/2978933
4. Argilés, J. M. (2001). Accounting information and the prediction of farm nonviability. *European Accounting Review, 10*(1), 73-105. DOI: 10.1080/713764592
5. ARIB. (2021). The Agricultural Registers and Information Board website. (2021) Retrieved from https://www.pria.ee/en/about-arib
6. Barakat, N., & Bradley, A.P. (2010). Rule extraction from support vector machines: a review. *Neurocomputing, 74*(1–3), 178–190. DOI: https://doi.org/10.1016/j.neucom.2010.02.016
7. Barnes, A. P., Hansson, H., Manevska-Tasevska, G., Shrestha, S. S., & Thomson, S. G. (2015). The influence of diversification on long-term viability of the agricultural sector. *Land Use Policy, 49*, 404-412. DOI: https://doi.org/10.1016/j.landusepol.2015.08.023
8. Barney, D. K., Graves, O. F., & Johnson, J. D. (1999). The farmers home administration and farm debt failure prediction. *Journal of Accounting and Public Policy, 18*(2), 99-139. DOI: https://doi.org/10.1016/S0278-4254(98)10018-2
9. D´Antoni, J., Mishra, A. K., & Chintawar, S. (2009). Predicting Financial Stress in Young and Beginning Farmers in the United States. Selected paper, Southern Agricultural Economics Association Annual Meeting, Atlanta. Retrieved from https://core.ac.uk/download/pdf/6833223.pdf
10. European Commission. (2016). Farm diversification in the EU. *Members' Research Service*. Retrieved from https://www.europarl.europa.eu/RegData/etudes/BRIE/2016/581978/EPRS_BRI(2016)581978_EN.pdf

11. European Commission. (2018). Ensuring viable farm income. Retrieved from https://ec.europa.eu/info/sites/info/files/food-farming-fisheries/key_policies/documents/cap_specific_objectives_-_brief_1_-_ensuring_viable_farm_income.pdf
12. Featherstone, A.M., Roessler, L.M., & Barry, P.J. (2006). Determining the Probability of Default and Risk-Rating Class for Loans in the Seventh Farm Credit District Portfolio. *Review of Agricultural Economics, 28*(1), 4-23. DOI:10.1111/j.1467-9353.2006.00270.x
13. Foster, B. P., Zurada, J., & Barney, D. K. (2010). Could Decision Trees Help Improve Farm Service Agency Lending Decisions? *Journal of Management Information and Decision Sciences, 13*(1), 69-91. Retrived from https://ir.library.louisville.edu/cgi/viewcontent.cgi?article=1361&context=faculty
14. Franks, J. R. (1998). Predicting financial stress in farm businesses. *European Review of Agricultural Economics, 25*(1), 30-52. DOI: https://doi.org/10.1093/erae/25.1.30
15. Fujikoshi, Y. (1993). Two-way ANOVA models with unbalanced data. *Discrete Mathematics*, *116*(1-3), 315-334. DOI: https://doi.org/10.1016/0012-365X(93)90410-U
16. Gestel, T. V., Baesens, B., & Martens, D. (2010). From linear to non-linear kernel based classifiers for bankruptcy prediction. *Neurocomputing 73*(16-18), 2955–2970. DOI: https://doi.org/10.1016/j.neucom.2010.07.002
17. Hansson, H., Ferguson, R., & Olofsson, C. (2010). Understanding the diversification and specialization of farm businesses. *Agricultural and Food Science. 19*(4), 269-283. DOI: 10.2137/145960610794197605
18. Jolly, R.W., Paulsen, A., Johnson, J. D., Baum, K. H., & Prescott, R. (1985). Incidence, Intensity, and Duration of Financial Stress among Farm Firms. *American Agricultural Economics Association, 67*(5), 1108-1115. DOI: https://doi.org/10.2307/1241382
19. Jouault, A., & Featherstone, A. M. (2006). Determining the Probability of Default of Agricultural Loans in a French Bank. Conference Paper, American Agricultural Economics Association Annual Meeting. DOI: 10.22004/ag.econ.21376
20. Katchova, A. L. (2010). Structural changes in U.S. agriculture: financial performance of farms in transition. Conference Paper, European Association of Agricultural Economists (EAAE). DOI: 10.22004/ag.econ.60965

21. Klepac, V., & Hampel, D. (2017). Predicting financial distress of agriculture companies in EU. *Agricultural Economics – Czech, 63*(8), 347–355. DOI: 10.17221/374/2015-AGRICECON

22. Kohv, K., & Lukason, O. (2021). What Best Predicts Corporate Bank Loan Defaults? An Analysis of Three Different Variable Domains. *Risks*, *9*(2), 29. DOI: https://doi.org/10.3390/risks9020029

23. Limsombunchai, V., Gan, C., & Lee, M. (2005). An Analysis of Credit Scoring for Agricultural Loans in Thailand. *American Journal of Applied Sciences, 2*(8): 1198-1205. Retrieved from https://thescipub.com/abstract/ajassp.2005.1198.1205

24. Lukason, O. & Laitinen, E. K. (2019). Firm failure processes and components of failure risk: An analysis of European bankrupt firms. *Journal of Business Research, 98*, 380-390. DOI: https://doi.org/10.1016/j.jbusres.2018.06.025

25. Lukason, O., & Andresson, A. (2019). Tax Arrears Versus Financial Ratios in Bankruptcy Prediction. *Journal of Risk and Financial Management*, *12*(4), 1-13. DOI: https://doi.org/10.3390/jrfm12040187

26. Lukason, O., & Camacho-Miñano, M.-M. (2021). What Best Explains Reporting Delays? A SME Population Level Study of Different Factors. *Sustainability*, *13*(9), 4663. DOI: https://doi.org/10.3390/su13094663

27.Lukason, O., & Hoffman, R., C. (2015). Firm failure causes: a population level study. *Problems and Perspectives in Management*, *13*(1), 45-55. Retrieved from https://www.businessperspectives.org/images/pdf/applications/publishing/templates/article/assets/6325/PPM_2015_01_Lukason.pdf

28. Lukason, O., & Käsper, K. (2017). Failure prediction of government funded start-up firms. *Investment Management and Financial Innovations, 14*(2), 296-306. DOI: http://dx.doi.org/10.21511/imfi.14(2-2).2017.01

29. Miller, L.H., & Ladue, E.L. (1988). Credit assessment models for farm borrowers: A logit analysis. A working papers in agricultural economics, 88-12. Retrieved from https://ageconsearch.umn.edu/record/178692/files/Cornell-Dyson-wp8812.pdf

30. Pannell, D. J., Malcolm, B., & Kingwell, R. S. (2000). Are we risking too much? Perspectives on risk in farm modeling. *Agricultural Economics, 23*, 69-78. DOI: https://doi.org/10.1111/j.1574-0862.2000.tb00084.x

31. Savitha, B., & Kumar K., N. (2016). Non-performance of financial contracts in agricultural lending: A case study from Karnataka, India. *Agricultural Finance Review*, *76*(3), 362-377. DOI: https://doi.org/10.1108/AFR-01-2016-0001

32. Sheskin, D. J. (2011). *Handbook of Parametric and Nonparametric Statistical Procedures* (5th ed.). Chapman and Hall/CRC.

33. Stulpinienė, V. (2011). The Concept, Causes, and Measurement of Farm Financial Distress. *Proceedings of the International Scientific Conference: Rural Development, 5*(1), 249-253. Retrieved from http://dspace.lzuu.lt/jspui/bitstream/1/2902/1/rural_developmen_2011_book1.pdf#page=250

34. Veganzones, D., & Severin, E. (2020). Corporate failure prediction models in the twenty-first century: a review. *European Business Review, 33*(2), 204-226. DOI: https://doi.org/10.1108/EBR-12-2018-0209

35. Zhou, Z.-H., Jiang, Y., & Chen, S.-F. (2003). Extracting symbolic rules from trained neural network ensembles. *AI Communications*, *16*(1), 3-15. Retrieved from http://129.211.169.156/publication/aicom03.pdf

Kokkuvõte

Selle magistritöö eesmärk oli ennustada põllumajandustoetusi saanud ettevõtete ebaõnnestumist. Analüüsimiseks kasutati logistilise regressiooni mudeleid, mis sisaldasid järgnevaid domeene: toetuste suhtarvud; finantsmuutujad; majandusaasta aruannetega viivitamine; maksuvõlgade muutujad ja eelneva nelja domeeni põhjal kombineeritud mudel. Kasutatud andmed pärinevad Põllumajanduse Registrite ja Informatsiooni Ametilt (PRIA). Finantsandmed koguti esitatud majandusaasta aruannetest Eesti Äriregistrist. Maksuvõlgade teave pärines Eesti Maksu- ja Tolliametist. Autorile teadaolvalt on see esimene uuring, mis ennustab põllumajandustoetusi saanud ettevõtete ebaõnnestumist.

Varasem kirjandus annab ülevaate sellest, et finantsandmed on kasulikud põllumajandusettevõtete ebaõnnestumise ennustamiseks (Barney, Graves, & Johnson, 1999) ning Altman ütleb (1968), et ebaõnnestumist saab ennustada kaks aastat enne sündmust. Ka see töö kasutab ennustusmudeli jaoks toetuse taotlemisele eelnenud kahe aasta andmeid ning maksuvõlgade jaoks eelnevate 12 kuulõpu andmeid. Autorile teadaolevalt ei ole põllumajandusettevõtete varasemates töödes kasutatud maksuvõlgade näitajaid ning majandusaasta aruannete hilinemisi. Küll aga on neid faktoreid uurinud Kohv ja Lukason (2021) ettevõtete maksehäire ennustamise mudelites.

Töös leiab kinnitust, et põllumajandusettevõtete varasemad maksuvõlad, ettevõtte vanus, subsiidiumimäär ja aastaaruannete hilisem esitamine on subsideeritud ettevõtte ebaõnnestumise prognoosimisel kõige täpsemad. Loodud ebaõnnestumise mudel suudab suhteliselt täpselt ennustada mitte-ebaõnnestunud subsideeritud ettevõtteid, samal ajal kõrvaldades umbes veerandi ebaõnnestunud subsideeritud ettevõtetest. Loodud prognoosimudelit saaks PRIA kasutada toetuste otsuste tegemisel ning selle tulemusel väheneksid valesti väljastatud toetuste poolt põhjustatud kahjud oluliselt. Isegi kui ennustusmudelit ei kasutata automaatse tööriistana, võiks PRIA toetuste hindamisel lisaks uurida, kas taotlejal on: varasemad maksuvõlgnevused, varasemad viivitused aastaaruannete esitamisel, suur toetuste määr ja ettevõte on vähe tegutsenud. Kuna need tegurid kirjeldavad rohkem ebaõnnestunud ettevõtteid, tuleks subsiidiumide määramise otsustamisel rakendada täiendavaid tingimusi. Kui tekkepõhised näitajad ei ole asjakohased, peaksid ettevõtted esitama täiendavad nõudepõhised aruanded - üksikasjalikemaid finantsandmeid (näiteks konto väljavõtte), mis annaksid täpsema pildi ettevõtte majandustulemustest. Töö piiranguks võivad olla ettevõtete eelvalik ja Eesti kontekst, mis võib tulemuste üldistamist parimal võimalikul viisil takistada.

Non-exclusive licence to reproduce thesis and make thesis public

I, Gerda Hermann (date of birth: 30.04.1988),

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:
1.1. reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and
1.2. make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

Prediction of subsidised firm failure with financial ratios, tax arrears and annual report delays,

supervised by Oliver Lukason and Tõnis Tänav,

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, **25.05.2021**