DE GRUYTER
OPEN

Dealing with Interpretability Issues in Predicting Firm Growth: Factor Analysis Approach

Ana Bilandžić, Jeger Marina, Nataša Šarlija
Faculty of Economics in Osijek, J.J. Strossmayer University of Osijek, Croatia

Abstract

Background: Alongside the theoretical progress made in understanding the factors that influence firm growth, many methodological challenges are yet to be overcome. Authors point to the notion of interpretability of growth prediction models as an important prerequisite for further advancement of the field as well as enhancement of models' practical values. **Objectives:** The objective of this study is to demonstrate the application of factor analysis for the purpose of increasing overall interpretability of the logistic regression model. The comprehensive nature of the growth phenomenon implies propensity of input data to be mutually correlated. In such situations, growth prediction models can demonstrate adequate predictability and accuracy, but still lack the clarity and theoretical soundness in their structure. **Methods/Approach:** The paper juxtaposes two prediction models: the first one is built using solely the logistic regression procedure, while the second one includes factor analysis prior to development of a logistic regression model. **Results:** Factor analysis enables researchers to mitigate inconsistencies and misalignments with a theoretical background in growth prediction models. **Conclusions:** Incorporating factor analysis as a step preceding the building of a regression model allows researchers to lessen model interpretability issues and create a model that is easier to understand, explain and apply in real-life business situations.

Keywords: factor analysis, logistic regression, interpretability, growth prediction model
JEL classification: C53, C55
Paper type: Research article

Introduction

Measuring and predicting firm growth is a vital topic of entrepreneurship research as it interconnects with numerous economic and management theories. Most of the studies in the field are oriented toward making theoretical progress in terms of setting the appropriate context and theoretical framework for studying firm growth. More recently, several studies have been devoted to methodological considerations related to the measure of firm growth (Shepherd et al., 2009; Janssen, 2009; Weinzimmer, 1998) and a selection of predictor variables (Sampagnaro, 2013; Kiviluoto et al., 2011; de Wit & Zhou, 2009). The current state of the field can be described as very fragmented in terms of definition of the growth variable. Inconsistency in dependent variable operationalization coupled with several other methodological issues (i.e. differences in time span over which the growth is modelled and frequent alternations between relative and absolute measures of growth) has led to inconsistencies in findings and implications for both scholars and policy makers. The authors of this paper argue that, in addition to these methodological considerations, interpretability of the model needs to be taken in account when building growth prediction models. Such models can serve as a tool that facilitates business decision making and therefore should be conceivable and easy to use by business people and entrepreneurs.

Regardless of the decision to measure growth as an increase in revenues, assets, market share or number of employees, the growth measure is usually defined as a binary variable. Logistic regression modelling is widely used for the analysis of multivariate data involving binary responses. It provides a powerful technique analogous to multiple regression and ANOVA for continuous responses. However, when working with highly correlated variables, logistic regression may provide results that are, from the theoretical perspective, very hard to interpret. The aim of this paper is to demonstrate the application of factor analysis as a preceding step to building logistic regression model in order to boost interpretability without compensating on accuracy and predictive power of the model. The goal of the study is to develop a model for estimating growth potential of small and medium-sized enterprises (SMEs) in Croatia based on predominantly financial data and some nonfinancial data noted in their financial statements. This study sheds a light to methodological difficulties in modelling enterprise growth and suggest a way of tackling those concerns. Additionally, further contribution of the paper can be found in the extraction of features that are relevant for predicting enterprise growth in the context of small and medium-sized companies. Those features are presented in a form of financial ratios that pinpoint main business aspects relevant for enterprise growth.

Previous studies used various approaches to tackle the issue of interpretability and accuracy of the prediction model. Schielzeth (2010) suggests some simple methods, such as centering and standardization of input variables or thoughtful removal of intercepts or main effects, to improve interpretability of regression coefficient in linear regression models. Furthermore, Li (2014) used a combination of principal component and logistic regression to distinguish accounting information distortion and achieve higher model accuracy. Similar approach was used in two other studies (Shengyuan, 2009; Kehong et al., 2006) that combined principal component and logistic regression but in a context of corporate financial distress prediction. Zhu et al. (2010) used principal component as a pre-processing method before applying logistic regression and discriminant analysis for credit risk estimation. Suleiman et al. (2014) used principal component as input for predicting applicants' creditworthiness in order to improve the predictive power of linear discriminant and logistic regression

models. Results showed that the use of principal component as input improved models prediction by reducing their complexity and eliminating data co-linearity.

Most of the research in high-growth enterprises is oriented on growth potential of large companies (Davidsson et al. 2006). Only recently studies have made a progress in identifying predictors of growth, but researchers are still far from an extensive comprehension of this topic. Factors influencing growth in small firms have usually been understood in terms of three main categories: characteristics of an entrepreneur, business and management practices, and institutional factors. Willingness to participate in situations with uncertain outcomes, mid-management experience (Cassia et al. 2009), education and entrepreneur's aspiration to grow (Kolvereid and Bullvag, 1996; Barringer et al., 2005) are selected as relevant growth predictors on an entrepreneur level. From the company's perspective age and size, strategic orientation (Barringer et al. 2005; Morone and Testa, 2008), innovation (Fischer et al., 1997), financial structure and productivity (Mateev and Anastasov, 2010) have strong positive relationship with growth potential. Finally, in the context of institutional factors, company's growth is under influence of tax system, regulated credit market conditions, employment security laws, low wage dispersion due to wage setting institutions, and public sector monopolization of the production of key services (Henrekson and Johansson, 1999).

Financial determinants of growth are present in models developed in the previous studies. Moreno & Casillas (2007) showed that rapid-growth firms are characterized by a lower availability of financial resources in the years immediately preceding their growth. This is consistent with Stevenson & Jarillo (1990) and Baum et al. (2001) who concluded that searching for and exploiting opportunities contributes to accelerated growth more than efficiently managing acquired financial resources. On the other hand, Becchetti & Trovato (2002) showed that availability of external finance and internationalization are positively related to firm's growth. In the context of transition countries, Mateev & Anastasov (2010) have suggested that firm growth is determined not only by the traditional characteristics of size and age but also by other firm-specific factors such as indebtedness, internal financing, future growth opportunities, process and product innovation, and organizational changes. Sampagnaro (2013) has identified the balance sheet ratios that enable managers to predict which enterprises are better candidates for a high-growth path. The study pointed out that firm size, firm age and, primarily, internal cash flows (despite bank loans), are of most relevance to the growth and success of a firm.

Methodological steps undertaken in this paper are the following: (i) development of the logistic regression model where predictors are financial ratios defined as observed variables; (ii) application of factor analysis on predictors in order to create factors; (iii) development of the logistic regression model where predictors are factors; (iv) comparison of both logistic regression models in light of their interpretability and predictability.

The structure of the paper follows the methodological steps and ends with the discussion on advantages and limitations of this approach, possibilities of application and suggestions for further research.

Data and Variables

The sample used in this research consists of 1492 privately-owned small and medium-sized enterprises (SMEs) in Croatia. They were chosen from the Financial Agency (FINA) data set that includes 53 434 SMEs which operated over the period from 2008 to 2013. For the purpose of this study, an enterprise is defined as high growth if it has average annualized growth in assets greater than 20% a year, over a three-year

period, from 2010 to 2013 (OECD, 2010). Out of total number of SMEs, 746 enterprises met this criterion. Development sample included 650 high growth SMEs, while validation sample consists of 96 high growth SMEs. The other 746 SMEs, which are not high growth, were selected randomly from the whole data set. They were divided in the same way as high growth enterprises.

Table 1
Descriptive Analysis of Predictors and the Statistical Significance of the Difference between High Growth and Non-high Growth Enterprises

Predictor	Non-High Growth		High Growth		
	Median	IQR	Median	IQR	p*
Liquidity ratios:					
current assets/total assets	0.745	0.595	0.839	0.418	<0.001
Leverage ratios:					
total debt/total assets	0.706	0.593	0.793	0.870	<0.001
current liabilities/ equity	0.473	1.910	0.348	3.062	0.047
Turnover ratios:					
total revenue/ total assets	3.781	14.166	9.613	26.26	<0.001
sales/total assets	0.855	1.57	1.657	2.516	<0.001
sales/working capital	0.438	4.323	0	6.072	0.907
(current assets – inventory)/sales	0.396	0.596	0.310	0.419	<0.001
365/receivables turnover	54.16	119.39	36.55	86.849	<0.001
Profitability ratios:					
net income/equity	7.61	33.25	23.64	56.172	<0.001
net income/sales	0.016	0.064	0.022	0.084	0.212
retained earnings/ total assets	0.045	0.345	0	0.778	<0.001
Other variables:					
non-tangible assets/total assets	0	0	0	0	0.349
% sales change	-0.094	0.437	-0.087	0.429	0.686
% total revenue change	-0.076	0.476	-0.077	0.487	0.629
% employees change	0	0.157	0	0.166	0.749
% assets change	-0.048	0.283	-0.052	0.272	0.436

* Statistical significance was measured according to Mann-Whitney test

Predictors for growth model are created for every enterprise in the data set for the period from 2008 to 2010. Total of 101 variables were created. They are grouped in 5 groups: liquidity (15 ratios), turnover (30 ratios), leverage (15 ratios) and profitability

(15 ratios). The fifth group consists of 26 variables which include industry sector, non-tangible assets and percentage change in number of employees, assets, profitability, sales and some other performance indicators. The following table displays the results of descriptive analysis of predictors included in the model.

In terms of industry affiliation, high growth firms included in the sample predominantly operate in ICT sector, finance and real estate, education, and agriculture. On the other hand, non-high growth firms are generally associated with trade, transport and storage, manufacturing industry, constructions, and hotels and restaurants.

Methods

In the process of developing a model with the binary dependent variable Y (probability for a firm to reach high growth or not) logistic regression was used. In simple terms, for one predictor x the logistic function would be:

$$y = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \tag{1}$$

The goal is to obtain β_0 and β_1. Because the above formula is not linear, through logistic transformation it becomes:

$$logit(y) = \beta_0 + \beta_1 x \tag{2}$$

where $logit(y) = \frac{y}{1-y}$, often called 'log odds'. Intuitively for more variables $logit(y)$ becomes

$$logit(y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n \tag{3}$$

Selection procedures forward and backward were used, and the selected variables were used with R built in function `glm()` to obtain our first model (Agresti, 2002).

To address the difficulties with interpretation of regression coefficients, factor analysis was conducted. Factor analysis is a procedure used to obtain a model with the following structure:

$$X_1 - \mu_1 = a_{11}F_1 + a_{12}F_2 + \cdots + a_{1r}F_r + U_1$$
$$\vdots$$
$$X_p - \mu_p = a_{p1}F_1 + a_{p2}F_2 + \cdots + a_{pr}F_r + U_p \tag{4}$$

where X_i, $i = 1,2, \dots, p$ are observed variables, and F_j and U_i, $i = 1,2, \dots, p; j = 1,2, \dots, r$ are unobserved. Equivalently, the set of equations can be written as

$$(\boldsymbol{x} - \boldsymbol{\mu}) = \boldsymbol{A}\boldsymbol{f} + \boldsymbol{u} \tag{5}$$

where A is the factor pattern matrix consisting of its elements a_{ij} which are called factor loadings, $\boldsymbol{x}$ is the $p \times 1$ vector of elements X_i, $i = 1,2, \dots, p$ and $\boldsymbol{\mu}$ is vector of their means. While $\boldsymbol{f}$ is the $r \times 1$ vector of elements F_j, $j = 1,2, \dots, r$, they are called common factors and are assumed to have mean 0 and variance 1. U_i, $i = 1,2, \dots, p$ are unique factors and are assumed to have mean 0, but variance σ_i^2, $i = 1,2, \dots, p$, they form the $p \times 1$ vector $\boldsymbol{u}$. Addionally, it is assumed that the unique and common factors are uncorrelated. So, by marking the covariance matrix of $\boldsymbol{x}$ with Σ, the previous equation turns into:

$$\Sigma = \mathrm{E}[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})'] = \mathbf{A}\mathbf{A}' + \boldsymbol{\Psi} \tag{6}$$

where $\boldsymbol{\Psi}$ is the vector of variances of U_i. Whereas the right side of the equation consists only of unobserved data, this process is not unique, and different factors can be obtained (Jobson, 2012).

To get the factor scores, F_j, $j = 1,2,...,r$ R was used and the function `fa()`, from the package `psych`. Not all variables showed correlation with at least one factor, so one by one was excluded from the factor analysis, until the desired result was reached. With the factor loadings from function `fa()` and some additional variables that covered profitability once again `glm()` was used to develop new logistic regression model.

Results

For the purpose of this study, data analysis and model development procedures can be divided into four steps. First, the standard logistic regression model was built based on predefined set of financial ratios. The model was juxtaposed to the underlying theoretical background, and inconsistencies in model results were identified. Second, factor analysis was applied on predictors and three factors were singled out as a result of that analysis. Third, a new prediction model was developed by using factors as predictors. And finally, a comparison of a prediction model without factor analysis and a prediction model with factor analysis was given and the results were evaluated in the context of both theoretical framework and methodological approach.

The first step covered development of logistic regression prediction model in a standard way with financial ratios set as observed variables. To reach an adequate level of prediction, the authors developed several prediction models that showed average performance. The final model consists of 15 variables and has satisfactory performance measures (total hit rate 64.65%, hit rate for growth firms 64.1%, hit rate for non-growth firms 65%; AUC=0.7, KS=40.26%; AIC=828.32). The structure of the model is presented in the following table.

The variance inflation factor (VIF) indicates moderate or high correlation among predictors (VIF values above 5 point to the presence of multicollinearity). In situations like this one, the direction of regression coefficient may not be in line with theoretical background or even sound business logic. In that context, several inconsistencies should be noted. First, some predictor variables have negative coefficients while the theory and sound reasoning suggest the opposite. For example, the Fixed-asset turnover has a negative regression coefficient suggesting that the increase of the turnover leads to decrease of potential to growth. Second, while the descriptive analysis shows that higher values of a particular indicator are characteristic for growing companies, the final model can show the negative value of respective coefficients, such as the Ratio of short-term liabilities to equity. Third, sometimes the sign of the coefficients can change in the process of model development which happened with the Return on equity and the Ratio of retained earnings and total assets.

To account for these inconsistencies, factor analysis on predictors was applied in the next step. Analysis of variance of eigenvalues showed that three factors could be generated. This represents 99% of total variance. The results of this procedure are shown in table 3.

Table 2
Results of Logistic Regression Model

Predictor variables		Regression Coefficients	VIF
Liquidity ratios:			
current assets/total assets		0.788**	1.114
Leverage ratios:			
total debt/total assets		0.032	257.99
current liabilities/ equity		-0.009*	1.033
Turnover ratios:			
total revenue/ total assets		$-1.02*10^{-5}$	1.059
sales/total assets		0.322***	1.716
sales/working capital		$2.79*10^{-7}$	1.018
(current assets – inventory)/sales		-0.002*	1.002
365/receivables turnover		-0.003*	1.024
Profitability ratios:			
net income/equity		0.001*	1.438
net income/sales		-0.015	1.021
retained earnings/ total assets		-0.007	254.53
Other variables:			
non-tangible assets/total assets		3.474**	1.042
% sales change		-0.028	1.003
% employees change		-0.347*	1.034
% assets change		0.194*	1.032
Industry affiliation*	Industry	-1.239	1.024
	Construction	-0.715	
	Trade	-1.509	
	Transport and storage	-1.883	
	Hotels and restaurant	-1.218	
	Information and communication	-0.896	
	Finance and real estate	-1.499	
	Other business activities	-0.916	
	Education, other services	-1.039	

Accuracy of the model: total hit rate 64.65%, hit rate for growth firms 64.1%, hit rate for non-growth firms 65%; AUC=0.7, KS=40.26%; AIC = 828.32

Table 3
Factor loadings of three extracted factors

Ratios	Abbreviation	Factor 1 (Turnover Factor)	Factor 2 (Liquidity Factor)	Factor 3 (Leverage Factor)
Current Ratio	**CuR**	-0.00012	**0.99866**	-0.0001
Leverage Ratio	**LR**	$3.04*10^{-6}$	$-1.7*10^{-7}$	**0.99874**
Total Equity to Total Asset Ratio	**TETAR**	$-3*10^{-6}$	$2.6*10^{-7}$	**-0.9987**
Quick Ratio	**QR**	$-6.3*10^{-5}$	**0.99866**	-0.00005
Total Asset Turnover	**TAT**	**0.99873**	0.00019	0.00021
Current Asset Turnover	**CAT**	**0.99811**	-0.00016	-0.00016
Sales to Total Asset Ratio	**STAR**	**0.99851**	-0.00004	-0.00005
Cash Ratio	**CaR**	0.0003	**0.84705**	0.00025

Looking at the factor loadings it could be noticed that variables are grouped in theoretically sound way assessing three groups of business performance indicators: business activity (turnover ratios), liquidity and leverage. The factors are defined by following equations:

$$F_1 = 0,333378\text{TAT} + 0,333174\text{CAT} + 0,333307\text{STAR}$$

$$F_2 = 0,496263\text{CuR} + 0,496264\text{QR} + 0,00745\text{CaR}$$

$$F_3 = 0,499363\text{LR} - 0,49936\text{TETAR}$$

Furthermore, in the third step, the factors were treated as predictors and the new logistic regression model was developed. Additionally, since profitability ratios were removed in the process of factor analysis, they were put back during the logistic regression model development together with some variable from the fifth group of variables. The logistic regression results are given in the table 3.

Table 4
Logistic regression model with factors as independent variables

Factor/Variable	Regression coefficient
Factor 1 (Turnover Factor)	160.3426
Factor 2 (Liquidity Factor)	8.2408
Factor 3 (Leverage Factor)	147.246
Intangible Assets/Total Assets	1.6228
Net Income/Sales	-0.0068
Net Income/Equity	0.0002
Change (%) in number of employees	-0.1437
Accuracy of the model: total hit rate 60.67%, hit rate for growth firms 56.1%, hit rate for non-growth firms 64.58%; AUC=0.64, KS=29.42%; AIC = 748.77	

The final step in the analysis was to compare logistic regression models with and without factor analysis applied prior to model building. Comparison is done according to model quality, interpretability and predictability. Based on hit rates, area under the curve and Kolmogorov-Smirnov indicators, the first model exhibits a slightly better performance in terms of model predictability. Nevertheless, according to Akaike information criterion (AIC), the second model shows a higher relative quality and therefore represents a better option when choosing a model for predicting growth. As regard to interpretability, all of the three issues introduced in the paper have been addressed. First, results obtained by combining factor analysis and logistic regression gave results that are more logical and in line with the theoretical framework (Sampagnaro, 2013; Segarra and Teruel, 2009). As it can be seen from the table 2, the potential for firm growth rises with an increase of liquidity, turnover, leverage and profitability, with a drop of return on sales and with a decrease in change in number or employees. Second, regression results coincide with the input data. Leverage ratios are higher in high growth SMEs and the same can be seen in the final model. Third, the results are consistent. Since one factor represents liquidity it implies that higher liquidity means higher potential to growth and it cannot happen that one liquidity coefficient is positively associated with the growth potential and the other negatively (which was the case in a regression model developed in a standard way).

Discussion and Conclusion

Logistic regression is a method of choice when analysing multivariate data involving binary responses. Still, sometimes logistic regression may result in models that are very difficult to explain, especially if derived from a large set of highly correlated variables. To overcome these issues without any great loss in information contained in data, the authors conducted factor analysis prior to regression model building. The goal was to improve interpretability of the model by identifying most important factors and reducing large number of variables, as well as lowering multicollinearity levels. Finally, the model created by combining factor analysis with logistic regression exhibits variables that are simpler and easier to interpret. Those variables, when observed separately, depict the main area of business performance – business activity, liquidity, leverage and profitability. Moreover, the influence of each variable on predicting growth is unambiguous and aligned with the underlying theory.

The first factor relates to the business activity, measure that put emphasis on efficiency of business operations by evaluating whether the assets structure is in line with the overall strategy. More precisely, the model confirms that the higher the speed at which a company converts its non-cash assets to cash assets, the better are prospects to grow (Mateev and Anastasov, 2010). Small and medium-sized enterprises have to match their strategy with the limited resources they have available, and conducting business activities faster, cheaper and at full capacity is a way of building competitive advantage over large competitors. The second factor depicts the value of various short-term assets in relation to the short-term liabilities. However, setting target values of liquidity ratios is not straightforward. On the one hand, enterprises should aim at higher levels of liquidity since keeping sufficient levels of liquid assets is prerequisite for financial success (which is confirmed by the model). On the one hand, excess amounts of current assets can be a sign of issues in the receivables collection or inventory obsolescence. The last factor, leverage factor, suggests that companies with a prospects for growth tend to use external financing (primarily bank loans) to fuel their growth. This finding confirms the importance of availability of external financing for companies aiming to grow (Becchetti & Trovato, 2002). However, scholars do not agree on this notion as some of the studies have singled out internal financing as relevant growth factor (Sampagnaro, 2013; Mateev and Anastasov, 2010). In the context of the data set we used, our results may be a reflection of the fact that SMEs in Croatia have very little options, aside from bank financing, for a long-term funding. Market of informal and formal investors such as angle investors and venture capital funds is underdeveloped in Croatia.

This study has certain limitations and they are predominantly related to the data itself. The models are developed on indicators that are measured on a firm level and are predominantly financial in their nature. However, firm growth can be influenced by many other factors that are related to the person of an entrepreneur as well as to the idiosyncrasies of the macroeconomic environment. Therefore, applying factor analysis on an expended data set, that includes different types of variables related to various units of analysis, would be a worthwhile activity.

Nevertheless, model for predicting company's growth can be a powerful strategic tool for managers and entrepreneurs, and is widely applicable in many areas of business decision making such as finance, management, marketing and sales. This only emphasizes the need for developing growth prediction models that are easy to understand and apply in business decision making, not only for researchers and academics, but for business people as well.

Growth prediction models can further be improved by focusing on two aspects: improving the existing data (strengthening the supervision and increasing the

reliability of financial statements), and expanding the data set (by including non-financial, entrepreneur-level and macroeconomic variables). All of that could lead to an increase of predictability of the model.

References

1. Agresti, A. (2002), "Categorical Data Analysis", Second Edition, John Wiley & Sons, Inc., Hoboken, New Jersey
2. Barringer, B. R., Jones, F. F., Neubaum, D. O. (2005), "A quantitative content analysis of the characteristics of rapid-growth firms and their founders", Journal of business venturing, Vol. 20 No.5, pp. 663-687.
3. Baum, J. R., Locke, E. A., Smith, K. G. (2001), "A multidimensional model of venture growth", Academy of management journal, Vol. 44 No.2, pp. 292-303.
4. Cassia, L., Cogliati, G. M., Paleari, S. (2009), "Hyper-Growth Among European SMEs: An Explorative Study", Available at SSRN 1389521.
5. Davidsson, P., Delmar, F., Wiklund, J. (2006), "Entrepreneurship as growth; growth as entrepreneurship", Entrepreneurship and the Growth of Firms, pp. 21-38.
6. de Wit, G., Zhou, H. (2009), "Determinants and dimensions of firm growth", EIM Business and Policy Research (No. H200903)
7. Fischer, E., Reuber, A. R., Hababou, M., Johnson, W., & Lee, S. (1997), "The role of socially constructed temporal perspectives in the emergence of rapid growth firms", Entrepreneurship theory and practice, Vol. 22, pp. 13-30.
8. Henrekson, M., Johansson, D. (1999), "Institutional effects on the evolution of the size distribution of firms", Small Business Economics, Vol. 12 No.1, pp. 11-23.
9. Janssen, F. (2009), "The conceptualisation of growth are employment and turnover interchangeable criteria?", Journal of entrepreneurship, Vol.18 No.1, pp. 21-45.
10. Jobson, J. (2012), "Applied multivariate data analysis: volume II: Categorical and Multivariate Methods", Springer Science & Business Media.
11. Kehong, G., Zhansheng, L. (2006), "Empirical Comparison among Different Predicting Models of Financial Failure for China's Listed Corporations", Journal of Shanxi Finance and Economics University, Vol. 5, pp. 129-133.
12. Kolvereid, L., Bullvag, E. (1996), "Growth intentions and actual growth: The impact of entrepreneurial choice", Journal of enterprising Culture, Vol. 4, No. 1, pp. 1-17.
13. Li, Z. (2014), "Research on Distinguish the Accounting Information Distortion Based on the Principle Components-logistic Regression Model", International Journal of Security and Its Applications, Vol. 8 No.4, pp. 37-50.
14. Mateev, M., Anastasov, Y. (2010), "Determinants of small and medium sized fast growing enterprises in central and eastern Europe: a panel data analysis", Financial Theory and Practice, Vol. 34 No.3, pp. 269-295.
15. Morone, P., Testa, G. (2008), "Firms growth, size and innovation: An investigation into the Italian manufacturing sector", Economics of Innovation and New Technology, Vol. 17 No.4, pp. 311-329.
16. Organization for Economic Co-operation and Development (2010), "High-growth enterprises: what governments can do to make a difference", OECD Publishing
17. Sampagnaro, G. (2013), "Predicting rapid-growth SMEs through a reversal of credit-scoring principles", International Journal of Entrepreneurship and Small Business, Vol.18 No.3, pp. 313-331.

18. Schielzeth, H. (2010), "Simple means to improve the interpretability of regression coefficients", Methods in Ecology and Evolution, Vol.1 No. 2, pp. 103-113.
19. Segarra, A., Teruel, M. (2009), "Small firms, growth and financial constraints", Xarxa de Referència en Economia Aplicada (XREAP)
20. Sheng-yuan, Y. A. N. G. (2009), "Study on financial predicament forecast of listed company in China", Journal of Kunming University, Vol.5, pp. 94-99.
21. Shepherd, D., Wiklund, J. (2009), "Are we comparing apples with apples or apples with oranges? Appropriateness of knowledge accumulation across growth studies", Entrepreneurship Theory and Practice, Vol.33 No.1, pp. 105-123.
22. Stevenson, H. H., Jarillo, J. C. (1990), "A paradigm of entrepreneurship: Entrepreneurial management", Strategic management journal, Vol. 11 No.5, pp. 17-27.
23. Suleiman, S., Issa, S., Usman, U., Salami, Y. O. (2014), "Predicting an applicant status using principal component, discriminant and logistic regression analysis", International Journal Of Mathematics And Statistics Invention (IJMSI), Vol. 2 No.10, pp. 5-15.
24. Weinzimmer, L. G., Nystrom, P. C., Freeman, S. J. (1998), "Measuring organizational growth: Issues, consequences and guidelines", Journal of management, Vol.24 No.2, pp. 235-262.
25. Zhu, K., Li, J.(2010), "Studies of discriminant analysis and logistic regression model application in credit risk for China's listed companies", Management Science and Engineering, Vol.4 No.4, pp. 24-32.

About the authors

Ana Bilandžić is a research assistant at the University of J.J. Strossmayer in Osijek, Faculty of Economics in Osijek, Croatia. She got her master's degree in mathematics at University of J.J. Strossmayer in Osijek, Department of mathematics and is currently a PhD student at University of J.J. Strossmayer in Osijek, Faculty of Economics, Croatia on the doctoral program Entrepreneurship and Innovativeness. Her research interests are focused on high growth entrepreneurship and mathematical modelling. Author can be contacted at anag@efos.hr

Marina Jeger works as an assistant professor at Faculty of Economics in Osijek. She got her MBA degree in corporate finance and entrepreneurship at Kelley School of Business in Bloomington, Indiana, USA, and her PhD degree in field of entrepreneurship at interdisciplinary doctoral program at J. J. Strossmayer University of Osijek. Her research covers several areas of entrepreneurship with the emphasis on entrepreneurial intentions, financial distress and firm growth modeling. Author can be contacted at marina@efos.hr

Nataša Šarlija is a full professor at the University of J.J. Strossmayer in Osijek, Faculty of Economics in Osijek, Croatia. She got her PhD from University of Zagreb, Faculty of Economics in Croatia on the modeling of credit risk for small and medium enterprises. Her research interests are: credit risk modeling, high growth entrepreneurship, credit analysis, financial modelling. Author can be contacted at natasa@efos.hr