DOI: https://doi.org/10.3126/pycnjm.v2i`1.81969

Characteristics of Fast Growth and Highly Successful Business – Problems in Approaches and Methods of Analysing the Data

Markku Virtanen

Abstract : *This paper focuses the problems in analyzing simultaneous growth and success of a firm. The refinement process of data will be explained in the course of several studies. The refined data consists of 348 growing firms where from 75 were fast growing and highly successful. We used regression, logistic regression and discrimination analysis in analyzing the data. Regression analysis reveals that growth and success are not surrogates. Cross-validated logistic regression and discriminant analysis classify well the cases in different categories. The robust analysis presupposes careful refinement of data, accurate model specification, and competent interpretation of the results.*

I. Introduction

Fostering of fast growth and highly successful businesses is seen as an important element of regional development. Different approaches have been used to identify the so called gazelles since they are thought to generate the majority of new jobs in the society. Many expectations will be loaded to the entrepreneurs and their businesses when society is changing. But these expectations may be overestimated because of the wide diversity of entrepreneurs and businesses. Thus, the identification of fast growth and highly successful businesses could decrease the gap between expectations and actual outcome.

From the policy perspective, it will be essential to be able to differentiate fast growth businesses since these businesses will be the most prominent portfolio companies of venture capitalists. On the other hand, improving access to venture capital is one very

Dr. Virtanen is Professor at Helsinki School of Economics. Hee can be reached at markku.virtanen@pyk.hkkk.fi

important policy target. Several studies have been conducted to identify growing firms as well as high or even hyper-growth businesses (Ardishvili et. al., 1998; Davidsson and Wiklund, 2000; Heimonen and Virtanen, 2007; Littunen and Virtanen, 2005; Littunen and Virtanen, 2006; Smallbone, Leigh and North, 1995; Virtanen and Heimonen, 2007). But identification of high growth businesses is not sufficient if we seek high success. Statistics from venture capital industry suggest that only one tenth of portfolio companies will be high successes and in one quarter of them the success will be good. However, in most of the studies growth performance has been connected with the success of the firm and in some studies growth has been used even as a surrogate of success (Birley and Westhead, 1990). This kind of approach could be defended in the studies like Smallbone, Leigh and North (1995). They used sample where high growth firms were associated with good performance so that only those businesses which showed some profit at the end of the period where classified as high growth businesses. On the other hand, in their study success was defined to be adequate if the firm demonstrated more than zero profit that cannot be seen as real sign of good profitability at the "lowest levels of success".

Virtanen and Heimonen (2007) concluded that high growth does not inevitably mean high success. The reason for this is obvious if you think about symptoms of growth. In order to grow a lot of funds will be needed and retained earnings may not cover the need for funds. But for example Markman and Gartner (2002) suggest that both sales and employment growth are unrelated to profitability in the case of extraordinary high growth businesses. Growth and success of SMEs can be approached from different perspectives depending of definitions of the concepts, their operationalisation and methods used in the analysis (Davidsson and Wiklund, 2000; Delmar, 1997; Markman and Gartner, 2002; Pasanen, 2003). This paper focuses different approaches in studying fast growth and highly successful businesses and use of different methods in analysing the data. Simultaneously the paper describes the learning process in refining the data and selection of proper methods in the analysis. The paper is organised so that first we introduce shortly research outline and research questions. Then we turn to the former literature especially those studies conducted to the problems of methodology and measurement of growth. Thereafter we will introduce the used methods and the results of estimations. Finally the conclusions and implications of the study will be given.

II. Research outline and research questions

In our previous papers, we have used both regression analysis (Virtanen and Heimonen 2007) and logistic regression (Heimonen and Virtanen 2007) in analysing almost the same data. Moreover, in Virtanen and Heimonen (2007) we used correlation analysis in studying the connection between growth and success. The inspiration for this paper was the fact that when using this kind of data you will be faced with the problems with the research design and objectives, the meaning of outliers, and the applicability of the research methods. Thus in this learning process the data should be refined to eliminate biases and to create a proper platform for a robust analysis. For example, to study growth and success simultaneously the distribution will be highly skewed. Thus

you should take into account a priori probabilities in logistic regression or in discriminant analysis (c.f. Moreno and Casillas, 2007). In our previous paper we studied success by using logistic regression analysis and concluded that the model was satisfactory even if the correct classification for successful businesses was less than 30 %. Professor Pekka Korhonen from Helsinki School of Economics advised us to use first discriminant analysis where at the first stage sample sizes of selected variables would be matched so that same amount of non-successful businesses would be selected randomly. Thereafter disrciminant analysis would be conducted with selected as well as with unselected cases.

Deriving from our experiences in this paper, we will focus growth and success simultaneously. The purpose of this paper is to test

1. What factors explain simultaneous fast growth and high success of firms?
2. What kind of impact the ignorance of a priori probabilities has on the results of the analysis?
3. What are the problems connected with different methods in analyzing differentiating factors of fast growth and highly successful SMEs and how does the selection of different analytical technique affect the results?
4. How could we improve the robustness of the data and used methods in the analysis?

In order to understand the research questions we would like to point out the dilemmas of using different methods in analyzing the data. When the regression analysis will be used properly we make the assumptions of normal distribution of data and error terms as well as equality of variances. We may affect these assumptions by scaling of variables and making logarithmic transformation and by use of dichotomic variables. However, when we for example change continuous variable to a dichotomic variable, we loose information and the interpretation of coefficient will be difficult.

The problem of explaining both high growth and high success simultaneously demands for use of dichotomic variable as a dependent variable. This demands for use of either discriminant analysis or maximum likelihood methods in analyzing the data. When we use logistic regression the interpretation of coefficients requires continuity of variables. Moreover, both in logistic regression and in discriminant analysis uneven distribution of different classes should be considered since uneven distributions mean differences in a priori probabilities. The main motivation for the study is to understand more profoundly the use of the data and different methods in analysing the performance of SMEs. In this analysis high performance means simultaneous fast growth and high success. We will use three different methods in analyzing the same data. The used methods are regression analysis, logistic regression analysis and discriminant analysis.

III. Former empirical studies

Wiklund (1998) emphasizes that prior research on entrepreneurship and SME growth focused performance studies have several problems which have their origin especially in

areas such as a lack of conceptualization and lack of use and development of theory, lack of integration of previous knowledge into research models and not sufficiently applied rigorous research methods. Further, maybe too narrow views of prior studies e.g. restricting a study to the growth of small firms by excluding other aspects of performance (e.g. success). Chandler & Lyon (2001) have reviewed 416 articles in entrepreneurship literature asking whether the methodologies and measurements employed in entrepreneurship research are sufficiently robust to foster paradigmatic growth and maturation. They evaluated the used research design and methodologies and categorized papers. About 70 % of the papers were empirical articles and regression analysis the most often used analytical technique.

Delmar (1997) studied methodological considerations in measuring growth and the pros and cons of different growth measures. by analysing the contents. He reviewed and analysed the contents of 55 studies and concluded that large array of different measures were used and that is why comparison among studies is difficult. Delmar (1997) found out that little attention was given to the choice of growth indicator, the chosen time period and the consequences of transformations of dependent variable. According to the titles of the analysed articles the concept of growth is fuzzy since 40 % of the titles referred to performance and also growth orientation and motivation was focus in some of the studies. Delmar (1998) noted that 47.3 % of the reviewed articles used regression techniques in the analysis and stepwise regression analysis was the most popular method (23% of the sample).

Davidsson and Wiklund (2000) demanded for more longitudinal research that has been lacking from entrepreneurship research. Since growth is a dynamic phenomenon this is a justified demand. However, several studies that use longitudinal data finally analyse the cross-sections – including our own studies – since the study periods are often too short to allow time series analysis.

Weinzimmer, Nystrom and Freeman (1998) was exceptional in the sense that they had solved the problem of too short time periods by using quarterly data. They studied measuring of organisational growth by using quarterly data from the years 1987 – 1991. Weinzimmer, Nystrom and Freeman (1998) used time series analysis to estimate beta so that size was regressed on time. However, some caveats could be raised especially because of the nature of the quarterly data. Even if they control seasonality and carefully justify the use of deflator these selections could be criticized. It is difficult to find a single indicator to correct seasonality since several firms, especially large ones, will be multi-branch businesses. There exist also huge differences within the same industry depending on the markets and products and / or services offered. Since the data points are the same for every business deflating the growth measures by the same deflator does not really improve the power of the analysis if different points of time will not be compared with each other. Moreover, the use of GDP price deflator does not reveal the branch specific effects of cost changes. In a stable economy the price changes between different quarters are probably small and that is why the robustness of the analysis does not inevitably require the use of deflator which could even be biased.

Littunen and Virtanen (2005, 2006) used similar variables and definitions as Smallbone, Leigh and North (1995) and discovered that those variables which quite successfully differentiated growing ventures from non-growth firms in the combined model did not include any statistically significant entrepreneur specific variables. They concluded that fit of the entrepreneur is necessary but not a sufficient condition for successful growth performance. The majority of those factors that differentiated growing ventures from non-growth companies were dependent on the strategic and operative choices of the entrepreneur.. In this study we will not be able to identify strategy changes which have been found to be important in differentiating growing businesses from non-growth firms (Littunen & Virtanen 2005). Moreno and Casillas (2007) used discrimination analysis and dichotomic dependent variable in differentiating high growth firms from non-high growth firms. They defined firm as high growth firm if its percentage of growth in 1998 – 2001 was more than 100 % higher than the median of its branch of industry.

Smallbone, Leigh and North (1995) is one of the studies where growth was connected with success. They used sample of 306 firms in eight specified manufacturing sectors in UK. They classified businesses as high, strong, and moderate growth, stable, and declining firms according to the terms of growth of real sales turnover, size and financial stability. High growth firms were associated with good performance so that only those businesses which showed some profit at the end of the period where classified as high growth businesses. However, when we would like to give some score to the level of the success the capability to break-even will not be sufficient to rate has the success been high or low.

Markman and Gartner (2002) used stepwise regression in analysing the connection between extraordinary growth and profitability of Inc. 500 companies. They tested a hypothesis that extraordinary high-growth rates – as measured by sales and employment – are negatively related to profitability. They did not find any statistically significant connection between growth and profitability. The caveats they propose include the exceptionality of the data and the so called problem of single-method bias. Markman and Gartner refer several times to those studies that point out the importance of measuring growth but do not pay much attention to the used operationalisation of profitability. Probably the results would have been different if profitability would have been continuous variable similarly as growth allowing a wider variance of profitability indicator? In their analysis the standard deviation (s.d.) of profit growth is 11 and s.d. of the growth of employment almost 4000 and sale more than 2300.

Thus the explanatory power of their model is almost non-existent. We do not argue that the relationship between growth and profitability in this case would have been as suggested in the null hypothesis but would like to point out the sensitivity of the analysis for operationalisation of variables and their measurement (c.f. Davidsson and Wiklund, 2000; Delmar, 1997). Pasanen (2003) reviews a large amount of performance studies and states that there is considerable variation in the criteria for success used in previous studies. A lot of studies are focusing on success and failure as well as key success factors of businesses.

IV. Data and methodology

In this paper we describe the refinement process of the data we have used in our analysis in different research papers. The data has been collected from different sources and thus there have been some overlapping observations. Incomplete information and problems of missing data have required restructuring and refinement of the data. The data sources we have used are "Successful firms in provinces of Finland" data base (Balance Consulting Oy 2007), and Voitto+ CD-rom data base (Suomen Asiakastieto Oy 2006 - 2008). The data characterizing innovativeness and input in R&D activities was collected from the patent data base, international patent database esp@cenet, and public financing from the sources of public funding.

In Virtanen and Heimonen (2007) the focus was on regional differences in two different regions. The data from growing and successful firms in the years 2003 – 2005 was collected from the above mentioned data bases. The firm that grew more than 10 % annually was classified as a growth business. The use of 10 % as the annual nominal growth of turnover corresponds Smallbone, Leigh and North (1995) definition of doubling turnover within 10 years. Doubling of turnover means slightly more than 7 % annual growth in ten years time. We have not deflated the figures since during the period 2003 – 2005 annual change in consumer price index was less than 1 %. Moreover, the closing month of accounting period may change from firm to firm which would skew the impact of annual inflation rate as a deflator. Using this classification criterion we identified 500 growing firms - 412 from urban and 88 from rural area.

In Virtanen and Heimonen (2007), the criterion for the fast growth was more than 30 per cent growth on the average annual sales in the years 2002 - 2005. This means that firms more than double their turnover, which is parallel to Birch,'s definition of gazelles. Birch (1979) defined gazelles as businesses, which increase the number of their employees more than 20 % annually in four consecutive years. Firms were classified as successful businesses should get a least 80 points of maximum 100 points in the used index. In Heimonen and Virtanen (2007) we used the same growth and success classification but combined "Successful businesses in provinces in Finland" data published in 2007 in the original data. Altogether we identified 466 and 101 growing businesses in urban and rural area respectively which means that the total size of the data was 567 businesses.

The closer look at the single observations revealed that we had to refine our data used in Virtanen and Heimonen (2007) and Heimonen and Virtanen (2007) in order to be able to focus really fast growth and highly successful SMEs. In several studies it has been noticed that if you measure growth with relative variables the size matters. Since some micro businesses were included in the data acquired from Balance Consulting we first removed all businesses that had less than 10 employees. Moreover, we eliminated all the firms which had more than 250 employees since they included several multinationals and their subsidiaries. The final step was to take away investment banks and finance and holding companies where the amount of assets were large compared to other companies with similar amount of personnel. Altogether we eliminated 219 companies.

The data of this study consists of 348 growing and successful SMEs. The data was collected from the years 2003 – 2005 which means that we have accounting data from the years 2002 – 2005. Similarly as in our previous paper we picked the growth businesses by using 10 % annual growth as a criterion for growing firms and at least 30 % annual growth as a gazelle. The analysis of distribution revealed that some of the businesses belonging to the category of good success (60 – 80 points, Virtanen and Heimonen, 2006) were succeeding very well. That is why we decided to choose 70 points as a limit for high success. Using this classification criterion we identified 273 NFG and/or NHS and 75 FG and HS businesses. In constructing the success index we applied international financial statements accounting standards and the instructions of Finnish advisory board of corporate analysis (YTN).

International financial statements accounting standards have been used a) to prevent random capital gains effect on annual growth and success classification of SME, b) to increase diverse branch of industry comparisons, and c) to create more holistic view of the firm level financial success including not only the profitability but also solvency and liquidity aspects of the growth firm. Success of the growth firm is defined by using the index constructed from the variables: earnings before taxes (percent), current ratio, return on investment (percent), equity ratio, net gearing (percent), repayment period of liabilities and business growth (percent). A firm should get at least 70 points of maximum 100 to be classified as a highly successful business. The definition of urban and rural areas follows the NUTS criteria. We use the density of population in relation to the density of population within the country as a criterion to classify the areas as urban and rural areas. We have selected one rural and one urban area where from we have collected data from official data bases.

We will use regression analysis (RA), logistic regression analysis (LRA) and discriminant analysis (DA) in answering the research question. The dependent variable in the RA-models will be growth of the firm measured by the average growth of annual turnover and success of the firm measured by the average success index during the years 2003-2005. LRA-models are based on the classification of growth SMEs and their comparison (0=Non Fast Growth, NFG, n=173; 1=Fast Growth, FG, n=175) as well as on the classification of fast growing and highly successful SMEs (0=Non Highly Successful, NHS, n=181; 1=Highly Successful, HS, n=167). Discriminant analysis is applied when growth and success variables are merged together and used in comparisons (0= NFG and/or NHS, n=261; 1= FG and HS, n=87).

The explanatory (independent) variables include for example age and size of the firm, branch of industry, radical innovativeness, public R&D funding (Appendix 2). Moreover, we will use region as an independent dummy variable in the growth and success estimations.

Results

Regression analysis

Regression analysis is not very much used in growth studies, which may be a consequence of the nature of the growth as a dynamic phenomenon. However, regression analysis has been used a lot in testing the performance measured by profitability of

businesses. Since we would like to explain both growth and success and operationalisation of the combined continuous variable for regression analysis would be difficult we decided to test the variables separately and analyse the similarities and differences. The hypothesis is that if growth and success are surrogates the same variables should be statistically significant variables in the model.

Stepwise regression was applied separately in growth and success models (Table 1). The best model for growth in the analysis included the variables logarithmic transformation of the amount of personnel, age, and innovativeness. Moreover profitability factor and branch of industry (trade) and region were discovered to be statistically significant explanatory variables. In the success model age and innovativeness were not statistically significant but the amount of auxiliary business names was included in the model. In addition, two dichotomic branches of industry variables (manufacturing and transportation and telecommunication) were statistically significant but they were different from growth model dummy (trade).

The results of growth model are parallel with e.g Delmar (1998) so that the younger the firm the higher growth but opposite Littunen and Virtanen (2006) and Virtanen and Heimonen (2007) who concluded parallel to Cooper (1993) that age of a firm will have a positive impact on the growth of a firm. But it should be noticed that we have used age of the firm, not age of the entrepreneur as the explanatory variable.

We would expect that the amount of radical technological innovation inputs will have a positive impact on the growth of a firm. It seems that the higher the total amount of radical innovation inputs the higher the growth of the businesses. Opposite to our previous paper (Virtanen and Heimonen, 2007) region seems to have some effect on the growth rate but not on success. Growth seems to be slightly higher in the population of capital region. The interesting detail is that the growth of the amount of personnel seems to have a negative effect on success but a positive impact on growth. Thus those businesses employing more during their development will have lower level of success. We could argue that in the short run the increase in employment would decrease the profitability of growing businesses.

All the variables appeared to be statistically significant at least at the $p < 0.05$ level in both models. In the regression analysis both in success and growth models profitability factor will be statistically significant. The sign of profitability factor suggest that the relationship with growth is inverse. This means that higher growth will be connected with lower profitability. This result does not support Markman and Gartner (1998) who did not find any connection between growth and profitability which could be seen as successful behavior. The explanatory power of the both models was quite good taking into account the cross section characteristics of the data and compared with other growth and performance studies (adjusted R2 for growth= 0.432, and for success = 0.423).

As a conclusion from the regression analysis we could answer the first question what factors explain simultaneous fast growth and high success of firms that the amount of personnel and profitability factor are the only variables that are statistically significant in

Table 1 : Regression models for growth and success

Dependent variabale: log of turnover growth	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Collinearity Statistics	
	B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
(Constant)	.583	.140		4.156	.000	.307	.859		
log growth of the amount of personnel	.703	.062	.500	11.297	.000	.581	.826	.911	1.098
log Age	-.262	.053	-.219	-4.893	.000	-.367	-.156	.893	1.120
log innovativeness	.267	.083	.139	3.209	.001	.103	.431	.950	1.052
Profitability factor	-.035	.014	-.104	-2.433	.016	-.064	-.007	.966	1.035
Branch of industry: Other industries vs. trade	-.098	.034	-.124	-2.872	.004	-.164	-.031	.959	1.043
Region	.072	.036	.088	2.009	.045	.001	.143	.923	1.083

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
6 step	,665f	.443	.432	.25657	.007	4.036	1	313	.045	1.776

Dependent variable: success index	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Collinearity Statistics	
	B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
(Constant)	72.607	1.356		53.546	.000	69.939	75.275		
Profitability factor	12.969	.978	.567	13.267	.000	11.046	14.893	.989	1.011
Branch of industry: Other industries vs. manufacturing	-13.231	2.655	-.216	-4.984	.000	-18.454	-8.008	.965	1.036
Amount of auxiliary business name	-1.098	.292	-.163	-3.764	.000	-1.673	-.524	.968	1.033
Branch of industry: Other industries vs. transportation and telecommunication	-9.953	3.058	-.142	-3.255	.001	-15.970	-3.936	.953	1.049
Growth of the amount of personnel	-.044	.018	-.109	-2.514	012	-.079	-.010	.960	1.042

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
5 step	,657e	.432	.423	.17.474	.011	6.320	1	314	.012	2.019

both models. Thus we may deduce that growth and success are not surrogates for each other (c.f. Virtanen and Heimonen, 2007). We should not mix the concepts of growth and success but they should be dealt with separately in this kind of analysis.

On the other hand, regression analysis will probably not be the proper way to explain growth and success simultaneously. The characteristics of the entrepreneurial market and assumptions of linearity, normality, continuity, and equality of variances among other things are in the conflict. Moreover, operationalisation of simultaneous growth and success will need some artificial measures where information will be loosed. In the following logistic regression and discriminant analysis simultaneous growth and success will be combined in a dichotomic variable.

Logistic regression analysis

Heimonen and Virtanen (2007) found out that the logistic regression model classified correctly 85.4 % of cases included in the analysis. NFG+NHS firms were classified successfully (97.9 %) compared with very low FG+HS grouping (22.1 %). We reported that criteria which measure the statistical applicability of the model suggest moderate fit of the model. Both Omnibus test of model coefficients and Hosmer and Lemeshow test of correct classification ($p > 0.05$) indicated that the model should be applicable in differentiating FG+HS businesses from NFG+NHS firms. However, there were severe problems with the data. In addition to multicollinearity problems of financial variables, the distribution of a priori probabilities was highly skewed. Only 16 % of the observations belong to FG + HS group. Thus the high classification rate does not defend our rating of fit of the model.

In Table 2 logistic regression analysis is presented first without taking into account a priori probabilities and thereafter with matched selection of NHS and NFG businesses. The overall classification rate of the model is almost 80 % but this is due to correct classification of NHS and NFG firms whereas the correct classification of HS and FG firms is very poor (26.7 %). It should be noticed that in logistic regression model it is always possible to achieve at least 50 % accuracy by simply setting the prediction for each observation to respond the most frequent outcome (Hoetker, 2007). Thus the model is not appropriate since it does not reach the level of a priori probability (0.5) of the HS and FG group.

When a priori probabilities are taken into account and similar amount of cases are selected to the analysis the results will change so that age will not any more be statistically significant variables. Moreover different dummies for branch of industry will be statistically significant in the forced model. The overall classification rate of the model is 63.8 %. Classification rate of selected cases is 62.3 % for N (FG + HS) firms and 65.3 % for FG + HS firms. Unselected cases all belong to N (HS+FG) and their classification rate is 66.3.

What about the interpretation of the coefficients of the model? In logistic regression model we may interpret only the coefficients of continuous variables. It could be argued

Table 2 : Logistic regression of combined growth and success variable

A) Variables in the Equation: no data transformation								
	B	**S.E.**	**Wald**	**df**	**Sig.**	**Exp(B)**	**95,0% C.I.for EXP(B)**	
							Lower	**Upper**
Amount of auxiliary business name	-.191	.109	3.046	1	.081	.826	.667	1.024
Age	-.057	.022	6.990	1	.008	.944	.905	.985
Growth of the amount of personnel	.006	.003	6.206	1	.013	1.006	1.001	1.011
Branch of industry: Services	-.889	.304	8.545	1	.003	.411	.227	.746
Business volume, debt ratio and profitability factor	-1.117	.598	3.483	1	.062	.327	.101	1.058
Venture capital factor	.606	.176	11.877	1	.001	1.833	1.299	2.586
Constant	-.243	.340	.513	1	.474	.784		

Classification Tablea				
Predicted Highly successful and high growth firm average 2003-2005				
		NHS < 70; NFG <30	HS >=70; FG >=30	Percentage Correct
Combined Success and Growth	NFG + NHS	236	10	95.9
	FG + HS	55	20	26.7
	Overall Percentage			79.8

B) Variables in the Equation: with data transformation								
	B	**S.E.**	**Wald**	**df**	**Sig.**	**Exp(B)**	**95,0% C.I.for EXP(B)**	
							Lower	Upper
	-.265	.127	4.367	1	.037	.767	.599	.984
Growth of the amount of personnel	.010	.004	6.245	1	.012	1.010	1.002	1.017
Branch of industry: Manufacturing	1.146	.553	4.293	1	.038	3.147	1.064	9.306
Branch of industry: Trade	.857	.454	3.565	1	.059	2.357	.968	5.738
Profitable capital structure	-1.161	.725	2.566	1	.109	.313	.076	1.296
Venture capital factor	.660	.238	7.721	1	.005	1.936	1.215	3.084
Constant	-1.953	.689	8.047	1	.005	.142		

Classification Tabled							
	Observed	**Predicted**					
		Selected Casesa			**Unselected Casesb**		
		HS AND FG		**Percentage Correct**	**HS and FG**		**Percentage Correct**
		NFG + NHS	FG + HS		NFG + NHS	FG + HS	
Combined Success and Growth	NFG + NHS	48	29	62.3	112	57	66.3
	FG + HS	26	49	65.3	0	0	0
	Overall Percentage			63.8			66.3

Table 3 : Discriminant analysis with (A) and without (B) matching a priori probabilities

A) Classification Results a,b (cross-validation of unselected cases)							
					Predicted Group Membership		
					NFG + NHS	FG + HS	Total
Cases Selected a.	Original	Count		NFG + NHS	50	27	77
				FG + HS	24	51	75
		%		NFG + NHS	64.9	35.1	100.0
				FG + HS	32.0	68.0	100.0
Cases Not Selected b.	Original	Count		NFG + NHS	124	45	169
				FG + HS	0	0	0
		%		NFG + NHS	73.4	26.6	100.0
				FG + HS	.0	.0	100.0
a. 66,4% of selected original grouped cases correctly classified.							
b. 73,4% of unselected original grouped cases correctly classified.							
B) Classification Resultsa,b (no data transformation)							
			FG + HS		Predicted Group Membership		
					NFG + NHS	FG + HS	Total
Cases Selected a.	Original	Count		NFG + NHS	50	28	78
				FG + HS	28	50	78
		%		NFG + NHS	64.1	35.9	100.0
				FG + HS	35.9	64.1	100.0
Cases Not Selected b.	Original	Count		NFG + NHS	117	57	174
				FG + HS	0	0	0
		%		NFG + NHS	67.2	32.8	100.0
				FG + HS	.0	.0	100.0
a. 64,1% of selected original grouped cases correctly classified.							
b. 67,2% of unselected original grouped cases correctly classified.							

that the more the firm has auxiliary business names the less probably it will be a FG + HS firm. Growth of the amount of personnel increases the odds to be classified as FG + HS firm. Branch of industry is dichotomic and thus discontinuous variable. In order to be able to deduce something about the signs of the factors the contents of those factors should be analysed first. As a summary of logistic regression analysis and answer to the question what kind of impact the ignorance of a priori probabilities has on the results of the analysis we may state that ignorance of a priori probabilities will have a significant impact on classification statistics. This means that the model does not predict the correct classification better than random selection. Even if the overall classification rate is almost 80 % the correct classification of the winners will be only 27 %. When a priori probabilities are taken into account the classification rate of FG and HS businesses will be 65 %. We could conclude that in order to be robust method in analysing skewed data cross validation should be used to take into account a priori probabilities of dichotomic groups. In LRA-models results should be interpreted differently than regression. LRA results should be interpreted as the function of estimated model. In other words the value of likelihood function expresses the probability or odds to belong the some predefined group but the interpretation presupposes continuous explanatory variables.

Discriminant analysis

If we would like to predict membership in some group we may use discriminant analysis In the discriminant analysis we will choose independent variables which we think would predict the membership in certain group.

In the discriminant analysis total cross-products variance will be partitioned into cross-products variance associated with differences between groups and within groups. The next step after finding out the overall relationship between groups and independent variables (predictors) is the analysis of discriminant functions that compose the relationship (Tabachnik and Fidel, 2007). The best discriminant functions are such that the gap between centroids of different groups is the largest.

Moreno and Casillas (2007) found that according to their definition more than 10 percentage of the total amount of firms were high growth businesses. They solve the problem of uneven distribution by dividing the non-high growth businesses into nine groups of equal size. Even if this segmentation of data solves the problem of uneven distribution it could be that selecting the same amount of non-high growth firms as there exist high growth firms randomly from the population of non-high growth businesses could have changed the results. Logarithmic transformations of growth of the amount of personnel and auxiliary business name, branch of industry dummy (services & Kibs =1) and working capital and profitability factor were discriminating variables in standardized canonical discriminant function in Model A (Table 3). In Model B logarithmic transformations of growth of the amount of personnel was not included in the discriminant function (Table 3).

Model A classifies correctly 66.4 % of the selected original grouped cases and 73.4 % of unselected original grouped cases. If we do not take into account a priori probabilities but both outcomes of dependent variable are as probable the classification rate decreases both in

selected original grouped cases (64.1 %) and in unselected original grouped cases (67.2 %). As a summary we may conclude similarly as in LRA ignorance of a priori probabilities means lower performance of the discriminant model. The obvious reason for this is that discriminant analysis is used for the classification purposes similarly as LRA also In DA the discrimant functions will be defined and the results should be interpreted to express straight intensity to belong to some group and difference between defined group centroids. In our analysis classification of cases is quite good compared to for example Moreno and Casillas (2007). Cross validation of data even improve the robustness of the model. Compared to LRA the results are parallel but in discriminant analysis we cannot tell anything about influence of single independent variables.

Conclusions and comparison of different analysis

The amount of personnel and profitability factor are the only variables that are statistically significant in regression models of growth and success. Thus we concluded that growth and success are not the same phenomenon but may even be inversely related. Regression analysis will not be the proper way to analyse high growth businesses if the opportunity for proper time series analysis will not be possible. The reason is that entrepreneurial market and growth by itself are dynamic and presuppose in most cases discontinuity, vague environment and holistic approach. If we ignore a priori probabilities and skewness of the data we may get invalid results even if the statistical indicators of the fit of the model and its characteristics will be satisfactory. There is no sense to use sophisticated statistical methods in analysing the data if it does not produce better results than a toss of a coin. The third question was that what are the problems connected with different methods in analyzing differentiating factors of fast growth and highly successful SMEs and how does the selection of different analytical technique affect the results. In Appendix 1 we have given characteristics of used statistical methods. These characteristics include modelling criteria and assumptions, evaluation criteria of models and challenges in interpretation. In all the models – as in all the statistical model analysis – the specification of the model is the basic problem. Model specification affects the validity of the results and in that sense it is of utmost importance for the robustness of the analysis. Operationalisation and measurement have an impact on reliability of the analysis. Selection of variables (Appendix 2) has an important role both from the perspective of validity and reliability.

How could the robustness of the data, and used methods in the analysis be improved. In this paper we have explained the refinement of data and compared the results of different analysis. It could be concluded that before using certain methods in analysing the data you should carefully investigate the characteristics of the data and be aware of the assumptions and aims of selected methods of analysis.

This paper is inspired by the learning experiences in analysing growth and development of growing businesses in regional context. In the course of the research process we have learned that there exist more myth than realities in the world of fast growth and highly successful businesses. These myths are driven by unconscious application of widely respected statistical methods in the problems and context where the use of these methods

will not inevitably be necessary. The same or even better results could be achieved by simple analysis of distributions and univariate statistics. However, careful analysis of data and its refinement may improve our understanding of parallel growth and success of businesses. Low and McMillan (1988) propositions to select a proper framework and define the purpose of the analysis should be remembered. The purpose and framework of the study affect the methodological opportunities and choices of the analysis. We have to acknowledge that chosen framework in this study has not enhanced our focus to exploit diverse method but we have learned through experience - through trial and error. This discovery has been very valuable and we would like to share the experience to save the extra time costs of the peers and our students.

References

Balance Consulting Oy (2007), Data base "Success firms in provinces", Eastern Finland and Uusimaa region.

Birch, D.L. (1979), *"The Job Generation Process",* Unpublished report prepared by the Massachusetts Institute of Technology Program on Neighborhood and Regional Change for the Economic Development Administration, US Department of Commerce, Washington, DC.

Birley, S. and Weasthead, P. (1990), Growth and performance contrasts between types of small firms, *Strategic Management Journal,* Vol. 11, 535-557.

Chandler, G. N., and Lyon, D. W. (2001), "Issues of Research Design and Construct Measurement in Entrepreneurship Research: The Past Decade" *ETP,* v25, n4 pp 101-112Cooper, A. C. (1993).

Challenges in predicting new firm performance. *Journal of Business Venturing* 8 (3), 241-253. Davidsson, P. & Wiklund, J. 2000.

Conceptual and empirical challenges in the study of firm growth. In D. L. Sexton & H. Landström (Eds.), The Blackwell *Handbook of Entrepreneurship:* 26-44. Oxford, Blackwell Publishers Ltd.

Davidsson, P. and Wiklund, J. (2001), "Levels of Analysis in Entrepreneurship Research: Current Research Practice and Suggestions for the Future", *Entreprepreneurship Theory and Practice,* v25, n4 pp 81-99

Davidsson, Per, & Low, Murray, B., & Wright, Mike (2001), "Editor's Introduction: Low and MacMillan Ten Years On: Achievements and Future Directions for Entrepreneurship Research, *ETP*, v25, n4, pp. 5 – 16

Heimonen, T. & Virtanen , M. (2007), The characteristics of fast growth and highly successful firms in rural and urban areas. Paper presented in RENT XXI Conference, November 22 - 23, Cardiff, UK.

Littunen, H. & Virtanen, M. (2005), Picking the winners – motivation and strategy as explaining factors of new venture growth., Paper presented in the *50th ICSB*

Conference, Washington, June 15 – 18, 2005

Littunen, H. & Virtanen, M. (2006), Differentiating Growing Ventures from Non-Growth Firms. *International Journal of Entrepreneurship and Management* (2006) 2: 93 - 109

Low, M. B. & MacMillan, I. C. 1988. Entrepreneurship: Past Research and Future Challenges. *Journal of Management,* Vol. 14, No. 2, 139 – 161

Moreno, A. M. & Casillas, J. C. (2007) High-growth SMEs versus non-high-growth SMEs: a discriminant analysis. Entrepreneurship & Regional Development; Jan, Vol. 19 Issue 1, p 69-88, 20

Pasanen, M. (2003), *In Search of Factors Affecting SME Performance, The Case of Eastern Finland.* Kuopio University Publications H, Business and Information Technology 1.

Sandberg, W. R. & Hofer, C. W. (1987) Improving new venture performance: The role of strategy, industry structure and the entrepreneur, Journal of Business Venturing, 2: 5-28.

Smallbone, D., Leigh, R. and North, D. (1995), The characteristics and strategies of fast growth SMEs, International *Journal of Entrepreneurial Behavior and Research,* 1(3), 44-62.

Suomen Asiakastieto Oy (2006), Voitto+, CD-rom. The newest financial statements data of 101 000 Finnish companies and groups.

Tabachnik, Barbara G. and Fidell, Linda, S. (2007), Using Multivariate Statistics (5 ed.), Pearson International Edition,

Virtanen, M. & Heimonen, T. (2006), The role of innovativeness in the growth and success of existing firms in Eastern Finland. Paper presented in the RENT XX – Research in entrepreneurship and small business Conference, Brussels, November 22 – 23, 2006

Virtanen , Markku and Heimonen, Tomi (2007), Comparative analysis of fast growing and highly successful firms in urban and rural areas, Paper presented in 52nd ICSB World Conference, June 13 – 15, Turku, Finland.

Appendix 1 : Characteristics of different methods

Statistical method	Modelling criteria and assumptions	Evaluation criteria	Challenges in interpretation
Regression analysis			
(RA)			
Dependent	Linearity	The fit of the model	Model specification
variable	Normality of variable	Explanatory power	Characteristics
Log Growth	Homoskedasticy	Confidence interval	of entrepreneurial market:
Success	Equality of variances	Multicollinearity	discontinuity and unstability
	Least square modelling	Singularity	Change and dynamism
	Degrees of freedom		Operationalisation of variables
			Measurement
Logistical regression			
(LRA)			
Dependent	Log linearity	A priori probability distribution	Model specification
variable combined	Normality of variable	The fit of the model	Classification and estimation
growth and success	Equality of variances	Classification statistics	power based on
FG +HS = 1	Maximum likelihood modelling	Cross validitation	probabilities and odd ratios
NFG + NHS = 0	Degrees of freedom		Operationalisation of variables
			Measurement
			Interpretation of coefficients
Discriminant analysis			
(DA)			
Dependent	Multinormal population	A priori probability distribution	Model specification
variable combined	Equality of variances of groups	Canonical discriminant	Classification and estimation
growth and success	Linear combination are compared	function	power based on
FG + HS = 1	to crossmatrix of the data	Classification statistics	disciminant function
NFG + NHS = 0	The amount of discriminant function	The gap between centroids	Operationalisation of variables
	(the amount of groups - 1)	Cross validitation	Measurement
	Degrees of freedom		

Appendix 2

Independent variables
Amount of auxiliary business names
Age of the firm
Amount of personnel
Growth of the amount of personnel
Turnover growth
Growth of the industry
Radical innovativeness
Public R&D funding
Turnover / employee
Gross income
Gross receipts / employee
Profit ratio
Quick ratio
Current ratio
Return on assets
Net gearing
Relative indebtedness
Working capital ratio
Currents assets/Turnover -ratio
Sales receivables
Accounts payable
Turnover
Net profit
Return on investment
Equity ratio
Research expenses
Development expenses
Intangible assets
Goodwill value
Other long term expenses
Dichotomic variables
Branch of Industry:
Other industries = 0, Manufacturing = 1
Other industries = 0, Construction = 1
Other industries = 0, Services including Kibs = 1
Other industries = 0, Trade = 1
Other industries = 0, Transportation and telecommunication = 1
Region
Rural area = 0, Urban area = 1
Factors:
Growth model
Profitability
Venture capital
Profitable leverage
Business volume
Success model
Profitable leverage
Business volume
Profitability
Combined model
Profitable capital structure
Solvency

Value Added Analysis: Evidence From Nepalese Commercial Banks

Ratna Man Dagnol

Abstract : *This paper examines the value added or wealth created by the Nepalese commercial banks. In total 12 banks are considered sample for the study, which has been partition into six banks in each group on the basis of the age of the sample banks. The old age banks are generated higher amount of value added in comparison to young age banks. Nepalese banks are seemed to allocate sufficient portion of value added to the future business expansion and growth, which provided ample size of capital in the context of Nepal. The study finds value allocates to employees, government, providers of capital and retained on business is significantly difference between the old and young age Nepalese banks. Value Added Statement (VAS) is a supplementary additional report of Profit and Loss Statement. Now a day, such value added statement is one of the new technique as regard a part of social responsibility accounting and reporting to the stakeholders. In the context of the Nepal, still there is no mandatory provision for preparing and reporting VAS to their stakeholders.*

I. Introduction

Commercial banks having independent and separate legal entity provide financial services to their clients and therefore, the income from banking and non-banking services are higher than that of cost of services keeping required provision against non-performing assets. The excess of the income over the cost of services is known as value added. The value added is needed to survival stability and growth of such entity within the society largely depends on the wealth created by it through the collective efforts of all the stakeholders- shareholders, providers of loan capital, employees and the government. All these stakeholders are the parties to whom the result of business operation is communicated. To satisfy the information needs of the users; the conventional financial accounting system generates data relating to financial performance through Profit and Loss Statement giving emphasis on the interest of shareholders, i.e., owners [Mandal and Goswami (2008)]. The Profit and Loss Statement does not provide any information

Mr. Dangol is Associate Professor at Nepal Commerce Campus, Tribhuvan University.
He can be reached at ----------@gmail.com