Company's Growth in Belgium, France, Italy, UK and Spain: A Search for a common Classification and Prediction Model for Growth on the Basis of Logit and Rule Induction.[1]

Arthur Limère[2]
LUC
Eddy Laveren[3]
UA
Koen Vanhoof[4]
LUC

ABSTRACT: We examine what factors drive company's growth in different European countries. We had some experience in similar but broader oriented research in the Belgian situation and we were now looking for a more universal core model that has good prediction power in different countries like France, Italy, UK and Spain with limited publicly available data. Remarkable is the degree of similarity between the resulting models in different countries and their simplicity in terms of required (accounting) data.
Persistent high profitability in the growth period, a rather good initial solvency (or high equity ratio) in the period prior to the period where growth was measured, but a degrading solvency afterwards due to debt increase were the best indicators for high growth.
Findings are consistent with pecking order theory of financing growth.

Keywords: growth; growth driver, pecking order theory, decision tree induction.

[1] This study was executed with the financial help of the Flemish government and the Fund for Scientific Research Flanders (FSR-project n°. G.0186.00).

[2] Prof. dr. A. Limère, Limburgs Universitair Centrum, Faculty of Applied Economics, Universitaire Campus, B-3590 Diepenbeek, Tel.: 011/26.86.03, e-mail: arthur.limere@luc.ac.be
[3] Prof. dr. E. Laveren, Universiteit Antwerpen, Faculty of Applied Economics, UFSIA-RUCA, Prinsstraat 13, B-2000 Antwerpen, Tel.: 03/220.40.86, e-mail: eddy.laveren@ua.ac.be
[4] Prof. dr. K. Van Hoof, Limburgs Universitair Centrum, Faculty of Applied Economics, Universitaire Campus, B-3590 Diepenbeek,e-mail: koen.vanhoof@|luc.ac.be

1. Introduction

Economic growth has always been on the agenda of the academic community. It is one of the important performance measures of companies and of our economic systems. As with human creatures and other life forms, growth is not an unlimited phenomenon and is most frequently studied in the context of life cycle models where it is supposed to be going through different stages (Greiner, 1972, 1998; Churchil & Lewis, 1984; Scott & Bruce, 1987).

A fascinating angel of incidence of research on growth has to do with the drivers of that growth. How are differences explained in growth potential of a firm? What conditions should be fulfilled to lead companies to successful growth?

Different financial factors have been stressed as being key determinants for the growth process of businesses such as profitability (Penrose, 1959; Marris, 1964), productivity (Giannakas et al., 2000; Seurat, 1999), cash flow (Carpenter-Petersen, 1998), capital structure and solvency (Durinck et al., 1997; Becchetti, 2002).

The relationship between the size of the company and growth is one of the most examined hypotheses in literature. Assuming that there exist determinants of growth essentially means that one refutes the stochastic growth process of Gibrat's Law. According to this theory, the growth rate or the proportionate growth of firms is independent of the current size of the firm. Numerous studies have been able to refute this law by detecting a negative growth/size relationship (Evans, 1987, Wagner, 1992, Variyam & Kraybill, 1992, Hart & Oulton, 1996; Becchetti, 2002; Correa, 2003).

Personnel related factors are management and personnel training (Sexton, 1997; Birley and Westhead, 1990; Cooray and Wijewadena, 1995; Leiponen, 1998; Gadenne, 1998) and labor flexibility such as temporary and part-time employment (Valverde et al., 2000; Gibb, 1991), Aparicio-Valverdi et al., 1997; Weiss,1993; Heather et al., 1996; Bielenski et al., 1992).

Other factors that are focussed on in literature as main factors and conditions for growth are research and product development (Lybaert, 1996; Mc Cann, 1991; Siegel, Siegel and Macmillan, 1993; Murray and O'Gorman, 1994; Klette and Griliches, 2000), firm location (Copus, 2001; Fujita et al., 1999) and demography (Venne, 2001). Governmental incentives as interest and capital subsidies and employment incentives are mostly received in a very critical way and are frequently considered to lead to suboptimal allocation of funds (Bergström, 2000; Leibenstein, 1966; Koevoets, 2000; Elias et al., 2000) and inefficiencies in promoting long-term growth.

The opposite of a successful growth history is failure and extensive research has been performed on conditions that lead to bankruptcy. Most research here is built on classification models like 'discriminant analysis' and 'logit regression'.(Altman, 1988; Ooghe-Verbaere, 1982; Ooghe-Joos-De Vos, 1993). More recent developments in the field of data mining techniques allow for even more detailed modelling on company failures. The success in this field suggests that similar techniques might be successful in explaining growth performances by defining groups of weak and strong growers and by selecting and applying classification methods to distinguish between those groups.

Extensive data are a prerequisite for good research efforts. Since a few decades extensive databases have become available on annual reports of firms. Specifically in a number of European countries like France, Italy and Belgium these datasets have also considerably improved for the smaller companies. In Belgium annual reports of almost all incorporated businesses are available in electronic form and outnumber the figure of 200.000. Though only a 10% of these corporations are of medium or large scale, whose reports are audited by an certified external accountant, these data offer an invaluable treasure of data for research. Even with the persisting believe in the existence of creative accounting practices, experiences with these research data were rather positive although some degree of filtering in the data is required.

In next paragraphs we will first document on an extended logit regression model for Belgium. Next we will derive a more compact core model for Belgium. After introducing a new approach of classification by means of decision tree induction, we extend the application to 4 other European countries with similar accounting publication duties. We give a number of descriptives and compare company growth with GDP-growth in all 5 counties. Classification of company growth by logistic regression and decision tree induction are executed and compared mutually and over the different countries.

2. Previous results: An extended growth prediction model for Belgium for the period 1994-1997 (Limère et al., 1999; Laveren et al, 2003)

We had some previous experience in search of growth drivers based on Belgian data sets of financial reports, extended with data from the social balance of the companies and demographic data.
30,270 Belgian (Flemish) companies were first ranked in terms of average growth in terms of total assets and added value (both equally weighted) over the period 1994-1997. A categorical variable CAVATA25 was then defined as having the value 0 for those companies belonging to the first quarter (WEAK GROWING) and the value 1 for those companies belonging to the fourth quarter (STRONG GROWING). After tens of variables, related to **Fout! Verwijzingsbron niet gevonden.** were tested in a stepwise procedure, a model with the following variables was retained with the best classification result of 81.95% of the cases (see Table 2). 5,464 cases belonged to the first and fourth quarter of growth and had observations on all independent variables.

Variable definition

LIQUIDITY measured by the acid ratio,averaged over 1994-1997;
TAX PRESSURE, being the taxes in the income statement devided by total added value and averaged over 1994-1997;
POPULATION AGE STRUCTURE , measured by means of the percentage of population that has the age below 40 year in the district of the company's location in 1994;
SOCIAL SECURITY COST, being total salaries and pension provisions over total added value and averaged over 1994-1997;
SOLVENCY, measured by the solvency ratio (equity over total assets) and averaged over 1994-1997;
PROFITABILITY, measured by return on total assets (RETTA = EBIT/TOTAL ASSETS) and averaged over 1994-1997;

TRAINING, measured by a dummy (0 if no training costs have been reported in the social balance, 1 if training costs are reported in the period 1996-1997)[5];
PART-TIME EMPLOYMENT, being the percentage of total work force that works part time and averaged over 1996-1997[5];
GOVERNMENTAL EMPLOYMENT INCENTIVES 1 through 5 being ‘employment agreement’, ‘maribel’, ‘job plan’, ‘company plan for redistribution of labor’ and ‘youth job plan’, each being presented by a dummy (0 if not participating in the incentive, 1 if participating in 1996 or 1997[5]);
INTREST/CAPITAL SUBSIDIES, measured by a dummy (0 if no subsidies have been received, 1 if interest and/or capital subsidies have been received in one of the years between 1994 and 1997);
Because of the short period and the late availability of the SocialBalance[5], no time lag was considered here between the independent variables and the dependent dichotome growth variable

Results

Regression results are shown in Table 1.
Social security cost (with a negative sign), solvency, profitability and one of the employment incentive (‘employment agreement’) were the most with growth correlated variables in the model.
Training is significant at a 1.5% level, employment incentive 2 at a 5.5% level only and interest/capital subsidies at a 4% level. All other variables are significant at a 1% level or better. Accuracy of classification is reported in Table 2.

Table 1

Variables in the Equation – extended model for Belgium

$$F(-\beta' x_i) = \frac{1}{1+\exp(\beta' x_i)}$$

$\beta' x_i = \beta_0 + \beta_1$ LIQUIDITY + β_2 TAX PRESSURE + β_3 POPULATION AGE STRUCTURE + β_4 SOCIAL SECURITY COST+ β_5 SOLVENCY + β_6 PROFITABILITY + β_7 TRAINING + β_8 PART-TIME EMPLOYMENT + β_9 EMPLOYMENT INCENTIVE 1 + β_{10} EMPLOYMENT INCENTIVE 2 + β_{11} EMPLOYMENT INCENTIVE 3 + β_{12} EMPLOYMENT INCENTIVE 4 + β_{13} EMPLOYMENT INCENTIVE 5 + β_{14} INTEREST/CAPITAL SUBSIDIES

[5] From the Social Balance which was only available in Belgium since 1996

Variabele	β	S.E.	Wald	Sig	R	Exp(β)
LIQUIDITY	0.3012	0.0712	17.8853	0.0000	0.0648	1.352
TAX PRESSURE	0.0537	0.0143	14.0960	0.0002	0.0565	1.055
POPULATION AGE STRUCTURE	0.1142	0.0315	13.1478	0.0003	0.0543	0.892
SOCIAL SECURITY COST	0.1168	0.0081	206.405	0.0000	0.2323	0.890
SOLVENCY	0.0340	0.0039	74.5095	0.0000	0.1384	0.967
PROFITABILITY	0.0847	0.0104	66.4267	0.0000	0.1304	1.088
TRAINING	0.4843	0.1981	5.9771	0.0145	0.0324	1.623
PART-TIME EMPLOYMENT	0.0180	0.0026	48.4700	0.0000	0.1108	0.982
EMPLOYMENT INCENTIVE 1	1.9439	0.1288	227.674	0.0000	0.2441	6.986
EMPLOYMENT INCENTIVE 2	0.2525	0.1314	3.6936	0.0546	0.0211	1.287
EMPLOYMENT INCENTIVE 3	0.9115	0.1543	34.8896	0.0000	0.0932	2.488
EMPLOYMENT INCENTIVE 4	3.5068	1.0094	12.0692	0.0005	0.0516	33.342
EMPLOYMENT INCENTIVE 5	0.4439	0.1740	6.5117	0.0107	0.0345	1.559
INTEREST/ CAPIT SUBSIDIES	0.3023	0.1471	4.2249	0.0398	0.0242	1.353
CONSTANT	5.6895	1.2668	20.1720	0.0000		295.745

$\chi^2 = 1645.997$ $p=0.0000$

Classification:

Table 2

Classification Table – extended model for Belgium

		Predicted		Percentage Correct
Observed		0	1	
CAVATA25	0	1140	244	82.37%
	1	249)	1099	81.53%
		Overall percentage		81.95%

One of the new research objectives was aimed at finding a more universal prediction model of growth that could be based on limited information, available in most annual reports as they are published in an international context. These ideas came after a closer look at the Belgian results. We came to the conclusion that up to 78% of classifications results where explained by as few as 2 to 3 key factors. This model is described in next paragraph.
Legal context of accounting and accounting principles tend to differ between countries although steady efforts are been made for more harmonisation in Europe. An interesting question here is whether similar drivers entail growth in industries over different countries as United Kingdom, Italy, France and Spain? The choice of these countries is no coincidence and is related to availability and extension of published accounts for a broad category of companies including SME's.
Cross-validation was not performed with an old-out sample but by applying the same analysis to a set of Walloon companies and a mixed set of Belgian firms. Results were very close to the ones reported above with a classification result of more then 80%.

A core model for Belgium for the period 1996-2000 ;new data and definition.

28,366 Belgian companies were ranked in terms of average growth of total assets over the period 1998-2000. To make the data set comparable as to scale of the companies with available data sets abroad, we limited the cases to have at least 3 full time equivalent employees. Only cases were retained of companies that published the annual report on a regular basis between 1996 and 2000. A categorical variable CATA25 was defined as having the value 0 for those companies belonging to the first quarter (WEAK GROWING) in terms of average growth in total assets over the period 1998-2000 and the value 1 for those companies belonging to the fourth quarter (STRONG GROWING).
After numerous models that were tested in a stepwise procedure on the basis of available accounting data and financial ratios we kept 3 independent variables:
ROE9800 being the average return on equity over the period 1998-2000, the growth period under study;
SOLV9897 being the average solvency over the period 1996-1997 which is the period immediately preceding the growth period under study;

SOLV9800 being the average solvency over the period 1998-2000, the growth period under study;
Notice the build-in time lag between the first observations on solvency and the growth period. The distinction between both periods as far as solvency is concerned happened a bit by coincidence but led to a very interesting result and interpretation: Initial status of companies before growth, as far as solvency is concerned, is positive but solvency degrades after the period of growth is entered, as result of this growth and its requirements in terms of financial structure.
Next we tried to classify the companies in both categories by the solely means of these financial features.

Results

Descriptives of these independent variables are shown in Table 5.
The results of the logistic regression are presented in Table 3 and Table 4. Overall 77.7% of the cases were classified correctly.

Table 3

Variables in the Equation - core model for Belgium

$$F(-\beta' x_i) = \frac{1}{1+\exp(\beta' x_i)}$$

$$\beta' x_i = \beta_0 + \beta_1\ ROE9800 + \beta_2\ SOLV9897 + \beta_3\ SOLV9800$$

	ß	S.E.	Wald	Sig.	R	Exp(ß)
ROE9800	0.048	0.001	1413.721	0.000	0.2653	1.049
SOLV9697	0.073	0.002	1297.843	0.000	0.2483	1.075
SOLV9800	-0.082	0.002	1543.359	0.000	-0.2703	0.921
Constant	0.014	0.038	0.131	0.718		1.014

χ^2 =3133.954 P=0.0000

ROE9800 = average returnon equity over 1998-2000[6]
SOLV9697= average solvency over 1996-1997[6]
SOLV9800= average solvency over 1998-2000[6]

Table 4

Classification Table - core model for Belgium

[6] Although multicollinearity between the independent variables is present (lower then 0.80) it is not that much of a problem since regression is only used for forcasting and classification, not for confidence intervals.

		Predicted		Percentage Correct
Observed		0	1	
CATA25	0	5559	1508	78.7
	1	1649	5467	76.8
Overall Percentage				77.7

CATA25 = 0: weak grower in total assets over 1998-2000
= 1: strong grower in total assets over 1998-2000

Interpretation

All regression coefficients are significant except for the constant term.
The classification variable Low-Growth/High-Growth CATA25 was positively correlated with the average return on equity during the growth period and the average solvency in the period preceding the growth period and negatively correlated with the average solvency in the growth period. Stepwise regression removed the variable of average return on total assets in the period preceding the growth period (RETTA9697) as well as return on equity in the same period (ROE9697), although they also had a positive coefficient when enforced into the model.
Companies seem to have better growth perspectives if they are able to and persistent in reaching good profitability performances during expansion. Their initial status before growth, as far as solvency is concerned, is positive but solvency degrades after the period of growth is entered, as result of this growth and its requirements in terms of financial structure. Clearly growth interferes negatively with solvency but the initial start in terms of solvency should be rather positive. Notice in the descriptives for Belgium in Table 5 how solvency in the period prior to the growth period (SOLV9897) on average is better for strong growers then for weakgrowers but decreases in the growth period (SOLV9800), while it increases for weak growers.

Table 5

Descriptives for Belgium of average growth and independent variables for weak en strong growers (%)

	avrg		stdv	
CATA25	0	1	0	1
Growth9800	-5.83	31.55	5.71	16.02
SOLV9697	34.32	34.81	21.76	21.80
SOLV9800	38.00	31.85	21.88	19.89
ROE9800	1.47	12.95	20.60	19.17

1stQ		median		3rdQ	
0	1	0	1	0	1
-7.80	20.00	-4.20	26.05	-1.94	37.58
17.20	17.45	29.75	30.50	48.10	49.00
20.57	16.27	34.40	27.70	52.67	44.20
-4.43	4.27	2.87	12.37	10.07	21.93

Another approach: rule induction

We wanted these results to be confirmed by applying new data-mining techniques. We looked for a better understanding of the very specific contribution of solvency before and during the growth period and applied a rule induction algorithm to the same data.

The purpose of a decision tree is to classify cases for a dependent variable on the basis of a set of rules for the independent variables. In this paragraph, the decision tree induction technique will be explained briefly.
A decision tree is a collection of *branches* (paths from the root to the leafs), *leafs* (indicating classes in a concept) and *nodes* (tests to be carried out).
A decision tree is built by means of recursive partitioning. This means that the sample is split up in different subsamples and these subsamples are further split up etc. The technique uses two sets of data, namely a training set and a test set. A decision tree is built on the basis of the training set and is tested by means of the test set.

In each stage of the building of the decision tree, the algorithm behind the technique will choose the best splitter, which is the variable (attribute) that splits up best the data in subsamples where one certain class of cases dominates. To determine which is the best splitter, the algorithm (in this study the c 4.5 algorithm is used – see Quinlan, 1986; 1993) will try every possible split of each variable. For each subsample, the best splitter will then be specified. This process proceeds until further splitting would not cause a significant improvement of the model. The algorithm is a Top-Down Induction of decision trees which means that the there will be no recursions in the tree. As a result the first used attributes in the algorithm are the most important ones
The splitting criterion is based on maximizing *information gain* , Gain(S, A) of an attribute A, relative to a collection of examples S of some target concept (weak/strong growth):, i.e. the expected reduction in entropy of this set A by knowing the values of attribute A.

$$Gain(S, A) \equiv Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

$$with \quad Entropy \equiv \sum_{i=1}^{c} - p_i \log_2 p_i$$

where Values(A) is the set of all possible values for attribute A, and S_v is the subset of S for which attribute has value v. p_i is the proportion of S belonging to class i (in casu weak grower/strong grower). |S| denotes the number of cases in set S.
This approach is based on information theory (Hunt et al., 1966).
The logarithm is base 2 and is a measure of the expected encoding length measured in *bits*. If the target attribute can take on c possible values, the entropy will be at maximum $\log_2 c$. If a node of the tree contains only cases of the same class, the entopy is equal to zero (minimum). *Entropy* can be seen as a measure of chaos that should be minimized This algoritm terminates when al subsets (nodes in the tree) are labelled or when no further attributes splitting the unlabeled sets are available.

Next to symbolic attributes, also numerical attributes are feasible. These numeric attributes undergo a binarization in the algorithm, that means thresholding their numeric ranges into pairs of subintervals to be treated as symbols.
After the decision tree is built, it will be pruned to avoid overfitting. The algorithm splits up the data set in subsets that will be smaller and smaller and the final subsets will be no longer representative for the population. As such, the model will incorporate structures that are found in the data set on which the model is estimated, but that are not representative for the population. To prevent this from happening, the tree will be pruned, which means that branches of the tree will be deleted and replaced by a leaf.
The test set is then used to find the best tree in terms of accuracy.

The main advantage of data mining techniques when compared to the more classical data analysis technique of regressions is that there are no assumptions for the underlying distribution of the data. It is a non-parametric classification technique. The advantage of decision tree induction to other data mining techniques is that it clearly shows which rules (and thus which variables) are used to classify the cases. As such, the importance of all variables in the classification of the cases can clearly be identified.
Decision tree learning methods are robust to errors, both errors in classifications of the training examples and errors in the attribute values that describe these examples of a concept (data noise). The training set may contain missing attribute values.

Results and interpretation of the tree

The found decision tree for Belgium, which has an accuracy of 70.91% is shown below.

Figure 1 BELGIUM

```
ROE9800 =< 8.033
   SOLV9800 =< 38.4
      SOLV9697 =< 33.75
         SOLV9800 =< 22.533
            ROE9600 =< -2.38 (1219.0, 0.662) -> 0
            ROE9600 > -2.38 (1314.0, 0.553) -> 1
         SOLV9800 > 22.533 (1195.0, 0.826) -> 0
      SOLV9697 > 33.75 (1003.0, 0.654) -> 1
   SOLV9800 > 38.4 (2917.0, 0.784) -> 0
ROE9800 > 8.033 (6725.0, 0.677) -> 1
```

ROE9800 and ROE9600 are the average return on equity over the total period 1996-2000 and the growth period 1998-2000.
If Return on Equity is larger then 8.033%, all cases are classified as belonging to the high growth group of firms (6,725 cases; 0.677 confidence).
If smaller then 8.033% classification happens in a more complex way and conditional on the variables SOLV9800, SOLV9897 and ROE9600. Similarities with the logit model come from requiring SOLV9897 to be **larger** then a certain limit value (here: 33.75%) and SOLV9800 (in the growth period) to be **smaller** then a certain value (here: 22.533%) for cases to be classified as growing fast, the latter in combination with a minimal average return over the total

period of –2.38%. Numbers of classified cases and confidences are mentioned in the tree as well as the classification rules suggested.
Cross-validation is performed in tenfold by dividing the sample in 10 equal parts and assigning nine portions to the training set and one portion to the test set. Portions are then interchanged and the best accuracy is selected.

Company growth in France, Italy, UK and Spain; a comparison with Belgium

We queried databases of annual reports in 4 other European countries over the period 1994-2000, except for Spain where the same data were only available since 1995. We preferred to have a somewhat longer period to better flatten out short-term cyclic movements in growth figures. The number of annual reports in each country is:

FRANCE	28,654
ITALY	14,614
UK	18,778
SPAIN	25,234

Logistic regression

Table 6 shows the results of the regressions for the 4 European countries, other then Belgium.We entered the same explaining variables. Because of the availability all data of Spain are lagging with one year. Remarkable are the very similar results as for Belgium. Classification results are given in

Table 7.
When using stepwise regression the solvency ratios were always in the best fitting equation with the same signs. The only difference was the profitability ratio ROE9600 that sometimes was replaced by ROTTA9600 or a combination of both. Signs were consistent and improvement in classification results only very slight compared to the reported classification results. In the latter case the partial regression coefficient of profitability was divided over both ROE and ROI..
All regression coefficients are significant, equal in sign but somewhat different in magnitude between the countries. Overall classification results are similar and of the order of 75%, except Spain that with 81.3 % outperform the Belgian result of about 78%. (core model)

Table 6

Replication of Logit regression over 4 European countries

$$F(-\beta' x_i) = \frac{1}{1+\exp(\beta' x_i)}$$

$\beta' x_i = \beta_0 + \beta_1\, ROE9600 + \beta_2\, SOLV9495 + \beta_3\, SOLV9600$ (france, italy and uk)

or $\beta' x_i = \beta_0 + \beta_1\, ROE9700 + \beta_2\, SOLV9596 + \beta_3\, SOLV9700$ (spain)

france	B	S.E.	Wald	Sig.	R	Exp(B)
ROE9600	0.045	0.001	1533.734	0.000	0.2374	1.046
SOLV9495	0.049	0.002	558.210	0.000	0.1511	1.050
SOLV9600	-0.063	0.002	872.063	0.000	-0.1986	0.939
Constant	-0.149	0.044	11.482	0.001		0.861
χ^2 =	2219.493		P=0.0000			
italy	**B**	**S.E.**	**Wald**	**Sig.**	**R**	**Exp(B)**
ROE9600	0.069	0.003	519.088	0.000	0.2826	1.072
SOLV9495	0.071	0.005	219.171	0.000	0.1832	1.073
SOLV9600	-0.104	0.005	443.149	0.000	-0.2610	0.901
Constant	0.345	0.062	30.481	0.000		1.411
χ^2 =	1106.843		P=0.0000			
uk	**B**	**S.E.**	**Wald**	**Sig.**	**R**	**Exp(B)**
ROE9600	0.032	0.001	569.856	0.000	0.2831	1.032
SOLV9495	0.031	0.002	164.816	0.000	0.1516	1.031
SOLV9600	-0.048	0.003	341.911	0.000	-0.2190	0.953
Constant	-0.202	0.077	6.807	0.009		0.817
χ^2 =	1169.557		P=0.0000			
spain	**B**	**S.E.**	**Wald**	**Sig.**	**R**	**Exp(B)**
R0E70	0.054	0.002	1266.404	0.000	0.3176	1.055
SOLV9596	0.072	0.002	878.044	0.000	0.2644	1.074
SOLV9700	-0.089	0.003	1211.805	0.000	-0.3107	0.914
Constant	-0.268	0.058	21.438	0.000		0.765
χ^2 =	3305.957		P=0.0000			

Table 7 Classification results for 4 European countries

france

		Predicted 0	Predicted 1	Percentage Correct
Observed		0	1	
CATA25	0	5262	1589	76.8
	1	2003	4886	70.9
		Overall Percentage		**73.9**

italy

		Predicted 0	Predicted 1	Percentage Correct
Observed		0	1	
CATA25	0	1898	554	77.4
	1	643	1583	71.1
		Overall Percentage		**74.4**

uk

		Predicted 0	Predicted 1	Percentage Correct
Observed		0	1	
CATA25	0	2555	472	84.4
	1	875	1309	59.9
		Overall Percentage		**74.2**

spain

		Predicted 0	Predicted 1	Percentage Correct
Observed		0	1	
CATA25	0	4336	687	86.3
	1	1020	3066	75.0
		Overall Percentage		**81.3**

The rule induction algorithm: results and interpretation

Just like with the Belgian situation we applied the rule induction algorithm. The 4 decision trees are given below.Each tree presents the decision rules with on the right hand side eventual classification (1 being strongly growing firms, 0 being weakly growing firms) and between brackets the number of cases and the significance expressed as a probability. Accuracy results of the test set are as follows:

France	70.1%
Italy	67.6
UK	69.5
Spain	72.0

Figure 2 FRANCE

- ROE9600 =< 14.638
  - SOLV9600 =< 46.95
    - ROE9400 =< 5.854 **(2838.2, 0.718) -> 0**
    - ROE9400 > 5.854
      - SOLV45 =< 37.49
        - SOLV9600 =< 29.356
          - SOLV45 =< 21.02 **(900.1, 0.589) -> 0**
          - SOLV45 > 21.02 **(524.8, 0.648) -> 1**
        - SOLV9600 > 29.356 **(671.5, 0.748) -> 0**
      - SOLV45 > 37.49 **(515.2, 0.644) -> 1**
  - SOLV9600 > 46.95 **(2014.6, 0.826) -> 0**
- ROE9600 > 14.638 **(6862.6, 0.7) -> 1**

Figure 3 ITALY

- ROE9600 =< 9.334
  - SOLV9600 =< 26.544
    - RETTA9600 =< -0.024 **(748.4, 0.7) -> 0**
    - RETTA9600 > -0.024
      - ROE9600 =< 1.81
        - SOLV9600 =< 13.678 **(752.2, 0.524) -> 1**
        - SOLV9600 > 13.678 **(484.9, 0.624) -> 0**
      - ROE9600 > 1.81 **(1239.9, 0.6) -> 1**
  - SOLV9600 > 26.544 **(1390.0, 0.786) -> 0**
- ROE9600 > 9.334
  - RETTA9600 =< 1.086 **(302.4, 0.565) -> 0**
  - RETTA9600 > 1.086
    - SOLV9600 =< 22.394 **(1429.5, 0.775) -> 1**
    - SOLV9600 > 22.394
      - ROE9600 =< 16.774 **(427.9, 0.521) -> 0**
      - ROE9600 > 16.774 **(531.6, 0.692) -> 1**

Figure 4 UK

- ROE9600 =< 23.63
  - SOLV9600 =< 29.894
    - RETTA9600 =< 1.906
      - SOLV9600 =< 10.812 **(541.9, 0.512) -> 1**
      - SOLV9600 > 10.812 **(747.7, 0.658) -> 0**
    - RETTA9600 > 1.906 **(744.3, 0.599) -> 1**
  - SOLV9600 > 29.894 **(3118.5, 0.742) -> 0**
- ROE9600 > 23.63 **(4236.6, 0.687) -> 1**

Figure 5 SPAIN

ROE9700 =< 18.863
SOLV9700 =< 31.358
SOLV9596 =< 29.09
SOLV9700 =< 16.585
RETTA9700 =< -0.75 **(504.2, 0.643) -> 0**
RETTA9700 > -0.75 **(932.7, 0.576) -> 1**
SOLV9700 > 16.585 **(1160.2, 0.703) -> 0**
SOLV9596 > 29.09 **(626.9, 0.666) -> 1**
SOLV9700 > 31.358 **(3977.6, 0.781) -> 0**
ROE9700 > 18.863 **(5415.4, 0.731) -> 1**

These classification results are clearly poorer then those attained with logit regression but give us a clear understanding of the decision rules used to classify cases, rules that can be interpreted. Two decision table trees, those of Italy and UK, only selected variables from the growth period (indicated with the name-extension “9600” standing for 1996-2000), while the other two use variables in their decision tree from the pre-growth period as well.. In 3 of the 4 trees, just like with Belgium, cases were classified on one hand side as being strongly growing , only on the basis of return on equity, having a value larger then 14.64% in the case of France (6862 or 48% of the cases) , larger then 23.63% in the case of UK (4236 or 45% of the cases) and larger then 18.86% in the case of Spain (5415 or 43% of the cases). No more conditions are imposed here. Remember that the profitability measure for Belgium had a turning point value of 8.03%. The fourth tree, that of Italy, is a little more complex on this side of the tree. It requires a profitability of equity of 9.33%, combined with a return on total assets of minimal 1.09%. One should be aware that return on total assets doesn’t yet reflect the possible financial leverage as opposed to return on equity, so we could expect here a smaller value. Additionally growth should be signalled here with a low solvency in the growth period (smaller or equal to 22.39%). If last condition is not fulfilled, return on equity should be minimal 16.77%. So far the lower side of the tree.
At the upper hand side all four trees are more complex and ask for more conditions to be fulfilled for being a growing firm. First of all in every of the four countries solvency in the growth period should now not exceed a maximum value of 46.95% in France, 26.54% in Italy, 29,89% in UK and 31.36% in Spain. We had already a similar condition above for Italy, but not to the same degree. The meaning of this requirement can be interpreted as being the price that should be paid for growth: giving in on solvency as a result of debt financing after internal generated funds are in shortfall.

Next to this condition the trees are very similar two by two. For Italy and UK, to be classified as strong growing companies, either profitability should exceed a certain relatively low value (1.81% and 1.09%) but measured by return on equity for Italy and measured by return on total assets for UK, or solvency in the growth period should still further be limited to lower values. (13.68% for Italy and 10.81 for UK, which suggests even more debt financing of the growth).
For France and Spain solvency in the pre-growth period should now be of a minimum level being 37.49% in France (21.02% if solvency in the growth period is equal or less then 29.36%) and 29.09% unconditionally in Spain. In France this is coupled at an extra profitability

requirement of 5.85% in terms of ROE. In Spain the last condition can be compensated by a minimum return on investment (RETTA9700) of -0.75 if solvency in the growth period is extremely low (<16.59%).

It has to be stressed that in many ways the results of rule induction are consistent with that of the logit regression. All variables appear with the same signs but sometimes in different combinations. Profitability remains the main driver for growth, mainly in the period of growth, but sometimes differently measured (ROE or RETTA). Companies that persist in being profitable in the growth period are most likely to belong to the category of strongest growers. Strong growing companies see their solvency decrease considerably in the growth period itself, as a result of more debt financing of their new investments.
However classification results are somewhat poorer then with logit regression.

Conclusion:

Main purpose of this contribution was aimed at a common prediction model for company growth in 5 European countries who compare very well in terms of accounting publication duties for SME's as well as for larger corporations. Taken into account that there are still differences in the degree of disclosure and in applied accounting principles, the model should be simple, slim and still have a good classification power between 'strong growers' and 'weak growers'. This dichotome variable was defined by means of 1st quarter and 3th quarter observations on average growth in total assets in the observed period. All annual reports of the five countries should be able to provide the explanatory data.
One of the research question was wether such a model would exist that would be quite equivalent in these countries. The empirical research was based on similar data sets from financial statements available in all 5 countries.

The application of logistic regression to this data , bore us a first model that gave similar results in different countries: Persistent high profitability in the growth period, a rather good initial solvency in the period prior to the period where growth was measured, but a degrading solvency afterwards were the best indicators for high growth. Classification results differed but mounted from a low of about 74% (Italy) till a high of 82% (Spain).
To understand better the awkward relation between the three dependent variables, we applied decision tree induction methods. Classification results were somewhat poorer but could be interpreted in a consistent way. Some of the previous mentioned variables appeared in these models in a conditional way upon each other. So is the condition on solvency depending on the degree of profitability, which seems to be the first selected variable in the algorithm for the 5 countries. Conditions on solvency in the prior and posterior period show the same positive and negative relation towards growth, but are introduced in the trees consecutively. Interpretations are similar as with the signs in the regression model. Higher growth lead to a giving in on solvency as a result of the need for funds and the pecking order theorema in finance which bear higher debt ratios, if internal funds are not available. These conditions are superfluous if profitability is high. Indeed, companies then are more capable to provide internally generated funds.

REFERENCES

Altman E. I. (1988), *The prediction of corporate bankruptcy: a discriminant analysis*, New York: Garland

Aparicio-Valverde, M. Kabst, R., Brewster C. and L. Mayne (1997), “Conclusion: The Flexibility Paradox.”, *Employee Relations*, 19 (6), pp. 596-608.

Becchetti, L. and G. Trovato (2002), “The Determinants of Growth for Small and Medium Sized Firms. The Role of the Availability of External Finance”, *Small Business Economics 19:* 291-306.

Bergström, F. (2000), “Capital Subsidies and the Performance of Firms.”, *Small Business Economics*, 14, pp. 183-193.

Bielenski, H., Alaluf, M., Atkinson, J., Bellini, R., Castillo, J. J., Donati, P., Graverson, G., Huygen, F. and J. Wickham (1992), “New Forms of Work and Activity: A Survey of Experiences at the Establishment Level in Eight European Countries”, Working Paper, European Foundation for the Improvement of Working and Living Conditions.

Birley, S. en P. Westhead (1994), “A taxonomy of business start-up reasons and their impact on firm growth and size.”, *Journal of Business Venturing*, 9, 7-31.

Carpenter, R., Petersen, B., (1998), “Is the growth of Small Firms Constrained by Internal Finance?”, Research Paper, Emory University, 33 p.

Churchill, N.C. en V.L. Lewis (1983), ″The five Stages of Small Business Growth.″, *Harvard Business Review*, Vol. June, 30-50.

Cooray, S. en H. Wijewardena (1995), “Determinants of Growth in Small Manufacturing Firms: the Japanese Experience*.”, Journal of Small Business Management,* vol. 33, n° 4, 87- 105.

Copus A. K. (2001), “From Core-periphery to Polycentric Development: Concepts of Spatial and Aspatial Peripherality.”, *European Planning Studies*, 9 (4), pp. 539-552.

Correa, A., M. Acosta, A. L. Gonzales and U. Medina, “Size, Age and Activity Sector on the Growth of the Small and Medium Firm Size”, *Small Business Economics 21:* 289-307.

Durinck, E., Laveren E., Lybaert N. (1997), "The impact of sales growth above a sustainable level on the financing choice of Belgian Industrial SME’s.", Working Paper, UFSIA, Antwerpen.

Elias, B., Termote, A., Van Eeckhoutte W., Omey, E. and S. Demeestere (2000), “Het Belgisch Tewerkstellingsbeleid inzake bijdrageverminderingen, hervedeling en creatie van arbeid: een juridische analyse.”, *Belgisch Tijdschrift voor Sociale Zekerheid*, 3e trimester, 42 (3), pp. 621-636.

Evans, D.S. (1987), ‘″The relationship between firm growth, size and age: estimates for 100 manufacturing industries.’″, *The Journal of Industrial Economics*, Vol. XXXV, pp 567-581.

Fujita, M. , Krugman, P. and A. Venables (1999), *The Spatial Economy; cities, regions and international trade*., Cambridge, MA: MIT Press.

Gadenne, D. (1998), “ Critical Success Factors for Small Business: An inter-industry Comparison.”, *International Small Business Journal*., Vol. 17, n°1, 36-56.

Giannakas, K., K.C. Tran and V. Tzouvelekas (2000), “Efficiency, technological change and output growth in Greek olive growing farms: a Box-Cox approach”, *Applied Economics*, 32, pp. 909-916.

Gibb, A. (1991), “ Appraising growth potential: key issues in the thresholds of growth of the independent business.”, in Rource, J. and S.

Greiner, L.E. (1972), “Evolution and revolution as organizations grow.”, *Harvard Business Review*, July-August, 37-46.

Greiner, L.E. (1998), ″Evolution and revolution as organizations Grow.″, *Harvard Business Review*, May-June, 55-67.

Hart, P., Oulton, N. (1996), 'Growth and Size of Firms', *The Economic Journal*, Vol. 106, September, pp 1242-1252.

Heather, P., Rick, J., Atkinson, J. and S. Morris (1996), "Employer's use of temporary workers.", *Labour Market Trends*, September, pp. 403-412.

Hunt, E. B. , Martin, J., & Stone, P. J., (1966), *Experiments in Induction*, New York, Academic Press.

Klette, T. J., and Z. Griliches (2000), "Empirical patterns of firm growth and R&D investment: A quality ladder model interpretation.", *The Economic Journal*, 110, April, pp. 363-687.

Koevoets, W. (2000), "Hoofdstuk 2 : De effectiviteit van vermindering van patronale bijdragen: micro-economische analyse van tewerkstellingseffecten van het jongerenbanenplan.", *Belgisch Tijdschrift voor Sociale Zekerheid*, 3e trimester, 42 (3), pp. 701-716.

Leiponen, A. (1998), "Competencies, innovation, and Profitability of Firms.", in Eliasson, G. , Green, C. Microfoundations of Economic Growth: a Shumpeterian Perspective, The University of Michigan Press, Michigan, 254-270.

Leibenstein, H. (1966), "Allocative Efficiency vs. 'X-efficiency'", *The American Economic Review*, 56 (3), pp. 392-415.

Laveren, A, A. Limère, K. Cleeren, and E. Vanbilsen (2003), "Growth Factors of Flemish Enterprises: an Exploratory Study over the period 1993-1997", *Brussels Economic Review – Cahiers Economiques de Bruxelles,* Vol. 46 – n°1, Spring 2003.

Limère, A., E., Laveren, R., Mercken, E., Vanbilsen (1999), 'Onderzoek naar de groei van de Vlaamse ondernemingen doorheen de gepubliceerde jaarrekeningen, afleiding van indicatoren voor een gericht overheidsbeleid.', Study executed under the authority of the Flemish Government, Diepenbeek.

Lybaert, N. (1996), "Information Use as Counter-Proof for the Stochastic Growth Theories.", LUC, Diepenbeek, Paper presented at the RENT X Conference Brussels, Belgium.

Marris, R.L. (1964), *The Economic Theory of Managerial Capitalism*, MacMillan.

McCann, J. (1991), ″Patterns of Growth, competitive technology, and financial strategies in young ventures.″, *Journal of Business Venturing*, Vol. 6, 189-208.

Murray, J.B. en O'Gorman, C., (1994), "Determinants of high growth: a study of the strategic attributes and the determinants of growth in small & medium sized enterprises.", working paper, EFER, 1994.

Ooghe, H., Joos, P., De Vos, D. (1993), "Risico-indicator voor een onderneming aan de hand van falingspredictiemodellen.", Accountancy en Bedrijfskunde, 18 (3), september, pp. 1-25.

Ooghe, H., Verbaere, E. (1982), "Determinanten van faling: verklaring en predictie.", RUG Accountancy, bedrijfskunde an beleidsinformatie, Gent.

Penrose, E.T. (1980), *The theory of the growth of the Firm*, Oxford, Blackwell.

Quinlan, J. R. (1986), "Induction of decision trees", *Machine Learning,*1(1)*, 81-106.*

Quinlan, J. R. (1993), *C4.5: Programs for Machine Learning,* CA: Morgan Kaufmann.

Scott, M. en R. Bruce (1987), "Five Stages of Growth in Small Business.", *Long Range Planning*, Vol. 20 , n° 3, 45-52.

Seurat, R. (1999), "Sustained and Profitable Growth", *Business Strategy Review*, 10 (1), pp. 53-56.

Sexton, D., Upton, B., Wacholtz, L. McDougall, P. (1997), ″Learning needs of growth-oriented entrepreneurs.″, *Journal of Business Venturing*, 12, pp. 1-8.

Valverde, M., Tregaskis, O. and C. Brewster (2000), "Labor Flexibility and Firm Performance.", *International Advances in Economic Research*, 6 (4), pp. 649-661.

Weiss, W.H. (1993), "The use of temporary and part-time help in the workplace.", *Supervision*, 54 (11), pp. 3-6.

Variyam J.N., Kraybill, D.S. (1992), 'Empirical evidence on determinants of firm growth'", *Economic Letters* Vol. 38, pp 31-36.

Venne, R. A. (2001)," Population Aging in Canada and Japan: Implications for Labour Force and Career Patterns.", *Canadian Journal of Administrative Sciences*, 18 (1), pp. 40-49.

Wagner, J. (1992), '"Firm Size, Growth, and Persistence of Chance: Testing GIBRAT's Law with Establishment Data from Lower Saxony, 1978-1989'", *Small Business Economics,* Vol. 4, pp 125-131