"An Empirical Taxonomy of Start-Up Firms Growth Trajectories"

Diambeidou, M.B ; Wertz, Vincent ; Verleysen, Michel ; Gailly, Benoît ; François, Damien

Document type: Communication à un colloque (Conference Paper)
AN EMPIRICAL TAXONOMY OF START-UP FIRMS GROWTH TRAJECTORIES

Mahamadou BIGA DLIAMEIDOU, Damien FRANCOIS,
Benoît GAILLY, Michel VERLEYSEN, Vincent WERTZ

www.crecis.be
AN EMPIRICAL TAXONOMY OF START-UP FIRMS GROWTH TRAJECTORIES

MAHAMADOU BIGA DIAMBEIDOU, DAMIEN FRANÇOIS, BENOÎT GAILLY

MICHEL VERLEYSEN, VINCENT WERTZ

Abstract

Dealing with the challenge generated by the study of the complex nature of new firm growth process, this article provides an original method that can accommodate, in a systematic way, the analysis of the new firm growth trajectories. Our approach applies principal component and cluster analysis to a large sample of 741 firms using multidimensional aspect of new firm growth over time and across sector to identify in a systematic way distinct growth stages. We then use sequence analysis and a Markov chain approach to extract and compare the trajectories of individual firms over time. This allows the identification of 4 distinct growth stages which were used to identify 7 typical growth trajectories that are valid for the majority of firms in our sample. This indicates that, based on empirical evidence, new firm growth trajectories is a complex process but nevertheless not idiosyncratic (nor completely random). Our findings indicate that this original systematic approach is useful for taxonomy development and therefore contributes to reduce the gap between the complexity of new firm growth process and the standards approaches often used to deal with it. They also have several implications and limitations that will be discussed.

Key Words

Entrepreneurship - Growth trajectory - longitudinal research method - Sequence analysis

Authors

Mahamadou BIGA DIAMBEIDOU is PhD candidate in Entrepreneurship at Université catholique de Louvain (UCL). For more details, see www.crecis.be

Damien FRANÇOIS is PhD and Research Assistant in Mathematical Engineering at UCL. For more details, see www.dice.ucl.ac.be/mlg/

Benoît GAILLY is Professor in Innovation Management and Strategy at UCL. For more details, see www.crecis.be

Michel VERLEYSEN is Professor in Machine Learning at UCL. For more details, see www.dice.ucl.ac.be/mlg/

Vincent WERTZ is Professor in Mathematical Engineering at UCL. For more details, see www.dice.ucl.ac.be/mlg/
Introduction

Over the past decades, new and small firm growth has received considerable attention from researchers and policy-makers around the world. New firms have been identified as engines of growth, innovation and wealth creation (Audretsch and Thurik, 2000; Birch, 1981; Davidsson, 1995; Davidsson Lindmark and Olofsson, 1998; Levie, 1997; OECD, 1994, 1998, 2002; Storey, 1994, 1995; Welbourne, 1997). Empirical evidence indicates that only a small proportion of firms accounts for a significant percentage of new job creation. Those firms often accelerate the development of new technologies and products that play a fundamental role in the prosperity of many countries (Birch, Haggetty and Parsons, 1997; Julien et al., 2001; Storey 1997). New firms are therefore a key element in regional economic development and represent as such an interesting research subject.

Despite their importance to regional development, knowledge about new firm growth is still scattered (Davidsson and Wiklund, 2000; Delmar, 1997) and little knowledge is available regarding how firms grow and perform over time (Geroski, 2001). This can be partly attributed to methodological problems, such as the difficulties experienced in identifying entrepreneurial firms. For example, to this concern Gibb and Davies (1990) argued that it is illusory to think that it would be possible to detect this kind of firm or to produce a complete ideal model. Other international studies concluded that a “typical” high growth firm does not exist (OECD, 2000; Delmar et. al., 2003).

From a theoretical perspective, scholars have shown that research has largely failed to generate cumulative results regarding new firm growth because there is a great variability in researchers’ use of growth conceptualization and operationalization (Davidsson and Wiklund 2000, Delmar 1997; Delmar et al, 2003; Murphy et al, 1996 Chandler & Hanks, 1993; Weinzimmer et al, 1998, Wiklund and Shepherd, 2005). Indeed from a theoretical perspective the phenomenon of entrepreneurial growth has been mostly studied within individual academic disciplines, which does not encourage an integrated and systemic analysis (Garnsey & Hefferman, 2005). Research was done from different theoretical imperatives such as industrial organization, resource-based view, strategic adaptation and evolutionary economic perspective. Thus, research from
one discipline tends to ignore important findings from competing schools. For example, since the original “theory of the growth of the firm” from Penrose (1959), where managerial resources played a pivotal role, various factors have been suggested as affecting growth. Some of them, such as environmental carrying capacity or market forces, are external to the organization (Aldrich, 1990; Singh and Lumsden, 1990). Others are internal, like capabilities, culture or strategy and have been mainly addressed from the resource-based view of the firm (Wernerfelt, 1984; Teece et al., 1997; Boeker, 1997; Zahra et al., 2000; Canals, 2000).

From an empirical perspective, growth remains a multi-faceted phenomenon, but this heterogeneous nature is often neglected by scholars. Despite the diversity of approaches in terms of indicators, formulae and time spans used to measure growth (Delmar, 1997), empirical research has also largely failed to generate cumulative results (Davidsson and Wiklund 2000; Delmar 1997; Delmar et al, 2003; Weinzimmer et al, 1998). The common explanation is the absence of consensus regarding which firm growth indicators should be used (Weinzimmer et al, 1998; Delmar, 1997; Murphy et al, 1996 Chandler & Hanks, 1993). Another explanation is that researchers often measure growth along a single dimension (Weinzimmer et. al., 1998) although this approach has been widely criticized as firm growth is heterogeneous in nature (Birley & Westhed, 1990; Delmar and Davidsson, 1998; Delmar et al, 2003). As a consequence, using a single measure of growth defined by a single criterion actually investigates only one particular kind of growth and the results are unlikely to be applicable to other forms of growth (Delmar and Davidsson, 1998).

Finally, most studies about growth tend to focus on specific sectors (for example high-tech) although the economic contribution of new firm growth appears to be spread across various sectors (Delmar et al, 2003). Indeed most papers looking at “promising firms” have focused on samples limited to new technology based firms, from sectors such as software products, telecommunications or biotechnology (Baldwin, Chandler and Papailiadis, 1994; Vyakarnam, Jacobs and Handelberg, 1997; Woywod and Lessat, 2001; Calvo and Lorenzo, 2001; Julien 2001).

In order to address those theoretical and empirical issues, recent entrepreneurship research argues that there is a strong need for a conceptual scheme and for longitudinal
growth studies (Busenitz et al, 2003; Chandler & Lyon 2001; Davidsson & Wiklund 2000, Delmar et al 2003, Garnsey et al. 2006; Pettigrew et al. 2001). The underlying assumption is that growth is a heterogeneous phenomenon that naturally happens over time; it should therefore be analyzed in a dynamic process perspective and across multiple organizational contexts. Indeed, while most new and small firm growth studies have focused on the explanation of the performance using cross-sectional data and/or have assumed that growth is an uninterrupted process, longitudinal approaches have shown that regular growth is the exception rather than the rule (Delmar et al., 2003; Garnsey et al., 2006; Garnsey & Heffernan, 2005; McMachon, 2001; OECD, 2002; Stam, 2003).

However longitudinal approaches generate methodological challenges which require new research methods (Huber and Van de Ven, 1995; Poole et al, 2000; Van de Ven, 1992; Van de Ven and Engleman, 2004), which in particular involve more than static comparisons between beginning and final states (Davidsson, 2004; Davidsson and Wiklund 2000). Based on their respective dissertations and works in the field, Davidsson, Delmar and Wiklund (2006, p.5) argue that “firm growth is a complex phenomenon. It is not uni-dimensional. It is hard to predict and assess. Further, it can manifest itself in various ways, and consequently it can have differential effects on several different levels”.

In this context, the purpose of this research is to present an original method that can accommodate, in a systematic way, the analysis of new firm growth trajectory based on a multidimensional construct of organic growth. More specifically, our objective is to establish whether the early growth of a firm is a process basically idiosyncratic i.e. related to the individual characteristics of each firm or typical growth trajectories exist and are adopted by a majority of firms. We analyzed the growth trajectories of 741 Belgian firms created between 1992 and 2002 and which have grown above micro-firm size. We developed and tested an original methodology allowing an empirical taxonomy of early growth trajectories across multiple sectors, integrating the multidimensional aspect of growth.
In the following sections we detail our research design and introduce the experimental setting. We then document the empirical results of this research and discuss major findings and limitation.

**Research design**

To deal with the challenges generated by the analysis of the complex nature of new firm growth process, we have elaborated a research design at the crossroad of entrepreneurship research and applied mathematics (see Lévesque, 2004, for the value added of the use of mathematics approach in entrepreneurship research, in term of theory development). More specifically, at a conceptual level, our research design integrates multidimensional and dynamic approaches of organic growth across sectors in order to conceptualize new firm growth process. The key aspects of this research design are addressed hereafter.

**Organic growth**

Although this aspect is often neglected by scholars, growth can be achieved in different ways (Delmar, 1997; Delmar et al., 2003; Levie, 1997; Roberts & Berry, 1985). The main illustrative example is organic growth versus growth by acquisition. However, Penrose (1959) emphasizes that organic growth should be more related to small firms and new ventures whereas growth by acquisition is a phenomenon more exhibited by older, larger firms in mature industries. In addition, according to Davidsson, Delmar and Wiklund (2006, p32) only organic growth is a legitimate interest for entrepreneurship research and remains a proxy for entrepreneurship.

Our research design therefore deals only with organic growth, in particular at the early stages of development of the firm.

**Multidimensional approach and ubiquity**

When considering the measurement of growth, there is as discussed in the introduction no consensus regarding which and how many indicators to use. Moreover, the majority of researchers do not justify theoretically their choice of variables, although those can have consequences on the results (Delmar 19996; Janssen, 2005). We therefore choose
for this research design to jointly use multiple indicators, based upon commonly used measures of firm size and financials.

Moreover, in the context of our research, we adopt the assumption that firm growth is an ubiquitous phenomenon. In other words, we will not limit ourselves a priori to specific sectors, on the basis of the assumption that firms develop in various manners whatever their sector and that the relevance of the sector dimension should be checked empirically a posteriori and not preconceived a priori.

**Growth trajectory**

As discussed above, growth is a process of change that needs to be studied over time (Davidsson, Delmar and Wiklund, 2006, p40). Indeed, Penrose (1995) saw growth as a cumulative process in which firm members build knowledge and competence. According to this author, firms are “[…] a result of development process […] in which interacting series of internal changes lead to increase in size accompanied by changes in the characteristics of the growing object” (Penrose, 1995:1).

In the entrepreneurship context, recent authors emphasized the underlying assumption of the growth process, suggesting that firm growth is driven by a “productive opportunity” (Penrose, 1995) in a cumulative process of interaction between the firm’s productive base and its market opportunities (Garnsey, 1995, Garnsey et al 2006). Thus, process studies of firms should examine interconnected causes, outcomes and further feedback effects (Van de Ven, 1992). We need longitudinal research because it mainly allows direct observation of change, causal statements, temporal context and feedback effects (Davidsson & Wiklund 2000; Chandler & Lyon 2001; Pettigrew et al. 2001).

As a consequence, our research design will focus on taking into account successive measures of the growth process as the firm evolves over time, rather than only considering the initial and end-state.

**Empirical Taxonomy**

As emphasized by most scholars in organizational studies (Archibugi, 2001; de Jong and Marsili, 2006; Ketcher and Shook, 1996; Pavitt, 1998; Rich, 1992), a useful
empirical taxonomy can reduce the complexity of empirical phenomena to a few constructs. Thus, a widely accepted and usable taxonomy is a fundamental element in the development of scientific body of knowledge (Sabherwal and Robey, 1993) and can serve as an empirical based framework to theory development. Indeed, previous researches suggest that, contrary to a typology considered as an individual creativity invention, taxonomy is an empirical classification tool for building the complex filing systems that allow both the ordering and retrieval of large amounts of data (McKelvey, 1975; Pugh, Hickson, and Hinings, 1969). Moreover, according to Rich (1992), a taxonomy is more than a simple classification of items into separate groups. It is a specific classification scheme that expresses the overall similarity between organisms in a hierarchical fashion. In addition, in their innovative small firm research, de Jong and Marsili (2006) emphasized that taxonomy classifies and labels many different items into groups or clusters that share common traits.

We therefore adopted in this research a taxonomy approach based on advanced applied mathematics tools, with the objective to attempt to reduce the complexity and therefore to improve our understanding of new firm growth trajectories. Hence our research design consists in mobilizing advanced applied mathematics tools in order to develop an empirical taxonomy of firm trajectories based on the evolution across sectors and over time of multiple size and financial indicators.

**Experimental setting**

Having specified our research design, we will present briefly in this section the methodology adopted to identify the typical growth trajectory, with regard to the choice of the sample and variables and to the methods of analysis adopted.

**Choice of sample and variables**

A valuable opportunity to address the key study issue in this chapter has been provided by the availability of the BEL-FIRST database developed by the Bureau van Dijk Electronic Publishing (BvD), one of Europe's leading electronic publishers of business information. Our research collected longitudinal financial information and demographic indicators about all Belgium firms. This study focusing on the initial growth of the firm,
the population considered here includes all the firms created after 1992 and still in existence in 2002 (N = 152,064).

Among these young firms, we selected all those which, since their creation, exceeded the stage of micro-firm level (as defined by the European Commission, 2003). This enables us to identify firms whose growth can be regarded as “promising” in the broadest sense, i.e. which contributed to economic development. This allows us to build a sample that goes beyond exceptional cases of very high and regular growth, often publicized but not at all representative of a “typical” firm. In order to eliminate most “false creation” cases (such as an existing firm turning part of its business into a subsidiary), we eliminated firms that already exceeded the size of a micro-firm at the time of their creation.

Regarding the choice of variables, we selected as our main measure of growth three indicators: Sales, Employment and Total Assets, which have all been considered as suitable indicators of growth (see Davidsson and Wiklund, 2000). Those indicators are combined with seven financial variables traditionally linked to firm performance (Value added, Operating income, Current income, Net income, Cash-flow, Working capital and Shareholders’ equity). In line with previous researchers (Davidsson and Wiklund 2000; Birley and Westhead, 1990; Weinzierl et al 1998, Wiklund, 1999), we think that the combination of multiple size and financial indicators gives richer information and therefore allows better investigating the firm growth process. Finally our data also included information regarding each firm’s main sector of activity, type of ownership and legal form. On the basis of this choice of variables, the firms for which available data were sufficient and coherent or could be reconstituted by simple interpolation were selected.

Methods of analysis

The method used to analyze the existence of typical growth trajectories consists in considering a firm growth trajectory as a sequence of states (corresponding to the successive years of existence) in a space with 10 dimensions (corresponding to the 3 size and seven financial indicators). In this space, mathematical tools for classification
and discrimination such as a principal components analysis\(^1\) (PCA) and empirical clustering based on density estimation (Cuevas, 2001) can be mobilized (see also François et al, 2004 for other examples of tools for classification). These tools make it possible to identify potential clusters, each cluster corresponding to firms in similar stages of development. Once these clusters are identified and tagged (“stage A”, “stage B”, etc...), the trajectory of a firm can be described as a sequence or Markov chains corresponding to the various stages it experienced successively. For example AAABABB represents seven years of the trajectory of a firm evolving between states close to the clusters “stage A” and “stage B”.

The firms having adopted similar growth trajectories will be characterized by similar sequences and those sequences can then be compared and analyzed through graphical interpretation, sequence analysis (Poole et. al.2000), Markov chain analysis (Bakerman and Gottman, 1986; Howard, 1971) and density distribution in order to evaluate the heterogeneity of growth trajectories and test the existence, validity and characteristics of typical trajectories.

**Empirical results**

We will present our results in four sections. The first section relates to the sampling results and to the validation of the use of multiple indicators and sectors. The second section presents the result of the clustering of the successive growth states into four “stages”. The next two sections explore the heterogeneity of growth trajectories and the existence and characteristics of typical trajectories through first a graphical analysis and then a systematic analysis.

**Sampling results**

As discussed above, our research is based on a sample including all Belgian firms created since 1992 and still in existence in 2002 (n = 152,064). From these firms, we selected those which grew above micro-firm size. There were 17,168 such firms identified in our sample. Those “promising” firms represented 6% of all the existing Belgian firms in 2002 and 11% of the firms created since 1992 and still in existence in

\(^1\) PCA is basically a projection method (Lawley and Maxwell, 1971) which can be used to develop cluster-based taxonomies (Evrard et al., 2003; de Jong and Marsili, 2006).
2002. However, they generated (in 2002) respectively 19% and 80% of overall gross job creation.

Among the firms selected, 33% had missing values regarding Employment, 53% regarding Sales and 17% regarding Cash flow. We excluded the firms which did not publish complete data for more than two consecutive years, or firms which published less than 40% of the data available for two consecutive years. Using those filters, the final data set included 741 firms.

Most sectors were represented in our sample (13 out of the 17 principal sectors included in standard NACE industry classification). The majority related to service industries (71%), while only 11% were related to Manufacturing and another 17% to the construction industry. Only 19% of the firms in our sample were technological and innovative firms. This confirms the relevance of our cross-sector approach.

Finally, the 741 firms in our sample were sorted according to the measure of size (employment, sales or total assets) along which they had grown above micro-firm size. 50% of the firm in our sample had reached only the employment threshold (more than 9 employees) while respectively 33% and 8% had reached only the sales and the assets thresholds. Less than 10% had reached more than one threshold and 2% had reached all three thresholds (they simultaneously had more than 9 employees, more than one million euro in sales and more than two million euros in assets). This confirms the relevance of our use of multiple indicators to measure growth (Delmar et al, 2003; Janssen 2005; St-Pierre et al, 2005), as using different indicators leads to different selections of “promising” firms.

**Clustering**

In order to explore the growth trajectories of the firm in our sample, we first tested the existence of clusters among the various states a firm experiences as it grows. We used for this purpose a Principal Component Analysis based on the successive absolute value

---

2 To identify the firms related to “high tech” sectors, we defined a conversion table between the Code NACE-BEL and the Code US SIC for the classification of industry sectors according to technological intensity (from the Bureau of the Census and Walcott (2001)). We then validated our approach by matching this with other existing classification approaches like OECD 1998 industrial sector classification and some new innovative firm classification indicators in the literature.
of our 3 size and 7 financial indicators for each firm. This analysis produced three principal axes, with a cumulated variance of 82% (the two first axes account for 72%). Other choices of variables were tested, in particular using relative rather than absolute financial values (ratios). Nevertheless, all these alternative choices proved less relevant in terms of restitution of information, i.e. generating a weaker cumulated variance.

The three principal axes identified enable us to represent all the successive states of the firms in a space with three dimensions. These axes represent composite variables which can be regarded as “latent dimensions” of the problem (Evrard et al., 2003), making it possible to apprehend fundamental dimensions of the studied phenomenon. The three axes are detailed in Table 1 below, according to their correlation with the 10 starting variables after a Varimax rotation.

***INSERT TABLE 1 ABOUT HERE***

Table. 1: Results of PCA analysis

Those results indicate that the first axis is more correlated with 4 financial variables (Operating Income, Current Income, Net Income and Cash flow) which can be linked to the “Performance” of a firm. The second axis is more correlated with 4 indicators (Employment, Sales, Value added and Total assets) which can be linked to the “Size” of a firm. Finally, the third axis is more correlated to the variables Working capital and Shareholder equity, which can be linked to the “Resources” of a firm.

As the “Size” axis identified through the PCA can be considered as the dependent variable of this study, we focused our graphical analysis on the “Performance” and “Resources” axes. We can therefore represent the successive states any firm in our sample goes through as it grows along those two axes, taking as a reference point the average value of the sample (Figure 1).

---

3 Our initial variables being standardized, the coefficient of correlation is a good indicator of the relation between a variable and a principal axis (Evrard et al, 2003).

4 This algorithm of rotation is based on the maximization of the coefficients of correlation of the most correlated variables (Hendrickson and White, 1964; Kaiser 1958).
Figure. 1: Two-dimensional representation of firm successive states

Hence the results of the clustering allowed us to identify three principal axes, of which one ("Size") can be considered as the dependant and two ("Resources" and "Performance") can be used to extract significant information relative to the evolution of each firm over time, across our sample. However those axes are only numerical constructs produced by the PCA, which can only approximately be related to actual dimensions of the firms and have no direct managerial interpretation for a given firm.

In order to test whether the two principal axes could be used in order to define meaningful clusters (i.e. whether they relate to managerial differences between actual firms), we measured the density distribution of all 10 size and financial variables and of four commonly used financial ratio (return on equity, return on asset, capital productivity and labor productivity) along the two axis.

As an illustration, the results of the density distribution for the variable Operating income are presented in Figure 2. The four curves represent the distribution of the variable Operating income for firms situated in the four quadrants of Figure 1 (corresponding respectively to above and below average value along the two principal axes). This figure shows that there appears to be a cutting point (at around 30,000 Eur of annual operating income) between the distribution curves related to the bottom-left and top-left quadrants (below average “performance”) and the distribution curves related to the bottom-right and top-right quadrants (above average “profitability”).

Figure 2: Density distribution of operating income along the principal axes
The analysis presented in Figure 2 indicates that the firms which achieve a higher value along the “performance” principal axis are indeed different from a managerial point of view, as their operating income will be significantly higher.

The application of this process to the 10 size and financial variables and to the four financial ratios indicate that the “performance” principal axis is significantly related to operating income, current income, net income, cash flow and labor productivity while the “resources” principal axis is significantly related only to shareholder equity. Combining those two axes and their managerial interpretation allows us therefore to identify four different stages a new firm can reach as it grows:

1. “Questions” are firms located at the bottom-left of Figure 1. They tend to combine lower than average operating income, current income, net income, cash flow and labor productivity (low performance) with lower than average shareholder equity (low resource). Their future development might at first sight seem at risk.

2. “Seeds” are firms located at the top-left of Figure 1. They tend to have lower than average operating income, current income, net income, cash flow and labor productivity (low performance) but higher than average shareholder equity (high resource). A firm associated to this profile could for example have been able to raise relatively important funds to ensure its initial development but needs time to improve its performance.

3. “Boutiques” are firms located at the bottom-right of Figure 1. They tend to have higher than average operating income, current income, net income, cash flow and labor productivity (high performance) but lower than average shareholder equity (low resource).

4. “Stars” are firms located at the top-right of Figure 1. They tend to combine higher than average operating income, current income, net income, cash flow and labor productivity (high performance) with higher than average shareholder equity (high resource). Their future development seems a priori promising.

Hence the PCA and the density analysis have allowed us to identify two axes along which four stages of growth can be identified which are both empirically valid (from the PCA) and correspond to actual managerial dimensions (from the density analysis); We will discuss in the next two sections how those two axes and four stages can be
exploited from a graphical and systematic point of view in order to test the heterogeneity of the growth trajectories of young firms and build a taxonomy.

**Graphical analysis**

This section aims to explore the existence of growth trajectories from a graphical analysis and by using the two first principal axes. Several representations of firm trajectories using the two principal axes we identified above are presented in Figure 3, where each box represents the successive states adopted by a given firm along the two axes. In this figure, similar trajectories have been gathered together. The first group (the top six boxes) represents rather linear trajectories, from high size /low performance to low size /high performance states. The second group gathers sigma-shaped trajectories. They indicate that the growth of those firms has not been smooth over the years, with some periods that may even correspond to decay. The third group (bottom line) presents angular trajectories going up first (increase in size) then bifurcating to the left (increase in performance).

***INSERT FIGURE 3 ABOUT HERE***

**Figure 3**: Examples of two dimensional projections of 18 firm growth trajectories.

Figure 4 presents an alternative graphical representation of growth trajectories within (Figure 4a) and across (Figure 4b) sectors. In that figure, the horizontal axis represents the successive years of development of the firm, while the vertical axis represents the value reached along the “performance” and “size” principal axes. Those two axes are aggregated into one single dimension by defining nine different categories, each related to a different value (low, medium or high relative to the average of the sample) of “performance” and of “size”.

***INSERT FIGURE 4a ABOUT HERE***

**Figure 4a**: 4 growth trajectories within the same sector (wholesale industrial)
These preliminary graphical analyses indicate that there appear to be at the same time a wide diversity of trajectories but also groups of trajectories that are similar and that are worth investigating further, in a more systematic way.

**Systematic taxonomy**

The results of the clustering presented above allowed us to identify four distinct growth “stages”, across which all firms in our sample evolve over time. The trajectory of a given firm can therefore be characterized by the sequence of successive stages it goes through. In order to identify those sequences in a systematic way, we used both a Markov chain approach and a sequence analysis approach. The corresponding results are presented hereafter.

i. **Markov chain approach**

The way firms in our sample move from one stage to another (dependencies) can be represented through a digraph (Figure 5), which details the probability that a firm starts in a given stage (boxes) and move from one stage to another (arrows). Such a digraph renders visible how stages are sequenced over time (Bakerman and Gottman, 1986).

The results presented in Figure 5 indicate that the probability to remain within a given stage is quite high (p > 0.61 for all for stages). This behavior can be interpreted as firm inertia. Moreover, “magic recoveries” and “catastrophes”, i.e. directly moving from low levels of performance and resource (“questions” stage, n°1) to high levels (“stars” stage, n°4) and reciprocally is very rare (respectively p < 0.02 and p < 0.03). However this representation of the data collected does not allow us to identify typical trajectories. This aspect will be discussed in the next section.
ii. Sequence analysis

In order to explore the existence of typical growth trajectories based on the characterization of firm trajectories as sequences of successive stages, we first considered only the four first years of existence of each firm. There were 602 firms in our sample (out of 741) for which sufficient data was available. The initial trajectory of each of those 602 firms can be characterized by a four-long sequence of stages (four years) among the four possible stages we identified through our clustering (“seeds,” “stars”, “boutiques” and “questions”), leading to 256 ($4^4$) theoretical possibilities. The distribution of the 602 sequences among the 256 possible ones is presented in Figure 6.

***INSERT FIGURE 6 ABOUT HERE***

**Figure 6: Distribution of growth trajectories observed.**

The high concentration of sequences observed in Figure 6 indicates that they are not uniformly distributed. In particular, only 115 different sequences have actually been observed, and only 22 correspond to more than 1% of the firms. Those 22 sequences collectively cover 71% of the firms and are presented in Table 2 hereafter.

***INSERT TABLE 2 ABOUT HERE***

**Table 2: Most frequent sequences.**

If one focuses on transitions between stages rather than on the time spent within each stages (the sequences 4333, 4433 and 4443 are then regarded as equivalent), 7 typical trajectories emerge from those 22 sequences, which collectively correspond to 59% of the firms in our sample. The resulting typical trajectories are presented in Figure 7 below.

***INSERT FIGURE 7 ABOUT HERE***

**Figure 7: Typical trajectories.**
Those 7 typical trajectories can be further split between four “stable” ones (a firm remains in a given stage over the time period considered) and three “unstable” ones. The 4 stable and 3 unstable trajectories are described hereafter (Table 3), by decreasing frequency.

***INSERT TABLE 3 ABOUT HERE***

Table 3: Taxonomy of growth trajectories

This taxonomy shows a great heterogeneity of growth trajectories both in terms of frequency and development paths. However the detailed characterization of each one of these trajectories would require individual case studies which exceed the framework of this research. Indeed, no significant correlations were observed (through \( \chi^2 \) - test) between the 7 typical growth trajectories identified and individual measures of growth (sales, employment or assets) or between the growth trajectories and particular demographic data (sector, type of ownership or legal form).

The results presented above indicate that, based on empirical evidence, new firm growth trajectories are neither “linear”, nor a random or idiosyncratic phenomenon, and that a taxonomy of seven typical trajectories corresponding to a majority of firms can be identified.

Discussion and conclusions

Growth has been extensively studied in the fields of entrepreneurship, strategic management and industrial organization. However most studies have concentrated on the explanation of growth using cross-sectional data or case studies and supposed that growth is essentially an uninterrupted process. However, the longitudinal studies of growth suggest that regular (or linear) growth is the exception rather than the rule.

With the aim to contribute to a better understanding of the growth process of new firms, we have presented an original methodological approach allowing the systematic analysis of growth trajectories based on a multidimensional construct of growth across sectors. This method made it possible to track systematically typical growth trajectories
and to identify 4 clusters which were used to identify 7 typical growth trajectories which corresponded with 59% of the new firms considered in our sample.

Our findings indicate that this original systematic approach is useful for taxonomy development and therefore contributes to reduce the gap between the complexity of new firm growth process and the standards approaches often used to deal with it. They also have several implications and limitations, that will be discussed hereafter.

**Implications**

Our empirical results suggest that new firm growth appears to be neither a continuous (or life-cycle based) nor idiosyncratic (or completely random) process. It can be adequately described through a limited number of typical growth trajectories that can be identified in a systematic way. This original contribution of our research has important theoretical and practical implications.

Our results imply that a cross-sectional approach can fail to capture the complex reality of the evolving new firm. They support the recent studies based on longitudinal approaches such as those of Delmar et al. (2003), Garnsey et al. (2006), Garnsey and Heffernan (2005), Mustar (2002), McMachon (2001) and Stam and Garnsey (2006).

In addition, our results show that organizational growth constitutes a multiform and cross-sector phenomenon by nature, which cannot therefore be reduced to a single dimension or a single sector. Thus, the results of our analysis support a multidimensional conceptualization of growth, contrary to what most researches used, particularly in the studies on the determinants of the growth process (Birley and Westhead, 1990). Moreover, they reinforce the recent work emphasizing the heterogeneity of new firm growth (Delmar et al., 2003; François et al. 2004; Janssen, 2005 and Weinzimmer, 1998). It raises the question of the relevance of the uniform quantitative approaches, focused only on the criterion of a high relative growth rate in sales and/or employment, adopted by policymakers and venture capitalists to evaluate the potential of a new firm.

Our findings are important from a theoretical perspective because their bring insight regarding how start-ups evolve over time. They contribute to our understanding and
appreciation of the heterogeneity of the growth trajectory phenomenon. Thus, researchers should develop more nuanced explanations of new firm growth process than the simple uninterrupted or “linear” dynamic growth process. Our findings indicate that nonlinearities exist in new firm growth trajectories and emphasize the rarity of very high growth trajectories.

In addition, this research provides a useful original taxonomy of new firm growth trajectories calling for explanations from case studies. Our taxonomy also extends previously developed taxonomies by taking into account the firm’s financial configuration as a complementary building block of the growth process scheme.

This study also provides a useful methodology contribution by showing the value added of the use of advanced applied mathematics to deal with the complex and dynamic nature of firm growth and therefore to contribute to theory development. We provided an original methodological approach based on systematic clustering and classification to derive an empirical taxonomy. We also provide a conceptual framework and helpful graphical tools for describing, characterizing and comparing new firm growth trajectories. In sum, our approach supports the recent work (Lévesque, 2004) emphasizing how mathematics can provide important contributions to current theories of management and organizations.

**Limitations**

This study has several limitations that should be noted. Firstly, in order to examine growth trajectories, we needed a significant number of longitudinal data which justified the use of retrospective secondary data limited to only one country. The nature of these data does not enable us to measure the effect on the evolution of new firms of some important qualitative factors such as strategies, entrepreneurial motivation and human capital.

Our research is also limited by the natural selection bias excluding failed firms in the sample. Therefore, we have no way of knowing what type of distinct growth trajectories such firms can exhibit, and how these growth paths might differ from the 7 typical growth trajectories identified in this research.
Although these limitations are important and must be taken into accounts, we nevertheless are convinced that this study should contribute to a better understanding of new firm growth process.

**Future research directions**

This research opens many future research directions. Firstly, the replication in various contexts of the developed empirical taxonomy and specific data analysis methodology should be considered.

Secondly, while this research can be considered as a first step towards a better understanding of the start-up growth trajectories, further research is needed to improve our understanding of the dynamic growth process of new ventures. It should explore which endogenous and exogenous factors might explain why start-ups follow the 7 identified typical growth trajectories. To this end, our retrospective approach and data analysis methodology should be complemented by other methods such as surveys and multiple case studies.

In this context, among other orientations, it is highly relevant to further our taxonomy by examining the relationship between innovative and technological sources and growth trajectories in high versus low technological industries. Moreover, this relationship controlled by environmental characteristics – at the level of internationalization for example - will add knowledge to our understanding of a new firm growth process.

Finally, while our findings provide empirical and methodological support in new firm growth research, one of the most important future research directions is to test the accuracy of the proposed taxonomy in the stability of the firm dynamic growth process beyond the limited period retained in this research.

**Acknowledgements**

This research is supported by a grant from the National Bank of Belgium. Part of this work is supported by the Belgian Program on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its authors. The work of D. François is funded by a grant from the Belgian F.R.I.A. Preliminary versions of this paper were presented at the European Summer University 2006 on entrepreneurship and entrepreneurship education research (Hameenlinna), at ERIM workshop “Perspectives on the Longitudinal Analysis of New Firm Growth” in 2005 (Rotterdam)
and, at CRECIS - Louvain School of Management seminar in 2006 (Louvain-la-Neuve). We thank Frédéric Delmar, Alain Fayolle, Saras D. Sarasvathy, Born Bjerke, Aard Groen, Paula Kyrö, Erik Stam, Elizabeth Garnsey, Frank Janssen, Regis Coeurderoy, Nathalie Delobbe and anonymous reviewers for their helpful comments.
References


OECD (1994), The OECD Jobs Study, Paris
OECD (2000), High-Growth SMEs and Employment, Paris
OECD (2002), High-Growth SMEs: Their Contribution to Employment and Their Characteristics, Paris


Table 2: Results of PCA analysis

<table>
<thead>
<tr>
<th>Axis</th>
<th>Empl. Sales Value added</th>
<th>Oper. inc.</th>
<th>Current income</th>
<th>Net income</th>
<th>Work. capital</th>
<th>Cash flow</th>
<th>Shar. equity</th>
<th>Total assets</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Perf.</td>
<td>0.12 0.03 -0.06</td>
<td>-0.47 -0.49 -0.51</td>
<td>-0.12</td>
<td>-0.44</td>
<td>-0.23</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Size</td>
<td>0.42 0.47 0.45</td>
<td>0.03 -0.01 -0.11</td>
<td>0.26</td>
<td>0.03</td>
<td>0.027</td>
<td><strong>0.50</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Res.</td>
<td>0.33 0.12 0.29</td>
<td>0.14 0.07 0.05</td>
<td><strong>-0.71</strong></td>
<td>0.20</td>
<td><strong>-0.47</strong></td>
<td>-0.04</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure. 1: Two-dimensional representation of firm successive states

Axis 1: “Performance”
- Operating income
- Current income
- Net income
- Cash Flow

Axis 2: “Size”
- Sales
- Employment
- Total assets
- Net added-value

Axis 3: “Resources”
- Working capital
- Shareholder equity

“Resources”

“Performance”
Figure 2: Density distribution of operating income along the principal axes
Figure 3: Examples of two dimensional projections of 18 firm growth trajectories.
Figure 4a: 4 growth trajectories within the same sector (wholesale industrial)
Figure 4b: 6 distinct growth trajectories from 4 sectors
Figure 5: Growth trajectories digraph.
Figure 6: Distribution of growth trajectories observed.
Table 2: Most frequent sequences.

<table>
<thead>
<tr>
<th>Sequences</th>
<th>Freq. (n = 602)</th>
<th>Cumul. Freq. (n = 602)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  S2222</td>
<td>22%</td>
<td>22%</td>
</tr>
<tr>
<td>2  S2111</td>
<td>5%</td>
<td>27%</td>
</tr>
<tr>
<td>3  S2221</td>
<td>5%</td>
<td>32%</td>
</tr>
<tr>
<td>4  S1111</td>
<td>5%</td>
<td>37%</td>
</tr>
<tr>
<td>5  S4444</td>
<td>5%</td>
<td>42%</td>
</tr>
<tr>
<td>6  S2211</td>
<td>3%</td>
<td>45%</td>
</tr>
<tr>
<td>7  S2224</td>
<td>3%</td>
<td>48%</td>
</tr>
<tr>
<td>8  S3333</td>
<td>3%</td>
<td>51%</td>
</tr>
<tr>
<td>9  S2223</td>
<td>2%</td>
<td>54%</td>
</tr>
<tr>
<td>10 S2244</td>
<td>2%</td>
<td>56%</td>
</tr>
<tr>
<td>11 S2444</td>
<td>2%</td>
<td>58%</td>
</tr>
<tr>
<td>12 S1311</td>
<td>2%</td>
<td>59%</td>
</tr>
<tr>
<td>13 S2113</td>
<td>1%</td>
<td>61%</td>
</tr>
<tr>
<td>14 S1333</td>
<td>1%</td>
<td>62%</td>
</tr>
<tr>
<td>15 S2242</td>
<td>1%</td>
<td>63%</td>
</tr>
<tr>
<td>16 S4222</td>
<td>1%</td>
<td>65%</td>
</tr>
<tr>
<td>17 S4333</td>
<td>1%</td>
<td>66%</td>
</tr>
<tr>
<td>18 S1131</td>
<td>1%</td>
<td>67%</td>
</tr>
<tr>
<td>19 S1133</td>
<td>1%</td>
<td>68%</td>
</tr>
<tr>
<td>20 S2122</td>
<td>1%</td>
<td>69%</td>
</tr>
<tr>
<td>21 S2233</td>
<td>1%</td>
<td>70%</td>
</tr>
<tr>
<td>22 S3111</td>
<td>1%</td>
<td>71%</td>
</tr>
</tbody>
</table>
Figure 7: Typical trajectories.

```
<table>
<thead>
<tr>
<th></th>
<th>“Resource”</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>“Seeds”</td>
<td>G</td>
<td>“Stars”</td>
</tr>
<tr>
<td>B</td>
<td>E</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>“Questions”</td>
<td></td>
<td>“Boutiques”</td>
</tr>
<tr>
<td>“Performance”</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
### Table. 3: Taxonomy of growth trajectories

<table>
<thead>
<tr>
<th>Description</th>
<th>Corresp. sequences</th>
<th>Freq. (n = 602)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stable trajectories</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Grow as a “seed”</td>
<td>S2222</td>
<td>22%</td>
<td>Firm having raised capital to reach breakeven</td>
</tr>
<tr>
<td></td>
<td>(high resources</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>but low performance)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D Grow as a “star”</td>
<td>S4444</td>
<td>5%</td>
<td>“Gazelle”; potential high-growth firm</td>
</tr>
<tr>
<td></td>
<td>(high resources</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>and high performance)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Grow as a “question”</td>
<td>S1111</td>
<td>5%</td>
<td>Firm experiencing troublesome growth</td>
</tr>
<tr>
<td></td>
<td>(low resources</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>and low performance)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C Grow as a “boutique”</td>
<td>S3333</td>
<td>3%</td>
<td>Profitable service firm with limited assets</td>
</tr>
<tr>
<td></td>
<td>(low resources</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>but high performance)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unstable trajectories</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E Grow from a “seed” to “question”</td>
<td>S2111, S2211, S2221</td>
<td>13%</td>
<td>“Seed” who burns capital before breakeven</td>
</tr>
<tr>
<td>G Grow from a “seed” to “star”</td>
<td>S2444, S2244, S2224</td>
<td>7%</td>
<td>“Seed” having developed towards profitability</td>
</tr>
<tr>
<td>F Grow from a “seed” to “boutique”</td>
<td>S2333, S2233, S2223</td>
<td>4%</td>
<td>“Seed” having lowered ambitions</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>59%</td>
<td></td>
</tr>
</tbody>
</table>