UNIVERSITA’ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI
“M. FANNO”

CORSO DI LAUREA MAGISTRALE IN
BUSINESS ADMINISTRATION

TESI DI LAUREA

“ENTREPRENEURIAL SPINOFFS: THEORETICAL AND EMPIRICAL ISSUES”

RELATORE:

CH.MO PROF. ANDREA FURLAN

LAUREANDA: JESSICA RAGAZZON

MATRICOLA N. 1130337

ANNO ACCADEMICO 2016 – 2017

Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere.
Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell'elaborato sono stati indicati nel testo e nella sezione "Riferimenti bibliografici" e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo alla pubblicazione originale.

Firma dello studente

____

INDEX

FIGURES

TABLES

INTRODUCTION

The creation of new firms has always been deemed an imperative for economic development (Schumpeter, 1934). Nowadays, it has become a hot topic due to a background of uncertainty and volatility and many scholars have tried to delve into the nature of *de novo* firms with the aim to identify which ones fit the environment better. Among these new entrants, *spinoffs* have received greater attention since they show considerably higher survival rates than other new start-ups. What discriminates spinoffs from other entrants is that their founders accumulated industry-specific experience when working for an incumbent firm operating in the same industry of their new ventures. Spinoffs are more likely to survive since they can exploit experience inherited from their parent firms (Klepper, 2001) in terms of market and technical knowledge, relationships with customers and suppliers and organizational routines and blueprints. Furthermore, many studies have exhibited that the bulk of new ventures are spinoffs, irrespective of the country or sector analysed.

The increasing interest on spinoffs has led to a proliferation of studies whose results and evidence have been synthetized, to the best of our knowledge, by two reviews authored by Klepper in 2001 and 2009. While the focus of the former is on theoretical aspects underlying spinoffs formation and performance and posits the analogy of spinoffs as children which inherit parents' knowledge as genes, the second emphasises mostly some empirical cross-border regularities on these new entrants, that are called "nine stylized facts" by the author. Klepper's second review unveils also a tremendous diversity in approaches and themes investigated by scholars up to that moment, among which their performance, their formation, the relationship with their parents and parents' characteristics conducive to the spawning event, the nature of the inherited intellectual capital and the extent to which they contribute to clusters development. Inasmuch interest on spinoffs is increasing, we expect that research *post* Klepper's review is likewise motley and while some analyses may have been conducted on existing strands of research, some others may have been performed on bloomed themes. Our endeavour is thus to identify whether spinoffs literature *post* this latest review can be broken up into different thematic groups, which represent the subfields of the knowledge. Therefore, we try to provide a systematization of the literature after 2009 to determine the state of the art of the discipline and where the literature is going. Researchers are likely to benefit from literature systematization since understanding the underlying processes and outcomes reveals the evolution of thought and provides an indication for the future (Culnan, 1986) and this is demanding in spinoffs literature given its heterogeneity and hectic disclosure.

Subfields detection is performed through quantitative tools such as *citation* and *co-citation analysis* which assume that papers linked by citation ties are more likely to belong to same thematic areas, whereas papers disconnected are less likely to delve into similar topics. Results from these analyses are only the starting point for subsequent procedures whose final output is a map of the knowledge under investigation.

Klepper's 2009 review ends with some research questions and among these the topic of innovation results one of the most compelling and debated by scholars. More specifically, it is still unknown to what extent spinoffs innovate more than other new entrants at the start-up phase and we try to solve this issue through an empirical analysis. Indeed, while some authors argue that spinoffs are inherently innovative new ventures, some others claim that industry-specific experience that their founders have accumulated during their previous working experience deters the innovativeness of the spinoff itself during the start-up phase. Our goal is thereby to determine whether spinoffs are incremental or radical innovators with respect to other new ventures, when they are created.

This thesis is structured as follows. In the first chapter, we provide an overview of the literature on spinoffs, by focussing mainly on the knowledge accumulated up to 2009, that is *pre* Klepper's latest review. Such overview highlights the major trends and studies on spinoffs topic, starting from a spinoff level of analysis, passing through a parent level of analysis and ending with a consideration on clusters and implications for the environment.

In the second chapter, we describe the methodology followed to perform the systematization of the literature on spinoffs *post* the aforementioned review. It encompasses some steps such as the retrieval of the relevant articles on spinoffs, citation and co-citation analysis and some statistical methods such as principal components analysis, cluster analysis and multidimensional scaling to obtain knowledge mapping.

In the third chapter, the results of the procedure followed are shown. Quantitative analyses are complemented by personal interpretation of articles' contents on spinoffs and the main subfields of research are presented, that is the topic investigated by authors on spinoffs literature *post* 2009.

In the fourth chapter, we delve into the relationship between spinoffs and innovation through an empirical analysis. A sample of *de novo* ventures founded by Alumni of University of Padova graduated between 2000 and 2010 is analysed for this aim. While in the first part of the chapter some descriptive statistics are shown, in the second part some logistic regressions will shed light on the degree of innovativeness of spinoffs with respect to other new ventures.

CHAPTER 1: LITERATURE REVIEW ON ENTREPRENEURIAL SPINOFFS

1.1 Introduction

The increasing interest on entrepreneurship has allowed the identification of different typologies of market entrants, among which spinoffs are playing a leading role. The bulk of new established firms are indeed spinoffs and this has drawn the attention of scholars and researchers towards this category of new entrants. In this chapter, we are going through a review of the literature on spinoffs, by emphasising their origin, their characteristics in terms of differences and similarities with other firms and what is still unknown on spinoffs.

1.2 Definition of *spinoffs*

Recent years have witnessed a great interest on the phenomenon of entrepreneurship, both at the scholars' level and at the policy makers' one, due to the unsteady economic and financial environment (Ferreira et al., 2017). New firms are, in fact, among the key drivers to foster economic growth; they are deemed outperformers with respect to older firms, being the latter characterized by organizational inertia (Eriksson and Kuhn, 2006). They are also job creators and enhance competition in the market (Dahl and Reichstein, 2007).

But what is the origin of these new firms entering the market?

It is worth pointing out that most of the entries occur when a market is experiencing an initial and growth stage, albeit some sectors may face following upward trends corresponding to technological development and to new practices which represent the time when new firms usually enter. A potential classification of new entrants may be tripartite and considering the extent to which new firms are created by and/or bound to the parent company: diversifying, parent-company ventures and *de novo* entrants. The first category embeds those firms entering new or consolidated markets that are formed, in general, by internal growth or acquisition (e.g. in the form of foreign subsidiaries). The second category is the result of a setting up process established by parent companies, different from diversification, and including joint ventures, franchises and parent spin-offs; parent spin-offs are deemed a hybrid between diversifying and *de novo* firms, since they are founded by a parent firm and at the same time are a separated legal entity. Among *de novo* entrants, start-ups and spinoffs are recognized and form separated legal entities (Helfat and Lieberman, 2002). The difference between these two latter entrants lies on the fact that while the former ones are created without previous ties with other industry incumbents, spinoffs are firms founded by former employees of incumbent firms that operate in the same industry in which the new firm is created (Agarwal et al., 2004; Klepper, 2001;

Furlan, 2016a); spinoffs' founders are independent of their former employers, albeit in some cases they could be financed and tied (Klepper and Thompson, 2005). Klepper (2009) associates that definition with the narrower concept of *intra-industry spinoffs*, while defining a *spinoff* as a firm created by ex-employees of incumbents, irrespective of the industry. For the sake of simplicity, we will refer to spinoffs henceforth with the meaning given by Klepper to intra-industry spinoffs.

It is also worth to underline that spinoffs topic has gained a stronger foothold recently, especially since when new ventures birth evidence discredited the widely accepted "garage belief". This popular way of thinking associated with entrepreneurship conjures up the image of a dropout young future entrepreneur who fiddles in her parents' garage and comes up with disruptive innovations that form the base for a new business opportunity (Chatterji, 2009; Furlan, 2016b). The "garage belief" is a blend obtained by other entrepreneurship images such as "the inspirational generation of innovative ideas, old-fashioned hard work and American ingenuity, bootstrapping resources to chase a dream, a rejection of the status quo, and the freedom of working for oneself" (Audia and Rider, 2005, p. 6). Although this cliché stems from successful innovative enterprises' founders *à la* Steve Jobs, it is not as common as thought (Audia and Rider, 2005). The empirical evidence has, in fact, shown that new ventures are founded mainly by entrepreneurs who are not wet behind the ears. A recent study conducted by the Department of Economics and Management of the University of Padova traces the typical new entrepreneurs' traits by analysing a sample of 450 firms located in the North-East of Italy in high and medium tech industries. Firms' founders are found to be educated (more than 50% are graduated), not very young (nearly 64% are older than 40 years old and only 6% of them are younger than 30) and to have worked several years in related industries (62.7% of the total) (Bettiol and Furlan, 2014). This entrepreneurial profile thus moves away from the dropout, young and without experience founders' profile commonly evoked by the "garage belief". The fact to be experienced appears to be a cross-border characteristic, as a study conducted by Audia and Rider (2005) on American start-ups exhibits that 91% of them operate in industries related to their founders' experience.

For these reasons, the dynamics underlying spinoffs deserve more explanation and investigation. Spinoff is a nuanced concept, according to the different settings which this term is employed in and thus it is broader than the definition provided before. For example, Ferreira et al. (2017) posit a taxonomy of the term, based on the different contexts in which it can be adopted, that is *corporate*, *academic* or *entrepreneurial*. The common denominator for these three concepts of spinoff is the fact that something new is established from an entity that is already existing (Wallin, 2012). A corporate spinoff is a separate unit, managed independently but created and partially owned by the parent firm and sometimes listed in a stock market which can be used by firms to fulfil corporate stakeholders' needs (Ito, 1995); the research on corporate spinoffs focusses on market performance and shareholder wealth (Ferreira et al., 2017). Spinoffs from public and academic sector are instead defined as "new, small, high technology or knowledge intensive company whose intellectual capital somehow has origins in a university or public research institution" (Callan, 2001, p.15). Lastly, entrepreneurial spinoffs are firms founded by employees of incumbent firms in an industry (Klepper, 2009) and will be the focus of this thesis. We will also rely on the classification provided by Ferreira et al. (2017) henceforth.

There is limited consensus also in establishing when a new venture can be deemed an entrepreneurial spinoff. For instance, Eriksson and Kuhn (2006) -and the same criterion is adopted by Andersson and Klepper (2013)- define a spinoff when at least 50 per cent of the employees in the new firm come from the same firm, even though they should represent less than 50 per cent of the total workforce of the former firm. By using the same Danish matched employer-employee dataset, Dahl and Reichstein (2007) classify a new venture a *spinoff* when at least two members of the management team of the new firm were previously employed in the same venture one year before the foundation, since most new firms in Denmark are more likely to be run by managers who are at the same time founders. Nonetheless, it can be argued that these different measures are not necessarily related to a lack of specificity around the definition of spinoffs, rather they may be ascribed to the incompleteness of the dataset which those authors rely on to conduct their studies, as founders' identification was hard. Research on spinoffs has dealt very often with the definition of spinoffs by considering two or more founders and more paid employees (e.g. Andersson and Klepper, 2013) even though a large part of new firms are actually proprietorships: to fill this gap, Furlan (2016a, p. 424) investigates the dynamics of firms founded by a single individual and defines *spinoff* a new venture "founded by persons who had previously worked as paid employees for a firm operating in the same industry as that of the new venture". We will rely on this definition for our analyses henceforth.

A further distinction is then usually made concerning the type of spinoffs, that is *pulled* or *pushed*: if the parent firm exits the same year in which the spinoff is born, the new venture is deemed pushed, otherwise it is pulled (Andersson and Klepper, 2013). The same distinction is made by Buenstorf (2009), who names *opportunity spinoffs* those firms created by the identification of opportunities to be pursued by former employees who have become entrepreneurs and *necessity spinoffs* those ventures triggered by adverse external factors that made less appealing the employment, as to reduce the opportunity costs of entrepreneurship.

This latter categorization is important not only for taxonomy purposes, but rather to understand empirical patterns associated with different types of spinoffs.

1.3 Where do spinoffs come from?

Since spinoffs play a leading role in economic growth, it is worth discovering the dynamics behind their formation. Klepper and Thompson (2005) summarize three categories of models, distinguishing them in three different camps, to explain the origin of these new entrants. The first category of models stems from the concept of information asymmetries and posits that an employee who made a discovery more valuable for the firm in which she is working at first, decided not to disclose her findings because of information asymmetries. Wiggins (1995) postulates a model in which an employee (the classic example is an R&D employee) is working to develop a discovery, whose successful development is dependent upon employee's efforts and whose success can be assessed by the employer. No payment will be done to the worker until the success of the discovery is actualized and, in any case, the employer may understate the profits, that will be higher the greater the efforts exerted by the employee; when "the likelihood of success is small, it is difficult to observe the labour output over long period and capital/labour ratio is low" and new enterprises are more likely to emerge (Wiggins, 1995, p.65): the payment will be greater the lower the odds of success and the length of time ties the owed payment down. These conditions are typical of path-breaking innovations and new lines of business and are conducive to spinoffs bursting out.

Another model proposed by Anton and Yao (1995), instead, posits that an employer can learn the discovery by her own even though the employee has not disclosed the discovery yet and the employer cannot establish whether a worker has made such discovery. Therefore, the employee can behave opportunistically by pretending to have made such discovery; the employer can contract a payment owed by the employee if the employer cannot learn the discovery and this payment is higher the more innovative (i.e. path-breaking) is deemed the discovery. Thus, spinoffs emerge to produce different types of products than their parents (Klepper, 2009).

The second category of models is based on a discovery event when a new employee is hired to work on a new project which has less value for the existing firm rather than for a new venture, as in the case when a new discovery may harm existing lines of the business or when it is far from the core business (Klepper and Thompson, 2005); thus, a new firm may be founded by the new employee. Even in this case, new ventures may emerge by producing different products than parent firms.

The third category of models refers to the role of employee learning in forming spinoffs (Klepper, 2001). They assume that all the firms produce the same type of products and are distinguished according to the type of knowledge, therefore, new entrants are only spinoffs. Employees accepts lower wages to be hired by more knowledgeable firms, since in the future they can leverage the knowledge of the market acquired as apprentices to start their new businesses. This model, differently from the previous outlined, posits that spinoffs start to produce the same products of their parents (Klepper, 2009).

Altogether, these models do not reflect what usually comes about for spinoffs. The first category of models, for example, assumes that employees do not reveal their discoveries, but this contrasts with the reality, since employers usually know the new ideas and refuse to implement them. The second category relies on the fact that firms are unable to capture the value of new ideas but this contrasts with the fact that the most successful companies, which are those that have better evaluative capacities, spawn more; further, it is postulated that new discoveries can cannibalize existing lines of the firm and this does not explain why, as shown empirically, new ideas have been started and discovered at the parent firm. Lastly, even though the third category of models confirm some of the empirical trends related to parent and spinoffs' performance, it cannot explain why these new ventures do not realize products similar to the parents' ones, but instead are used to implement rejected ideas on variants of parents' products (Klepper and Thompson, 2005).

More recent theories have been proposed to give an explanation on reasons behind spinoffs formation and they seem to confirm empirical trends. They share the common idea that employers are not able to fully grasp the actual value of new ideas or the value of their employees and this causes similar treatment to all the employees, with the result that best performing employees start their new businesses (Klepper, 2009). Klepper and Thompson (2010) theorize that firms are actors composed by many individuals that at first have the same ideas; then they receive different signals on what the firm should do to continue its business and this usually occurs when the parent is in decline. These signals reflect disagreements between the firm and employees with new ideas implying that the larger the disagreement, the more likely the employee will leave the parent to found a new venture. The validity of this model is undeniable: the evidence shows that spinoffs result from disagreements on new ideas and, when these latter are proposed by the most talented employees, spinoffs are outperformers and their parents are superior performers as well. Furthermore, this model contributes to the explanation on the spawn off time: since at the beginning of the parent's life all the employees share a common view, disagreements are near to zero and therefore no spinoff occurs; later, when the parent firm is no longer an outperformer, disagreements arise and then, after dissidents leaving the parent, the rate of spawning declines.

Buenstorf (2009) builds a framework that can be complementary to the one just described on the origin of spinoffs. This model conjectures the existence of external events that trigger new firms' foundation. When these events are adverse and weighing on parents' performance, *necessity* spinoffs arise; on the contrary, when events are related to opportunities identification, *opportunity* spinoffs are spawned. This framework partially breaks away with Klepper and Thompson's one as it acknowledges that disagreements are not the only driver for spinoffs formation.

1.4 Evolutionary economy, population ecology and knowledge inheritance

Employee learning theories have proved to have limited explanatory capacity on some spinoffs' trends: the empirical evidence has in fact shown that spinoffs perform better than other kinds of start-ups (e.g. Agarwal et al., 2004; Delmar and Shane, 2006; Chatterji, 2009) but learning theories can only evaluate what is the starting point of a new venture and therefore cannot predict the performance. This latter, instead, can be explained by other two theories, that is *population ecology* and *evolutionary economics*. Population ecology assumes that organizations are dominated by inertia, both for internal and external constraints that makes hard a dynamic adaptation to the environment (Hannan and Freeman, 1977); for this reason, only those firms that enter the market with knowledge and resources fitter to the environment can succeed, while the others die (Dencker et al., 2009). A different perspective is the one proposed by Nelson and Winter (1982) in the evolutionary economics theory: the core of this theory is the concept of *routine*, defined as "what is the most regular and predictable about business behaviour" (p.15). Routines are therefore identified as organizational characteristics, similar to what genes represent in biological evolutionary theory, since they persist in the entity and by shaping the possible behaviour of the firm (the actual one results also from the interaction with the environment), they are also selectable (as firms endowed with better routines are more likely to succeed in the environment) and inheritable (these routines are passed down by "today" firms to "tomorrow" ones, which will present the same characteristics). Thus, evolutionary theory maintains that a firm's pre-entry resources and knowledge have an impact on its adaptation capabilities and firms that succeed in adapting have higher chances to survive (Dencker et al., 2009). Each line of the business and each function require these routines to be developed and their first installation depends on founders. Founders are likely to adopt rules that shaped their previous experience and therefore when a new venture is created, the new entrepreneur bases its routines on the previous employers' ones; as Nelson and Winter (1982) point out, the memory of new ventures is embedded in the organizational actors and employees gain knowledge on routines simply by belonging to a specific organization (Dahl and Reichstein, 2007).

Klepper (2001) provides a model, stemming from the analogy parent-spinoff with parent-child and from the evolutionary theory to infer the fate of the new firm. Organizations can reproduce when one or more employees transplant organizational routines to the spinoffs and only those routines more appropriate to the environment will provide a competitive edge with respect to other start-ups. Consequently, spinoffs are going to inherit a subset of routines from the parent firm, that are more relevant to compete in the market. This can explain why the lower is the gap between the industry of the new firm and the parent's one, the more likely is the fitness of parent's routines for the offspring. As a matter of fact, spinoffs differ from the other categories of entrants also because they need to capitalize the knowledge they accumulated from their parents instead of looking for the most profitable markets which to compete in (Klepper and Sleeper, 2005). This model laid the foundations for the *knowledge inheritance theory* which posits that a new venture's founder transfers the stock of knowledge accumulated during past working experience to the new firm that resembles the parent one (Agarwal et al., 2004; Klepper and Sleeper, 2005; Furlan, 2016a).

Knowledge inheritance theory works well for intra-industry spinoffs: knowledge transfer is effective when spinoffs enter the same industry of the parent. The founder has industry-specific knowledge -both technology and market related- including blueprints and practices that can help her shape the new organization, at least in the previous phases and, according to Stinchcombe (1965), it is likely to have a long-term influence. Other start-ups, instead, lack this knowledge specificity and are forced either to hire employees with industry skills or to learn by doing. The transfer of knowledge from the founder cannot, however, be compared to the one from a hired employee: the former in fact, promotes a full and more effective knowledge transfer between organizations and, at the same time, is more interested in implementing the best practices leading to profits. While new hired employees may prefer not to lose their non-replicated knowledge, as it represents the source of the power within an organization, founders do not discriminate between their own and their ventures' aims and thereby allow a lavish dissemination of knowledge throughout the new organization (Agarwal et al., 2004). This theoretical background on spinoffs can be also used to rebut the general assumption that *de novo* entrants, and more specifically spinoffs, are endowed with a fewer resources at the initial stage with respect to parent-company ventures and diversifying entrants as defined earlier (Helfat and Lieberman, 2002).

1.5 Knowledge inheritance and effects on firms' performance

It is worth noting that the kind of knowledge inheritance theory refers to focusses not only on technological aspects but also on market awareness: relationships with suppliers, sales techniques and detection of market opportunities are more likely to be successful when they come from consolidated practices and routines that the founder has transferred to her new venture and that have proved to be winning in the past (Dahl and Reichstein, 2007). While individual characteristics of entrepreneurs are relevant on the foundation of their new ventures, experience accumulated at the parent firm on how to acquire financial resources, to gain social capital, to create routines, to identify business opportunities and intellectual property's management are a vital springboard (Chatterji, 2009).

Altogether, this can contribute to overcome the *liability of newness*, a phenomenon that Stinchcombe (1965) associates with every new venture, at the start-up phase. This concept supposes that new firms are expected to die because they lack experience and trust and, if the firm survives during the initial stage, the experience learning curve mechanism is such that the firm can develop organizational routines and practices that allow its survival and growth in the future. Spinoffs do not have to learn any organizational roles, criteria to take decisions, specialized skills and other new elements that when are missing can contribute to liability of newness; at the same time, they can rely on trust of established social relations and therefore have more ties to customers and suppliers than other start-ups (Phillips, 2002). The trust dimension can also be investigated from the access to capital viewpoint: venture capitalists are less likely to be reluctant to fund experienced entrepreneurs rather than to fund entrepreneurs with no previous experience (Chatterji, 2009). Consequently, it may be inferred that spinoffs are more likely to survive than other start-ups. Despite the main focus on high-tech industry of previous studies, the empirical evidence has, indeed, confirmed this conclusion: in semiconductors (Klepper, 2009), disk drives (Agarwal et al., 2004), lasers (Klepper and Sleeper, 2005), medical devices (Chatterji, 2009), fashion (Wenting, 2008) and law industries (Phillips, 2002), Danish private sector (Eriksson and Kuhn, 2006; Dahl and Reichstein, 2007) and Italian manufacturing industry (Furlan, 2016a) spinoffs have higher survival rates with respect to other start-ups, that is have a lower hazard rate. The empirical patterns are consistent with the fact that resources developed after entry impact on the survival of new firms, but success is mainly driven by initial resources and knowledge (Helfat and Lieberman, 2002) which can overcome better the liability of newness as mentioned before.

The evidence has also shown that spinoffs' founders leverage their specific knowledge gained by previous job rather than general experience. When creating their spinoffs, employees from incumbent firms bestow the fit resources and knowledge upon their new firms in a way that other founders do not do; the superior performance of spinoffs may in fact be traced back to this phenomenon (Buenstorf and Fornahl, 2009). In medical device industry, technical knowledge is not the driver of spinoffs' superior performance: nontechnical knowledge from the parent firm, that is regulatory and marketing knowledge and opportunities detection are the determinants of competitive edge (Chatterji, 2009). Spinoffs' outperformance is not limited to the early stages of new ventures, rather it concerns post- entry performance as well: pre-entry experience has revealed to have a long-term influence on firm survival after the entry and this may unveil that initial choices, inherently affected by pre-entry experience, play a main role in technological and market conditions, even though firms can change and adapt to environment's needs (Helfat and Lieberman, 2002).

Pre-entry knowledge and management experience have been investigated also as indirect mechanisms that positively affect spinoffs' survival through learning activities. Learning activities are based on routines and occur when an organization adapts its routines or beliefs because of either direct or "experiential search", that is either by trial-and-error experimentation or by learning from others' practices. Scholars have been interested by the effects of early-stage business planning and post-entry product line change as two learning mechanisms. In the first case, business plans are written before the launch of the new venture and the consequences on performance are not clear, while the second case refers to a change in products or services and as such the result of learning through experience. It was found that early-stage business planning negatively affects spinoffs' survival in that planning activity may hinder the receptiveness of the organization and less able founders expect to offset their inability with this activity; despite this, pre-entry knowledge and experience positively impact on the benefits of early-stage business planning. Product line changes, instead, positively affect the survival and the effect of pre-entry knowledge and experience is positive as well. In sum, pre-entry knowledge plays a critical role in firms that try to fill the gap between their existing resources and the required ones by means of learning activities and it should be considered for survival prospects (Dencker et al., 2009).

Among studies on spinoffs' performance, Delmar and Shanes's one (2006) is noteworthy, as it provides an analysis that encompasses not only the survival dimension, as most of the studies have done, but it also focusses on the sales of the firms as measure of economic performance, confirming that industry experience has a positive effect on spinoffs' performance, albeit with marginal decreasing returns (see paragraph 1.5.4).

Finally, spinoff's status has proved to enhance the quantity of venture capital and to receive higher valuations at the last round of financing with respect to other categories of entrants (Chatterji, 2009). However, the effects of pre-entry knowledge on new firms' survival still represent a “fruitful avenue for future research” and new studies are thereby desirable (Helfat and Lieberman, 2002, p.753)

1.5.1 Effects of parent firms’ performance on spinoffs survival and birth rate

The lower hazard rate of spinoffs deemed as an unconditioned dogma may be misleading. Since spinoff performance is shaped by pre-entry experience of founders, which in turn is affected by parent knowledge, it can be inferred that the quality of the parent firm shapes the quality of the spinoff and therefore its performance. In other words, the quality of the organizational experience affects new venture’s success and performance (Dahl and Reichstein, 2007). An unsuccessful parent may, in fact, transfer routines and knowledge that are not appropriate to compete in a specific market by means of the mobility of employees (Phillips, 2002) and therefore the failure of the spinoff relying on such routines is a self-fulfilling prophecy. As Klepper (2001) claims, more successful parents will generate more successful spinoffs because these latter rely on better routines inherited from their parents. Later, Klepper (2002) finds that US automobile new producers are more able to compete when founders worked several years in a leading company of the sector.

Several studies have dealt with the issue of an unhealthy parent leading to a lower spinoff performance and have used different methods to identify this condition. The different methods adopted, however, share the common premise that a parent firm near to failure is run by organizational routines that no longer fit the environment and therefore the performance of the parent firm can be deemed a proxy for the quality of knowledge affecting new venture’s survival. Wenting (2008) for example, approximates the number of years in which the parent company has survived until the time of spawning to the quality of the firm itself and has proved that more successful spinoffs are likely to be generated by successful parents (in terms of survival years). Survival as proxy for firm’s performance is also used by Buenstorf and Klepper (2009), who prove that leading firms in Akron tire industry are characterized by higher survival rates and by a higher spawning rate; the progeny is also outperformer in the region. Eriksson and Kuhn (2006) distinguish the performance of the different types of spinoffs according to the occurrence of a “push” factor: a spinoff is deemed *pushed* if the same year in which the spinoff is founded the parent stops its operations and activities and therefore the new entrepreneur is “pushed” to create a new venture to continue working. The authors demonstrate that this type of spinoff has higher hazard probabilities than spinoffs whose parents continue their activities and has not lower hazard rates than other start-ups. Dahl and Reichstein (2007) reach the same conclusion by categorizing parent firms that survive after the founding year of the progeny and parents that die after spinoffs foundation; the result is that spinoffs from surviving parents have the highest survival rates, spinoffs from exiting parents are the worst performers and the group of start-ups, founded by inexperienced entrepreneurs, are middle-performers among the two groups. Spinoffs from German automobile producers showed higher survival probability than non-spinoffs, with the longevity of the former positively related to the number of years their parents were active (Klepper, 2009). Finally, Phillips (2002) does not find a statistically significant relationship between firm age and firm chances of survival, albeit he demonstrates that new entrepreneurs leaving parents near to failure are less likely to build successful organizations. Although some dissident opinions claim the importance of necessity spinoffs (read: pushed) as critical players in the market, the superior performance of opportunity spinoffs (read: pulled) is acknowledged (Buenstorf, 2009). Consequently, we may state that it is not the existence of pre-entry knowledge *per se* that determines the fate of a new venture, but the type of knowledge as result from the quality of the parent firm.

It has been also argued that inheritance as mechanism to explain the superior performance of spinoffs can be a red herring. Another interpretation of the relationship between a positive performance of the parent and a positive performance of the spinoff concerns with the fact that more successful parents are magnets for better employees that, when leaving, can rely on their talent to build successful firms (Wenting, 2008). Thus, the previous working experience should be ruled out, advocating the superiority of the parent screening process as key for the success of spinoffs. Even though this opinion cannot be wholly neglected, the evidence points out to a stronger inheritance process than a screening one (Chatterji, 2009). As a matter of fact, when employees are hired, they learn practices and routines simply by being part of an organization (Nelson and Winter, 1982) and thereby the learning component is a driver to explain the success of new firms. Furthermore, the evidence shows that employees who found new ventures in the same industry of their parent perform better than those entering different ones (Chatterji, 2009) and this may be a signal of the dominance of inheritance and learning mechanisms over talents' selection. Notwithstanding, the complementarity of inheritance and screening processes should not be ruled out and for this reason further investigation is required.

The quality of knowledge of the parent firm, has provided useful insights also concerning the rate at which these firms spawn off. The concept of inheritance can explain this phenomenon as well: members which fit the industry more have the highest reproduction rates (Klepper and Sleeper, 2005) and the evidence confirms this pattern. In automobiles, lasers, disk drives, tires and semiconductors industries, outperformer parents are characterized by a higher spawning rate and this is consistent with the idea that better parents are the best training place for employees willing to start their new businesses (Klepper, [2001, 2002, 2009]).

The evidence also shows that firms with superior technology and that are early entrants in the industry spawn more spinoffs. The larger number of offspring by leading firms emphasizes that the superior environment of these latter is the ideal condition to make employees learning about practices and routines to found their new ventures (Buenstorf and Klepper, 2009).

Notwithstanding, spinoffs are not idiosyncratic to successful firms; declining parents are also likely to spawn off as mentioned before. The slower economic performance and some organizational changes are indicators of a situation of crisis that are likely to be followed by higher spawning rate (Klepper, 2001). Among organizational changes, the appointment of a new CEO and firm's acquisition by other firms -either belonging to the same industry or not- increase the likelihood of spinoffs (Klepper, 2009). Eriksson and Kuhn (2006), by considering the shift of the CEO to be a signal of internal chaos, find that a recent appointment has a strong probability to be followed by employees' departure to found new firms. Concerning firm's acquisition, instead, the new management may not grasp discoveries and opportunities previously identified by R&D (Klepper and Sleeper, 2005) and this may cause employees' leaving. This is consistent with the theory of disagreements postulated by Klepper and Thompson (2010), which implies that organizational changes can create disagreements and these latter give rise to new ventures. The model confirms also some empirical regularities such as the time of spawning: more successful companies spawn more and the time of spawning usually occurs when the parent is declining and during its "middle age". The rate at which parents spawn off in many cases is U-shaped, indeed: it rises until a certain point -corresponding to the point in which the parent is 14-15 years- and then it declines. By analysing the longer-lived parents in laser industry, Klepper and Sleeper (2005) find that the highest spinoff rate occurs between the age 11 and 15 and the same pattern was observed in law firms, autos and semiconductors. This may be ascribed to the amount of knowledge provided by the parent firm: the increasing experience and knowledge accumulated by the firm provides fertile ground for new business opportunities and an increasing spinoff rate, then, little by little, this knowledge becomes embedded in physical capital and thereby the transfer of knowledge is hindered. Klepper (2001) explains the declining pattern by assuming that if employees have not left the parent until the middle age, they are very likely to remain.

The quantity of knowledge is crucial in explaining the spawning rate insofar it is judged jointly with other characteristics. Agarwal et al. (2004), for example, find that the abundant knowledge developed at the parent firm enables opportunity recognition by employees, thereby increasing the spinoff generation even though this relationship is not straightforward. The exploitation of business opportunities is dependent on the degree of utilization of such knowledge: when the firm is more prone to focus on either the technological know-how -new discoveries and technological breakout- or market pioneering know-how -the commercialization of such discoveries before competitors-, opportunities are identified but not exploited. This has a two-fold consequence: on one hand, missed exploitations are very likely to create frustration among employees and a gap between their aspirations inside the firm and the current strategy emerges; on the other hand, employees mature confidence on founding their new ventures as business opportunities exist. When instead the incumbent enters a new market, thereby developing both the technological and the market related sides, employees are dissuaded from the creation of their new enterprises. Thus, the number of potential spinoffs is directly proportional to the knowledge of the parent firm, but the actual number of spinoffs is higher when the firm does not exploit opportunities (Buenstorf, 2009). Whilst it is difficult to trace both technological and market inheritance during spinoffs' formation (Chatterji, 2009), it may be stated that when the value creation of the firm is not matched with value appropriation, spinoffs are very likely to be founded.

1.5.2 Effects of spinoffs on parent firms

A spinoff event implies the mobility of employees from a parent firm to a new venture. Considering what has been said before, whenever this event happens, a transfer of skills and routines occurs. By studying law firms in Silicon Valley, Phillips (2002) claims the importance of the parent-progeny transfer. In his study, he concludes that the consequences of the transfer should not be limited to a mere spinoff's performance viewpoint, but should be assessed also from the side of the parent firm, that becomes deprived of some human and social capital, skills, resources and ties to both customers and suppliers, that are vital for its survival. As a matter of fact, the transfer of resources and routines provides both benefits to the offspring with respect to other new ventures that lack a parent firm -and therefore lacking some industry-specific experience- and drawbacks in terms of performance and survival for the parent. What is more, new founders who cover higher-ranked positions at the parent firm have relationships with external resources and are crucial for internal knowledge and social capital; thus, their departure is likely to harm more the firm than lower-ranked employees' leaving. Indeed, the quantity of routines and practices which the entrepreneur can rely on are an increasing function of her perceived importance inside the organization. Two trends are thereby observed: the first one is that when new ventures are founded, higher-ranked employees increase the likelihood of survival of their new firms because of the huge amount of resources and skills transferred from the parent firm, whereas the second trend is that the higher the previous rank of founders, the higher the likelihood of failure of the parent. The probability of failure is also higher the more the offspring is similar to the parent: new ventures occupying the same niches of their parents are likely to be a greater threat than offspring competing in other markets. It has been finally observed that the effects of the spinoff event ebbs over time, as the parent rebuilds its internal structures and routines which has been deprived of; when this process is concluded, the parent becomes competitive again. This latter trend was studied also in disk drive industry: as soon as the parenting event appeared, parents exhibited a decline in their technological knowledge, followed by an improvement with respect to technological frontier the greater the better the technological endowments of spinoffs (Klepper, 2009). Even though the overall effect on parent firm which has experienced a spawning event is positive, parent technological performance differs according to the time in which such event occurs and to the technological state of spinoff. Over time, parents with successful spinoffs outperform parents with no spinoff and this effect is magnified for increasing values of spinoff performance. This may be attributed to a further adaptation to the environment after spawning, which has dampened the organizational inertia of the parent; furthermore, because of inheritance dynamics, outperformer spinoffs signal to the labour market the reputation of their parents as fitter incubators for entrepreneurship (McKendrick et al., 2009).

1.5.3 Parents' characteristics conducive to spinoffs' process

The importance of the parent firm has been claimed by scholars who have tried to investigate how different characteristics of former workplace influence spinoff process. Among these features, the size of the parent firm has always been a hot topic. By studying disk drive industry, Franco and Filson (2000) demonstrate that the size of the parent firm is irrelevant to explain the spinoff formation. Agarwal et al. (2004) show that size is positively related to spinoffs' birth. Eriksson and Kuhn (2006), instead, emphasize that the size of potential incubators is negatively associated with the probability of spawning and Andersson and Klepper (2013) reach the same conclusion when they consider MNEs. In automobiles and semiconductor industries larger firms spawn less spinoffs per employee (Klepper, 2009). These latter results may be coherent with the fact that bigger firms try to fulfil employees' career aspirations and therefore are less likely to attend their leaving.

The effect of size on spinoff's performance has been debated as well. Andersson and Klepper (2013) find that the larger the size of the parent, the higher the probability of spinoff to be outperformer. Phillips (2002) obtains the same results, albeit less statistically significant; this relationship is, in fact, less emphasized when the transfer from the parent firm is higher, that is when the previous rank of employees is higher. This finds support in the theory: larger firms are not necessarily good routines providers as they rely on bureaucratic and complex structures that do not address the malleability and flexibility needs of smaller and younger firms.

Nonetheless, the size may be helpful for lower-ranked employees, who know relatively little of the market and can thereby benefit from the status and influence associated with a larger parent firm. This is the reason why higher-ranked employees who absorb and transfer parent routines may benefit less from parent's status than lower-ranked workers.

Parents that are early entrants in an industry are more likely to spawn better performer spinoffs (Klepper, 2009); this can be explained by positing that early entrants are superior performers than other firms in the industry (Klepper and Simons, 2005) and parents' performance predict somehow spinoffs performance. Franco and Filson (2000) show that early entrants are more likely to spawn early spinoffs entrants and this contributes to explain the survival rates associated with these kinds of spinoffs. Spinoffs are also more likely in young and growing industries that rely less on capital intensity (Klepper, 2001) and this confirms that older firms (older than 15 years) are more capital intensive and therefore spawn less, as mentioned before. Finally, the breadth of parent's business has been investigated: what emerges from the studies of Klepper and Sleeper (2005) is that firms with broader product lines have more spinoffs and each product location represents a distinct source of spinoff; this is in line with the inheritance theory as spinoffs tend to employ the knowledge developed at the parent. Agarwal et al. (2004), instead, find that the presence of the parent in many segments deters the formation of new ventures, perhaps because opportunities have been identified and exploited at the same time by the parent itself and no edge is left to spinoffs.

1.5.4 Effects of founding team experience on spinoffs' performance

The lower hazard rate of spinoffs has been proved by aforementioned studies. However, authors most of the times neglected the role of the amount of industry-specific experience on the performance. Thus, it stands to reason to understand whether the experience from multiple individuals that compose the founding team is desirable and whether the years of working experience in the same industry foster economic growth of the venture.

For what concerns the first question, the empirical evidence has shown that spinoffs with multiple founders perform better than those founded by one entrepreneur but this fact should not be ascribed to a more specific technical knowledge, rather to a marketing and administrative one. Spinoffs with more than one founder address more marketing and administrative issues with respect to single proprietorships and this has revealed to be a winning strategy; multiple founders do not have more technical skills, rather greater skills related to marketing and administration. Founders show to leverage more their experiences constrained by their positions and organizational roles than technological specialties of their parents. Therefore, multiple experiences are desirable for spinoffs performance (Klepper, 2001). Phillips (2002), instead, finds that multiple parents may hinder new venture's survival because of the costs associated with conflicts, negotiation and organization of different ventures' models that are idiosyncratic to each parent; progeny with a single founder is, thereby, expected to have lower hazard rates. These results sharply contrast with the theory that multiple founders bring different resources and routines that can be useful when running a business. Eriksson and Kuhn (2006) seem to reconcile these two findings, by demonstrating that when many employees from the same workplace found a new venture, this new firm has a lower hazard rate. Thus, the importance of the number of workers to predict spinoff's performance should be assessed by considering jointly the amount and the kind of the knowledge transferred. It may be also stated that whenever multiple founders leave the same parent firm to become entrepreneurs, the transfer of resources and skills is such that the parent is harmed more than in the case in which only a single employee leaves it (Phillips, 2002). As mentioned before, the more fit routines are transferred -and therefore more high-ranked employees leave-, the better the performance of the new venture.

The second topic of interest regards how the amount of experience of the founder, usually measured by working years, affects the fate of the new venture. As a matter of fact, studies use a dummy variable to identify the spinoff status, which assumes value 1 when the entrepreneur had previous working experience in the same sector and 0 if not. The amount of this experience has been, therefore, disregarded and it is unclear whether the dummy variable provides a satisfactory explanation for the impact of experience. Some studies have tried to delve into this topic, beyond the assumption that the more employees gain experience, the more they transfer routines and therefore they found outperformer businesses (Klepper, 2001).

Stemming from the conjecture that experience of founding team improves the performance of new ventures, Delmar and Shane (2006) study two components of experience, that is *industry experience* as employees and *start-up experience*. Industry experience, defined as previous work in the industry, is expected to drive performance as it provides new entrepreneurs with industry rules, customers and suppliers relationships and work practices; likewise, start-up experience, defined as earlier firms' foundation, bestows firm organizing, opportunity detection and acquisition of resources upon new founders who are more likely to be successful entrepreneurs. The authors find that both experience components enhance a longer survival and an increase in sales of the new ventures, albeit in different ways. Experienced entrepreneurs are more likely to found firms with higher survival rates, but this positive effect is largely due to the presence of such experience rather than the amount of that; sales, on the other hand, are higher when entrepreneurs have founded at least 4 firms than firms founded by no-experienced founders. The consequence is that a little experience provides a buffer that allows the survival of the firm but not enough to generate positive income. This study demonstrates that the amount of industry-specific experience has positive effects on firm's performance, but these effects have decreasing marginal returns. The same conclusion is reached by Furlan (2016a) who hypothesizes that previous working experience can certainly provide benefits for spinoff's survival but too much experience may be translated into a replication of the business and therefore may not foster its growth. Moreover, Dahl and Reichstein (2007) find no support to the thesis that a high degree of industry-specific experience can influence the probability of survival. Industry specificity has revealed to be a key determinant of spinoffs' survival in the short term, but after four years the gap between spinoffs and other start-ups becomes slimmer (Eriksson and Kuhn, 2006). Therefore, industry experience can affect the likelihood of survival of spinoffs in the very first years but the relationship between the amount of experience and spinoffs' performance is not linear.

The length of tenure was delved into also for predicting the rate at which new firms are spawned. Specifically, there is an inverse relationship between the working time at the former employer and the likelihood to start a new venture. Furthermore, for founding a new venture alone general skills are required; this suggests that the number of previous jobs augments the probability to found a spinoff. On the other side, when several employees are hired, the new organization is endowed with different skills and the founder is very likely to have changed a fewer jobs (Eriksson and Kuhn, 2006). This concept emphasizes the importance of knowledge inheritance mechanism as driver for spinoff generation and the necessity to bestow different capabilities upon new firms for achieving a superior performance.

1.5.5 The choice of the market where to enter

The choice of the market where new firms enter is shaped according to their different nature (i.e. parent-company ventures, spinoffs, start-ups or diversifying entrants), as reported earlier. For example, diversifying entrants are more likely to enter geographic locations or market niches in which the gap between the existing firm resources and the ones required by the industry is not so wide. Whenever this gap is large enough, the choice is to enter different markets *via* parent-company ventures, which have less ties to the parent firm than diversifying entrants. When the focus is on the experience of spinoffs' founders, instead, *de novo* ventures are analysed. Both start-ups and spinoffs share the fact that pre-entry knowledge influences the choice of the market. While in start-ups entrepreneurs enter different markets than the ones which they were previously employed in, they can leverage the knowledge accumulated on suppliers and customers as valuable resource. Entrepreneurial spinoffs start their business endowed with a pre-entry knowledge that is quite similar to the one of the parent (Helfat and

Lieberman, 2002) and therefore they enter industries similar to the ones of the parent. In laser industry, Klepper and Sleeper (2005) find that almost all the spinoffs entered markets closely related to the ones served by parents; while at the beginning spinoffs entered narrower markets where the parents operated, then they started to produce in different albeit related markets, by engaging in activities different from their parents. In semiconductors, the initial production of spinoffs was a subset of their parents' one as well (Klepper, 2009). The importance of pre-entry experience was revealed also in disk drives industry, in which some spinoffs entered new submarkets before their parents; whilst this is consistent with theories on exploitation of new discoveries by employees, it was also observed that innovative efforts usually occur at the parent firm (Klepper, 2001; Helfat and Lieberman, 2002). In this respect, authors debate the innovative nature of spinoffs' entry with respect to other new firms and opinions are usually divergent (e.g. Agarwal et al., 2004; Klepper, 2009).

By studying spinoffs of Intershop in Jena in e-commerce software industry, it emerged that the focus of these new firms was on a specific kind of e-commerce software and their business model was different from the one of their parent, albeit founders admitted having relied on the specific knowledge accumulated during on-the-job-learning. Spinoffs' founders, in fact, exploited more the specific knowledge accumulated during their working experience rather than a more general one. Depending on the role inside the former organization, new founders were more likely to establish ventures whose focus was on issues learnt during their working experience. For instance, one of Intershop's spinoffs that was founded jointly by the former head of quality and the legal advisor entered software quality control segment. On the contrary, workers involved into the strategic decision-making process have shown a higher likelihood to develop new firms closely related to parent's business model (Buenstorf and Fornahl, 2009). This may represent a further proof of knowledge inheritance importance on spinoffs' formation and market choice.

It is notable also that in industries characterized by high heterogeneity of submarkets -as in the case of lasers and software- and given the fact that specific on-the-job learning is a driver of firms' diversification, spinoffs are boosted by the specialization on a specific niche or submarket and therefore are not deemed a threat by the parents.

Furthermore, the entry of spinoffs has been investigated also from an environmental point of view: in laser industry, spinoffs are more responsive to adverse conditions than favourable ones, that is they are less likely to enter when hostile environmental conditions occur and at the same time more positive conditions do not foster spinoffs formation (Klepper and Sleeper, 2005). In sum, unfavourable market conditions have higher effects than favourable ones (Klepper, 2001).

1.5.6 Spinoffs and clusters' formation

Spinoffs have also been studied in the context of industry clustering, which has proved to be one of the hottest topics concerning spinoffs. The main issues are in fact related to their location in clusters and industrial districts -groups of highly related industries that operate in a specific region (Delgado et al., 2014)- and the identification of the determinants of the performance in these areas. Clustering in populated areas is attractive for new entrepreneurs who can rely on social ties, knowledge and confidence in order to accumulate resources for new firms (Sorenson and Audia, 2000).

Detroit, Silicon Valley and Akron areas have been studied because of the concentration, respectively, of automobile, semiconductor and tire producers, that gave rise to an enormous proliferation of new ventures. The evidence shows that firms that are incumbents in these clusters spawn more; in Silicon Valley and Massachusetts, publicly traded firms spawn more companies funded by venture capitalists; moreover, in industrial districts there is a disproportionate entry of spinoffs and the market share of these latter plus the one of their parents after the parenting event is larger than the pre-existing share of the parent, implying that spinoffs are not a zero-sum phenomenon (Klepper, 2009). However, scholars tried to investigate whether the concentration in some areas occurred because of the positive externalities created by an industrial district, the so-called *Marshallian externalities* or because of other reasons. Marshallian externalities consist of a trinity of benefits which firms in a specific area can be influenced by, namely a local pool of skilled labour, local suppliers' linkages and local knowledge spillovers. Each of these drivers is related to cost and productivity benefits with increasing returns the closer the geographic proximity; besides, transaction costs are reduced and specialized institutions at the local level can very often arise (Delgado et al., 2014). The empirical evidence, however, exhibits decreasing returns for these externalities, especially during the later stages of the life cycle of the industry, confirming what Marshall identified as a limitation on the agglomeration that relies on only one industry for its economic development (Potter and Watts, 2011).

Spinoffs are deemed important players for the formation and growth of clusters (Klepper, 2009) and two examples are offered by spinoffs of Fairchild Semiconductor that fostered the growth of Silicon Valley (Klepper, 2002) and spinoffs of tire leading firms in Akron (Buenstorf and Klepper, 2009). Concerning the former, Klepper (2009) emphasizes that Silicon Valley development was fostered by leading semiconductor spinoffs of leading firms. Semiconductor is in fact an industry in which firms must make difficult choices on which technologies to develop and spinoffs emerged mainly between 1957 and 1986, when technical and market uncertainties caused disagreements both at the strategic and at the management level of incumbent firms; when parents were not persuaded by undertaking some ideas, spinoffs founders exploited these opportunities and this reconciles with the theory of disagreements mentioned earlier (Klepper, 2009).

The second case is related to the cluster formation in Akron, a city in Ohio which has become famous for the agglomeration of tire producers during the last 40 years. According to the conventional view, Akron capitalists attracted the firm Goodrich to move its rubber plant in Akron to escape from the Eastern competition and to exploit the proximity to automobile producers in Ohio and Michigan. The development of Detroit area fostered the agglomeration of tire suppliers in Akron; the geographic advantages jointly with agglomeration economies therefore would explain firms' clustering in that area. Nevertheless, the evidence hardly reflects this conventional view: first, new entrants were less than the ones that should be expected if agglomeration economies provided huge benefits; second, entrants mainly came from that region and had ties with local firms (they were spinoffs), with a very few firms coming from other regions; third, evidence of agglomeration economies lacked. Moreover, entrants from Akron were outperformers than entrants coming from different places and this emphasizes the role of inheritance theory to explain the outperformance of spinoffs than other entrants; as a matter of fact, if agglomeration economies had existed, all the entrants would have benefited from them, without pointing out differences in performance. Several spinoffs were founded by employees of leading parents which tended to locate near the latter. Their performance and rate of formation implied that Akron region was shaped first by capabilities inherited by firms rather than agglomeration economies. Thus, the development of the cluster is based on the establishment of some early leading entrants which provided a knowledge background for successful spinoffs that tended to be founded near them; this self-reinforcing process created a breeding region for outperformer firms in the industry (Buenstorf and Klepper, 2009). The same conclusion is reached by Wenting (2008) by investigating the fashion industry: as spinoffs locate near their parents, the concentration of the industry in the region is led by the creation of fitter routines by means of spinoffs. The reason why spinoffs tend to locate near their parents is a key component in cluster formation and may be rooted in the geography of social structure: as people tend to establish geographic networks of contacts, the established ties are such to constrain individuals to move limitedly, giving rise to a geographic inertia. Since the costs for relocation and new ties establishment is high, people are more likely to stay in a fixed place and, when they become entrepreneurs, their new firms are expected to arise there. Furthermore, given that employees tend to work near their workplace and given the geographic inertia, spinoffs' founders are likely to found firms near their parents and this would explain why some geographic areas are populated by many firms and, specifically, by many spinoffs. As a matter

of fact, entrepreneurs coming from different locations may move to crowded areas because of legitimacy purposes or expected benefits from those areas; nonetheless, the empirical evidence has found that a transfer from current location is quite rare (Sorenson and Audia, 2000).
One exception to this pattern occurs when parent firm's attitude towards its spinoff is hostile: such behaviour hinders cluster formation (Furlan and Grandinetti, 2016) that is an engine of regional growth. New employees relocating in the same place of their parents can benefit from information shared with previous employers and colleagues and these conditions spur new ventures to breed within cluster boundaries. This reconciles with the findings of Buenstorf and Klepper (2009): the cluster is mainly driven by spinoffs formation and performance. By studying, instead, the automobile industry in Detroit cluster, Klepper (2007) finds that leading spinoffs coming from leading companies are concentrated in that area and enter there more than other new businesses; firms in Detroit area are characterized by a higher performance of spinoffs and this is linked to spinoffs' process rather than agglomeration economies. Other categories of new entrants exhibit indeed comparable hazard rates between Detroit area and elsewhere locations and this may rule out the hypothesis of agglomeration economies as driver of region's growth. Despite these results, spinoffs in Detroit area show lower hazard rates than other spinoffs located elsewhere and this evidence may imply that spinoffs' process is intertwined with agglomeration economies. The same conclusions are reached by Buenstorf and Fornahl (2009), by analysing the spinoffs of Intershop Communication AG in the software industry in Jena. Intershop started to spawn when it did not pursue some opportunities, that in turn were the base for spinoffs formation. These conditions, together with expected growth trends allowed the incredible proliferation of new firms. This cluster development was also supported by the network of spinoffs based on ties, interactions and knowledge spillovers that frequently occurred. These interactions seemed to mirror the conventional agglomeration economies and their benefits. Founders, in fact, could count on the same background in Intershop and on the familiarity among them which represented a source of regional social network. Still in this case, spinoffs were the drivers for the growth and the performance of the region, but this process was reinforced by the effects of agglomeration.
While more crowded areas can be affected by more fierce competition, many scholars posit the existence of a mix of competition and cooperation among them (Sorenson and Audia, 2000). This may be coherent with the results provided by Klepper (2007) and Buenstorf and Fornahl (2009), implying that the explanation behind clusters' development should not be *aut* spinoffs *aut* agglomeration economies related. As Klepper (2009) hints, further knowledge on cluster formation dynamics should be gained. Agglomeration economies may have been overstated and inheritance theory may not be enough to explain some empirical trends. Rather, further

studies on the interactions between the two phenomena may shed light on the topic and thereby be reconciled with previous analyses.

1.6 Implications for policy makers and non-compete covenants

The outlined evidence and theories pinpoint putative positive effects from spinoffs formation. Given the fact that spinoffs survive longer than other entrants and are willing to undertake ideas and opportunities that their parents only lope along with, the benefits for the entire society are undeniable. Most of the scholars acknowledge the crucial role of spinoffs to technological development and to the growth of regions and society. Nevertheless, some dissidents argue that parents may be harmed from this process and thereby protection measures should be devised and enforced. The supporters of this view claim that parents who fear employees parting from the workplace are less likely to engage in R&D activities, as new founders can exploit opportunities previously identified by parents. Non-compete covenants and trade secrets are provided to deter spinoffs' formation in some states, being respectively a ban on either working for a competing firm for a period or founding a competing firm and a ban on intellectual property disclosure or exploitation in the future (Klepper, 2009). As mentioned earlier, some scholars tried to disentangle this issue and demonstrated that, whilst at the beginning the parent firm is undermined by spinoffs formation, later it catches up (e.g. Phillips, 2002). Spinoffs formation is thus an impetus for parent firms to improve and develop those profitable projects that are reluctant to carry out. In Silicon Valley and California, where non-compete covenants are not allowed, spinoffs have indeed brought about social benefits; if these places are representative, such agreements should not be thereby wished for (Klepper, 2009). The benefit for the whole economy is likely to be higher if firms are not allowed to restrict potential spinoffs competition by means of those agreements (Dahl and Reichstein, 2007).

Since spinoffs are expected to be successful entrants given the accumulated experience of their founders, it is necessary to direct the attention of the managers towards the importance of pre-entry resources. Learning post-entry and the capability to adapt to the environment are constrained by the quality of pre-entry knowledge and the success of the new venture is clearly dependent upon the match between required resources and actual ones. These represent a key determinant for the fate of the new firm (Helfat and Lieberman, 2002). Likewise, studies on industry-specific experience have revealed that an increase in this variable may not be followed by an increase in spinoffs performance. Thus, policies for spinoffs support should take a closer look to the context of the founders and the one of the parent firm to be more effective (Furlan, 2016a).

Finally, as cluster formation is expected to be dependent on spinoffs formation, provided that a neutral or lenient attitude is shown by the parent, policies should dampen an eventual parent hostile behaviour (Furlan and Grandinetti, 2016).

1.7 Spinoffs in a nutshell: stylized facts and Klepper's 2009 review

Spinoffs are an intriguing phenomenon characterized by a spread interest among scholars but also by unresolved issues. The process behind spinoffs creation is perfectly summarized by Moore and Davis (2004), who describe their experience in Silicon Valley pinpointing that spinoffs are not good at creating things, rather they are more able to exploit them. In fact, what they call the "Silicon Valley effect" consists in the fact that every new idea is an opportunity for a new firm. As mentioned earlier, this is consistent with the empirical evidence in most of the industries and with the theoretical models briefly outlined.

Table 1.1 summarizes some of the topics on spinoffs discussed earlier and the contribution of authors considered: effects on spinoffs are examined by considering parent characteristics and founding team; then, effects on the parent firm given spinoff characteristics are reported; finally, the effects on clustering given parent features are synthetized.

Table 1. 1: Synthesis of the effects related to spinoffs process. Source: personal elaboration

EFFECTS ON SPINOFFS			
Level	**Variable**	**Spinoff survival**	**Spinoff rate**
Parent level	Performance (higher performance)	Increase (Buenstorf, 2009; Klepper [2001, 2002, 2007, 2009]; Eriksson and Kuhn, 2006; Buenstorf and Klepper, 2009; Dahl and Reichstein, 2007; Furlan, 2016a; Wenting, 2008)	Increase (Klepper [2001, 2002, 2007, 2009]; Klepper and Sleeper, 2005; Buenstorf and Klepper, 2009) Increase when organizational changes (Eriksson and Kuhn, 2006; Klepper and Sleeper, 2005)
	Age (older parent)	Increase (Wenting, 2008; Eriksson and Kuhn, 2006; Dahl and Reichstein, 2007) Statistically insignificant (Phillips 2002)	Increase (until 15 years), then declines (Klepper, 2001; Klepper and Sleeper, 2005)
	Size (bigger parent)	Increase (Andersson and Klepper, 2013) Increase when lower- ranked employees leave (Phillips, 2002)	Decrease (Eriksson and Kuhn, 2006; Andersson and Klepper, 2013; Klepper, 2009) Increase (Agarwal et al., 2004) Insignificant (Franco and Filson, 2000)
	Knowledge accumulated	Increase when it is not only technical (Chatterji, 2009; Agarwal et al., 2004)	Increase when parents do not exploit opportunities (Agarwal et al., 2004)
	Early entrant	Increase (Franco and Filson, 2000; Klepper and Simons, 2005; Klepper, 2009)	Increase (Franco and Filson, 2000; Klepper and Simons, 2005; Klepper, 2009)

EFFECTS ON SPINOFFS			
Level	**Variable**	**Spinoff survival**	**Spinoff rate**
	Young and growing industries	-	Increase (Klepper, 2001)
	Broader product lines	-	Increase (Klepper and Sleeper, 2005) Decrease (Agarwal et al., 2004)
Founders level	Multiple founders	Increase (Klepper, 2001) Decrease (Phillips, 2002) Increase when coming from the same parent (Eriksson and Kuhn, 2006)	-
	Amount of experience	Increase (Klepper, 2001; Dahl and Reichstein, 2006) U-shaped trend: increase followed by a decrease (Delmar and Shane, 2006; Furlan, 2016) No significant relationship (Dahl and Reichstein, 2007)	Increase when single founder (Eriksson and Kuhn, 2006)

EFFECTS ON PARENT FIRM		
Level	**Variable**	**Parent firm performance**
Spinoff level	Performance (higher performance)	U-shaped trend: decrease followed by an increase (Phillips, 2002; Klepper, 2009; McKendrick et al., 2009, Wade and Jaffee, 2009)
	Higher rank of employees	Higher probability of failure (Phillips, 2002)
	Similar scope	Decrease when spinoff occupies the same market niche (Phillips, 2002)

EFFECTS ON CLUSTER DEVELOPMENT		
Level	**Variable**	**Cluster growth**
Parent level	Performance (higher performance)	Increase due to spinoffs formation (Buenstorf and Klepper, 2009; Klepper, 2007; Buenstorf and Fornahl, 2009)
	Hostile attitude	Decrease due to less spinoffs (Furlan and Grandinetti, 2016)

One of the biggest contribution to the topic of entrepreneurial spinoffs is Klepper's 2009 review, which wraps up the most relevant theories and evidence on this theme up to 2009.

This review identifies nine *stylized facts*, that is patterns with a general validity in almost if not all studies on some relevant aspects of spinoffs, which have been already reported earlier and that can be deemed a synthesis on spinoffs evidence in industries analysed:

1. Better performing firms have higher spinoff rates. Larger firms spawn less start-ups per employee

2. Acquisition by industry and non-industry incumbents are likely to generate spinoffs around the time of acquisition. Spinoffs are more likely when the CEO has recently changed
3. The rate at which new entrepreneurs venture out increases with the age of the firm until the latter is 15, then it declines
4. Spinoffs perform better than other entrants
5. The better the performance of the parent, the better the performance of the spinoff
6. Spinoffs tend to commercialize products that are a subset of the parent's ones
7. After spawning, parents experience an increase in the hazard rate of exit, followed by an increase in survival rate
8. Firms located in a cluster have a higher spawning rate
9. Spinoffs enter clustered areas and their market share lumped together with the parent's one is higher than the previous market share

Klepper's 2009 review ends with some questions, each one providing an avenue for future research on spinoffs. The first one regards whether employees exploit knowledge from the parent and which kind of knowledge can give them a competitive edge. Further studies are, in fact, required to assess whether spinoffs perform better because of the knowledge inherited or because better firms hire better employees that in turn found better spinoffs.

A second question deals with the extent to which organizational culture and firm size are conducive to new firms' performance and foundation by employees.

The third and most debated question is related to the externalities that spinoffs formation can bring about. The main dilemma is whether spinoffs are "rapacious plunderers" or "paragons of innovation" (Klepper, 2001, p.639) and to what extent they seize parents' discoveries. Further, it is asked whether a potential spinoffs threat dampens investments in R&D at the parent level and what are the effects of the usage of non-compete covenants with respect to places where these are not allowed. Moreover, the literature debates on the extent to which spinoffs innovate more than other *de novo* firms, given the fact that some scholars claim their inherent innovative behaviour at the start-up phase whereas others maintain the knowledge inheritance theory as predictor of less innovation.

Another question concerns with organizational changes that spur spinoffs' formation: an analysis on what exactly triggers new ventures at the organizational level is required.

Finally, questions surround the relationship between cluster formation and spinoffs. Clusters see higher spawning rate but much remains to be discovered on the motivation leading to such proliferation and how the society can benefit from them.

1.8 Conclusion

Klepper's contribution provides a useful synthesis of the literature on spinoffs, highlighting gaps and drawbacks of the knowledge accumulated, such as inconsistent studies and the focus on a limited array of industries of analysis. Several authors have also tried to contribute to spinoffs' knowledge from different points of view, ranging from a more industrial oriented (as in the case of Buenstorf [2009] and Klepper [2001, 2002, 2007, 2009] towards a managerial one (as Helfat and Lieberman (2002) and Agarwal et al. (2004)). Motives behind spinoffs formation, elements affecting spinoffs performance and rate of spawning and dynamics inside clusters are some of the topics that entrepreneurial spinoffs literature dealt with up to 2009.

It stands to reason to expect that the proliferation of papers on spinoffs has not been deterred by Klepper's latest review, which can instead be considered a useful starting point for scholars who investigate entrepreneurial spinoffs topic and are willing to pursue research areas identified by previous researchers. Given the motley nature of spinoffs research highlighted by Klepper's 2009 review, it is likewise reasonable to expect that studies after 2009 have occurred in different thematic areas, some of which identified by the aforementioned review.

Our goal is thereby to *provide a systematization of the literature after Klepper's latest review, in order to identify, if possible, different thematic areas that correspond to subfields of research that authors have, albeit unwittingly, investigated. Furthermore, the increasing focus on spinoffs demands a systematization of the literature post 2009, which, to the best of our knowledge, has not been performed yet*.

Such systematization should not only fix up papers on spinoffs after 2009, categorizing them in distinct areas of research, but also identify where the literature is going and what are the research areas that have appealed more in the very recent years.

CHAPTER 2: METHODOLOGY OF LITERATURE SYSTEMATIZATION

2.1 Introduction

The increasing interest by researchers and scholars in entrepreneurial spinoffs, and in new ventures in general, encourages a deeper understanding of the knowledge accumulated heretofore on this topic. A systematization of the literature is usually performed through quantitative techniques and it allows researchers to grasp the thematic specialties which a knowledge domain is composed by. In the present chapter, we describe the methodology followed to identify thematic subdomains of entrepreneurial spinoffs literature *post* Klepper's 2009 review.

2.2 Knowledge mapping and steps of the procedure

In most cases, scholars of different research fields try to observe and analyse the themes and their evolution (Ferreira et al., 2017) while gaining insights into emergent patterns (Chen et al., 2010) and providing directions for future research (Culnan, 1986) by mapping the cognitive structure of a field. *Mapping the knowledge* is defined by Garfield et al. (1978, p. 192) as "an attempt to arrive at a physical representation of fields and disciplines -and, at a lower level, of individual papers and scientists- in which the relative locations of entities is depicted"; this allows the definition of subdomains or specialties in each field of research and the results can be exploited by scholars and policy makers (Özçinar, 2015). This idea stems from the *Citation Network* concept, developed first by Garfield et al. (1964): the underlying assumption is that publications in a research field form the network of the knowledge; being a network a collection of nodes joined together in pairs by edges (Newman, 2010), publications represent the nodes and citations between documents represent the edges, thus the relationships between papers, in such a network (Sorkun and Furlan, 2016). Shedding light on the structure of the knowledge requires therefore the depiction of the relationships created by publishing scientists and the conception of the research field as a mosaic of different specialties, whose identification is not trivial (Small and Griffith, 1974). Our endeavour is thereby to map the intellectual field of spinoffs, without claiming to identify a specific number of specialties: we cannot assert *a priori* that some subdomains will be found, but we may posit that different facets of spinoffs have been investigated; whether these aspects represent different subdomains must be ascertained.

The quest for subdomains is generally accomplished through bibliometric techniques. *Bibliometrics* can be defined as a quantitative analysis on scientific publications (Thelwall, 2008) which relies on statistical and mathematical tools to study patterns in the use and publication of documents (Diodato, 1994). Bibliometric methods therefore allow the objectivity and the quantifiability of the results, avoiding subjective biases (Nerur et al., 2008). According to Thelwall (2008) there exist two types of bibliometrics: *evaluative* and *relational*; while the former enables the comparison among the contributions of different groups of scholars, the latter attempts to find relationships within research, discover research opportunities and detect the emergence of national and international co-authorship patterns. Relational bibliometrics subsumes different techniques which are based on different units of analysis (documents, authors, words and journals).

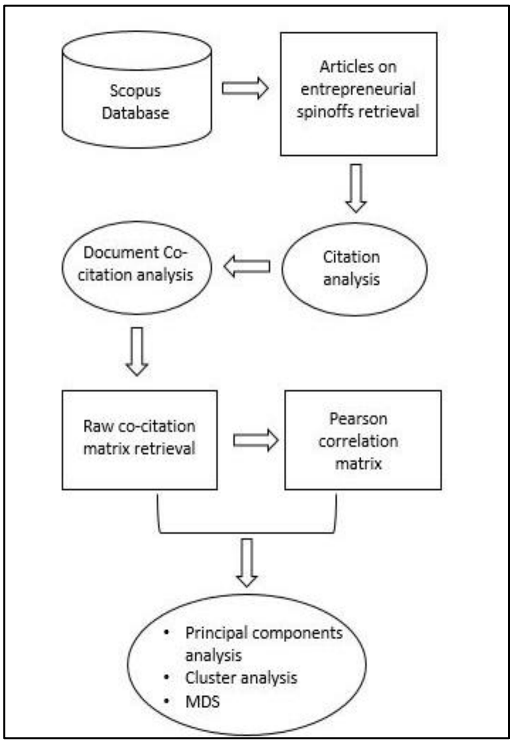
Citation analysis and co-citation analysis are among the most used relational bibliometric techniques for mapping the structure of the knowledge (e.g. McCain, 1990 analysed the intellectual structure of macroeconomics; Özçinar, 2015 mapped the teacher education domain; Ramos-Rodríguez and Ruíz-Navarro, 2004 and later Nerur et al., 2008 examined the structure of the strategic management field) and we also made use of them in our study. While citation analysis allows the identification of the most prominent authors or documents in a research field (Ramos-Rodríguez and Ruíz-Navarro, 2004), co-citation analysis provides the map of such a field by recognizing clusters or subgroups of authors or documents that are more similar from their references (Hsiao and Yang, 2011). Co-citation is in fact a measure of association between two authors or papers, which is computed through the frequency with which a pair of authors or papers have been cited together by subsequent authors or papers (Small and Griffith, 1974). The premise is that the more two documents or authors are jointly cited, the more they share similar contents and knowledge (White and Griffith, 1981) and this exhibits a natural and objective way for clustering them (Small and Griffith, 1974); authors and papers perceived as related tend to group together when mapped, whereas those perceived as different do not (Culnan, 1986).

However, to perform citation and co-citation methods in our study, the units of analysis needed to be chosen. McCain (1990) clearly distinguishes between *author-based analyses* and *document-based* ones, despite the acknowledgement of common assumptions and techniques between the two. While in author-based analyses the authors of the documents are the units of analysis which form the network of the knowledge and are then grouped in different clusters or specialties (McCain, 1990), in document-based analyses, the documents are the units of analysis and the network of knowledge they establish provides clusters of documents (Hsiao and Yang, 2011). We relied on document-based analysis for some reasons: first, the network of the knowledge is not based on authors themselves but on the topics of their publications and therefore on their documents; second, clusters of authors can contribute to identify different school of thoughts but not necessarily different subject areas: while documents are less likely to belong to different subfields of research, authors are likely to investigate different topics (Özçinar, 2015), thereby the relationship between author and cluster may be not univocal; finally, references in document analyses are expected to reveal more than the ones provided by author analyses, given the specific nature of the use of document-based references (Chen et al., 2010).

Bibliometric analysis is usually performed once a research field has reached a certain degree of maturity, by splitting the period of interest in sub-periods of analysis; it is also true that bibliometric tools are less likely to have been applied when the topic is quite novel (Ramos-Rodríguez and Ruíz-Navarro, 2004). Entrepreneurial spinoffs topic is not as mature as other research fields which have already been investigated through a bibliometric lens; nonetheless, the recent years witnessed a copious proliferation of papers which can justify a quantitative analysis on the literature, that is the purpose of this thesis. As reported in the previous chapter in fact, the study by Ferreira et al. (2017) shows an increasing emphasis on entrepreneurial spinoffs with respect to the other two kinds of spinoffs and, to the best of our knowledge, no bibliometric study has dealt with entrepreneurial spinoffs yet.

The goals of our analysis are therefore three: first, finding the seminal papers on entrepreneurial spinoffs and the most influential authors; second, understanding possible subfields inside this research field and interlinkages among them and finally designing future agendas.

Fig. 2. 1: Steps of analysis.
Source: adapted from McCain (1990)



Our endeavour to depict the intellectual structure of entrepreneurial spinoffs follows a stepwise approach, similar to McCain's (1990) and based on other bibliometric studies undertaken in different research fields (e.g. Nerur, et al., 2008; Ferreira et al., 2017). The process is illustrated in Fig. 2.1. First, the main documents on entrepreneurial spinoffs from 2009 to 2017 are retrieved from *Scopus* database, then the bibliometric tools *citation analysis* and *document co-citation analysis* are performed and finally their outputs are converted into a raw co-citation matrix and a correlation matrix and then used as starting point for quantitative analyses -*principal component analysis*, *cluster analysis* and *multidimensional* *scaling* – performed by using the software UCINET (Version 6.630[1]). These three techniques aim at grouping the documents on entrepreneurial spinoffs and, in wider terms, understanding the structure of the literature on this topic. All these steps will be described in detail in the following sections.

2.3 Selection of the documents

The selection of the pool of documents (called *source documents*) was performed through a Boolean search into the bibliographic database Scopus, which has proved to provide good coverage (Thelwall, 2008) and it was accomplished through some choices. First, the keyword string was used with the Boolean operator "OR" for the research in "title, abstract and keywords": the searched terms were "spin-off", "spinoff", "spin-out" and "spinout", both in plural and singular form, as all of these have been employed in the literature to identify *de novo* ventures founded by ex-employees of incumbent firms operating in the same industries of the new firms (e.g. Agarwal et al. (2004) use "spin-out", Chatterji (2009) uses spin-off, Klepper (2009) uses "spinoff"). The choice to search into "title, abstract and keywords" was due to the fact that these elements in databases usually provide a good synthesis of articles' contents. Second, the period was limited between 2009 and July 2017: the choice of the starting period is coherent with the fact that in 2009 Klepper wrote the latest review and synthesis regarding entrepreneurial spinoffs up to 2009 and it is deemed a seminal paper (see paragraph 1.7). Third, the scope of the research in Scopus was limited to the subject areas "Business, management and accounting", "Economics, econometrics and finance" and "Social sciences" because articles on spinoffs are expected to be found in these fields. Fourth, similarly to the study of Ramos-Rodríguez and Ruíz-Navarro (2004), only articles published in journals were considered, since they represent a sort of "certified knowledge", subject to the review of researchers.

2.4 Citation analysis

Counting citations to publications in the literature is one of the ways through which bibliometric studies can provide statistical distributions (Culnan, 1986). Citation analysis is one of the techniques adopted in bibliometric studies and it is the starting point to conduct a co-citation analysis. Its assumption is that authors cite those papers or authors they deem relevant in their research field (Ramos-Rodríguez, Ruíz-Navarro, 2004); stated differently, citations are surrogates for the influence of information from some documents on other documents and therefore they measure the dependence of scholars on previous works (Culnan, 1986), enhancing at the same time the reliability of the citers (Ramos-Rodríguez, Ruíz-Navarro, 2004).

[1] Developed by Borgatti et al., 2002.

Citation analyses are performed by counting the frequency of citation of one document in the references of all the other documents and this entails the recognition of the most influential papers (Ferreira et al., 2017).
The output of citation analysis is in fact the identification of highly cited documents, whose contents are milestones that can represent the way to understand a topic (Moed, 2005).
While some papers conduct citation analysis (e.g. Ramos-Rodríguez and Ruíz-Navarro, 2004; Backhaus et al., 2011) starting from the identification of a pool of documents whose references form the set of elements on which citation and co-citation analyses are performed, we followed instead the method adopted by Hsiao and Yang (2011): according to this latter, the analyses are conducted on the source documents retrieved themselves and not on their references. As a matter of fact, the focus of this study is to provide a snapshot of the intellectual structure on spinoffs after Klepper's 2009 review, identifying the possible subdomains, or specialties, of this knowledge; the objects of our analysis are therefore those articles that have been disclosed since that date and not their references, as these latter may deal with methods and topics unrelated to spinoffs that have been cited for each specific article's purpose. In other words, finding the most influential papers among the references of documents on entrepreneurial spinoffs and clustering such references to detect subdomains of knowledge can be misleading. Furthermore, in other bibliometric studies, documents subject to citation and co-citation analyses are the ones with the highest citation rates: this may cause a biased selection of the papers, as frequently cited works for methodological purposes may exclude from the analysis more thematically coherent and less cited works. Following the method of Hsiao and Yang (2011), once the source documents are defined, the set of papers citing these documents can be retrieved. The database Scopus was employed also for the identification of citing papers for each article. This procedure resulted in 553 citing papers and some of these cite more than one source document.
To perform citation analysis, some procedures were followed. First, for each citing document the bibliographic references were retrieved from Scopus and Microsoft Excel files were compiled with data on authors, years of disclosure and words from the titles as reported by Scopus output. Then, the statistical software R was employed starting from these Excel files in order to recognize the perfect correspondence between the list of references of citing documents and keywords from titles, authors and years of the source documents. The output of this procedure was a *citation matrix*, which displayed on the first row the source documents retrieved -the set of cited papers- and on the first column the set of citing papers; the cells of the matrix were accordingly filled with 0 and 1 by R: a cell was filled with 1 if the citing paper reported the correspondent column paper among its references and 0 if not. Since R recognizes

only exact symbols, a manual normalization was then performed, with the aim of providing higher reliability and accuracy to the automatic results (some inconsistencies were found, such as the case of one citing paper whose references reported the wrong year of disclosure of a source document, resulting in an erroneous 0 in the citation matrix and therefore corrected).

2.5 Co-citation analysis

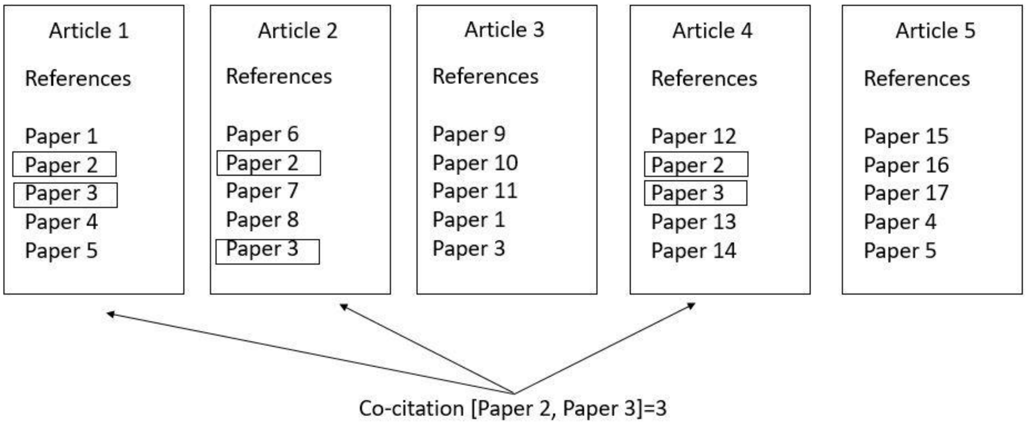
As stated in the previous sections, *co-citation* is a relationship established by citing documents (Small, 1973) which identifies ties between papers that authors in a subdomain of knowledge deem important (Small and Griffith, 1974). Since co-citation measures the strength of the relationship between couple of items (in our case two articles) it provides an indication of their intellectual proximity (Ferreira et al., 2017) and therefore it is deemed a valid approach for exploring the intellectual structure of a research field (Nerur, et al., 2008; White and Griffith, 1981). Co-citation is calculated in fact through the frequency with which a couple of papers are cited together by the citing set; when a strong co-citation exists between the pair of articles, many new papers are more likely to cite jointly them as well in the future (Small and Griffith, 1974). The upshot of this process is a network of relationships created by authors through the diffusion of knowledge (Nerur, et al., 2008).

It is worth specifying that co-citation must not be confused with bibliographic coupling: this latter refers to the situation in which two papers cite one or more papers in common and a higher similarity between these two papers occurs the more citations they have in common. Bibliographic coupling is static as it is dependent only upon citations contained in paired documents -and therefore it cannot be changed-, while co-citation is a relationship recognized by current scholars and therefore it is likely to change, as new discoveries are made (Small and Griffith, 1974). Further, co-citation embeds opinions of different authors and for this reason is a more reliable indicator than bibliographic coupling (Small, 1973).

In Fig. 2.2 the co-citation procedure is depicted by means of a simplified example: given a set of five citing articles and considered their references, the co-citation count results in 3 for Paper 2 and Paper 3, since they are cited jointly three times (by articles #1, #2 and #4); likewise, co-citation for Paper 1 and Paper 3 results in 2 (they are cited jointly by articles #1 and #3) and the same frequency is observed for Paper 4 and Paper 5 (they are jointly cited by articles #1 and #5).

Among all the couples of papers, Paper 2 and Paper 3 are characterized by the strongest co-citation link since they exhibit the highest co-citation count and therefore they are more likely to deal with the same topic of research; on the other hand, Paper 1 and Paper 4 are cited just by article #1: these two papers are more likely to belong to different clusters of topics or at least to deal with more dissimilar themes than Paper 2 and 3, according to co-citation. Paper 1 and 3 are instead characterized by a higher frequency than Paper 1 and 4 and therefore their intellectual proximity is closer.

Fig. 2. 2: Co-citation count. Source: adapted from Ramos-Rodríguez and Ruiz-Navarro (2004)



The output of co-citation analysis is usually a square symmetric raw data matrix with identically ordered papers on columns and rows (McCain, 1990) which reports the frequency with which a column paper and a row paper are jointly cited by the set of citing documents.

Deciding the values to be put in the diagonal cells of the matrix is a quite debated topic (McCain, 1990; Culnan, 1986; Nerur et al., 2008). While some authors argue for treating the diagonal cells as "missing values" (e.g. McCain, 1990; Ramos-Rodríguez and Ruíz-Navarro, 2004), some others (e.g. Culnan, 1986; Nerur et al., 2008, Hsiao and Yang, 2011; Özçinar, 2015) choose to create an *artificial value* computed by the sum of the three highest frequencies of that specific column or raw -i.e. the highest co-citation frequencies of a specific document- and dividing it by two, thereby considering the relative importance of the document in a subfield of research (White and Griffith, 1981). Since neither methodology compromises the results evidently (McCain, 1990) and the artificial value seems to be widely applied in the most recent co-citation analyses, we opted for this latter choice.

It is also worth restating that co-citation analysis was performed on the set of cited and citing articles identified by the procedure outlined by Hsiao and Yang (2011): usually in fact, bibliometric studies start from the identification of papers on a particular topic and consider their references as the cited documents upon which co-citation analysis is conducted; unfortunately, this reveals two major drawbacks: on one side, the cited articles can deal with topics unrelated to the ones under investigation and they can be cited for reasons other than content similarity; on the other hand, the size of the total cited articles can be quite large and for technical reasons is hardly possible to analyse all the documents. This is the reason why some bibliometric analyses (e.g. McCain, 1990; Ramos-Rodríguez and Ruíz-Navarro, 2004) establish a cut-off citation threshold for considering only the most cited documents and coping with this problem.

The procedure to obtain the raw co-citation matrix for papers on entrepreneurial spinoffs encompassed several steps, some of which are reported in other studies on document co-citation analysis (e.g. Small and Griffith, 1974; Hsiao and Yang, 2011):

- Each of the retrieved articles was paired with every other document within the set of articles: this resulted in n(n+1)/2 pairs;
- The software R was used in this phase of the analysis. Starting from the citation matrix described in the previous section and using Microsoft Excel files, n-1 sub-matrices were created, each one corresponding to a specific article, with a decreasing number of columns. The first column of sub-matrices reported the 553 citing papers, whereas the first row reported the pairs of articles of the sample, keeping fixed in each pair the article whose sub-matrix was referred to (for example, the first sub-matrix, related to the first article on spinoffs, exhibited in the first row pairs formed by Paper #1 with every other paper; the second sub-matrix the pairs formed by Paper #2 with every other paper, except for Paper #1, since the pair Paper #2 and Paper #1 had already been computed in the first sub-matrix; the third sub-matrix exhibited the pairs formed by Paper #3 with every other article, except for Papers #1 and #2 which were already been analysed, and so forth). The decreasing number of columns was thereby due to the fact that whenever a pair of articles had been previously analysed, it was not included in the following sub-matrices. The cells were filled with 1 and 0, according to whether a row citing article cited both the papers on entrepreneurial spinoffs correspondent to each specific column of the sub-matrices (1) or not (0).
- The co-citation count was arranged by creating a square symmetric matrix using Microsoft Excel: it reported as first row and first column the identically ordered papers on entrepreneurial spinoffs. The off-diagonal cells were filled in the following way: since each cell represented the number of times the correspondent column and row papers were cited together by the set of citing papers, that space was filled with the sum of the column in the sub-matrix in which there was that pair of papers (i.e. the co-citation count between Papers #1 and #2 in the co-citation matrix was the sum of the column reporting the pair Papers #1 and #2 in the first sub-matrix, retrieved by the procedure using R; the co-citation count between Papers #7 and #8 in the co-citation matrix was the sum of the column reporting the pair Papers #7 and #8 in the seventh sub-matrix, retrieved by the procedure using R). This mechanism made the lower and the upper parts of the matrix identical (McCain, 1990). It is worth pointing out that the more a column in a sub-matrix reported values 1 (the more citing papers cited the pair of spinoffs papers correspondent to that column), the higher was the

number in the specific cell in the co-citation matrix. The diagonal values were instead filled with the artificial value mentioned earlier.

- The raw co-citation matrix reported co-citation counts of papers on spinoffs; nonetheless, some of these articles have not been cited yet (according to the information provided by Scopus database) and therefore their co-citation with every other article is always 0. The articles so identified have been removed from the Excel file, leading to a smaller square matrix.
- This smaller matrix was further adjusted by virtue of a threshold: as total co-citation frequency of some articles was very low, only papers with a total co-citation value higher than or equal to 3 were retained. The aim was to avoid biases related to the involvement of articles that could have been untied to the other ones, i.e. that may have failed to be grouped within a thematic group on spinoffs as they were too different from all the others and thus outliers. Consequently, only articles with more relationships or strong ties with other papers have been considered for more accurate results. This resulted in a further smaller raw co-citation matrix.
- A manual check was performed to ensure the reliability of the automated procedure.

The set of cited documents on entrepreneurial spinoffs in raw co-citation matrix (32 by 32) was deemed an appropriate number for all the analyses performed and Hsiao and Yang's methodology (2011) was deemed more coherent for our clustering purposes.

2.6 Pearson Correlation matrix

Retrieving the co-citation matrix is a critical step in bibliometric studies and, in some cases, it is used as input for the following quantitative analyses leading to knowledge mapping (Culnan, 1986; Nerur et al., 2008; Özçinar, 2015; Ferreira et al., 2017). Indeed, co-citation matrix is a proximity similarity matrix: the higher the number in a cell, the more similar two documents are (Leydesdorff and Vaughan, 2006). Nonetheless, some other authors argue for the application of Pearson correlation matrix as input for the analyses with respect to raw co-citation matrix (e.g. McCain, 1990; Ramos-Rodríguez and Ruíz-Navarro, 2004; Hsiao and Yang, 2011). Pearson correlation is deemed, in fact, an acknowledged measure of similarity; the correlation matrix resulting from raw co-citation becomes a matrix of proximity values that signals the similarity or dissimilarity between couples of elements (McCain, 1990) -in our case articles-. Supporters of Pearson correlation highlight two major advantages with respect to raw co-citation. First, for any pair of documents, Pearson correlation coefficient is a measure that exposes the similarity of their co-cited profiles with respect to all the other articles, thus it is not the mere frequency with which two works are jointly cited: this means that when two articles are always co-cited along with a third one and rarely with the others, they will have a high positive correlation and will be perceived as similar by the citing set; at the same time, correlation copes with problems of "scale" between highly cited documents and less frequently cited ones (McCain, 1990; Ramos-Rodríguez and Ruíz-Navarro, 2004).

The formula used to compute Pearson correlation coefficient *r* is the following:

$$r = \frac{N(\Sigma XY) - (\Sigma X)(\Sigma Y)}{\sqrt{[(N\Sigma X^2) - \Sigma(X^2)][(N\Sigma Y^2) - \Sigma(Y^2)]}}$$

To convert the raw co-citation matrix into a Pearson correlation matrix using Pearson correlation coefficient formula, for a given pair of articles A and B, each element of vector X will be the co-citation frequency of article A with every other article and each element of vector Y will be the co-citation frequency of article B with every other article and N is the number of articles subject to co-citation less 1, as the self-citation is meaningless. Pearson correlation is always between -1 and +1 (He and Hui, 2002). Since it measures the strength of the relation between couples of articles, values equal or close to +1 mean a strong tie between the elements of the pair, while the lower the values the weaker their bond.

On the other side, there is no shortage of supporters of raw co-citation matrix with respect to Pearson correlation as input for quantitative analyses aiming at clustering. In the study conducted by Ahlgren et al. (2003), Pearson correlation coefficient fails the test in stability of similarity measurement between authors in an author co-citation analysis. Pearson correlation is criticized also by Leydesdorff and Vaughan (2006) who claims the inherent proximity nature of the raw co-citation matrix, implying the uselessness of further similarity measures' application to build a proximity matrix.

Since Pearson correlation is featured by both advantages and drawbacks, the decision to choose either raw co-citation matrix or Pearson correlation was taken after a discussion with Stephen Borgatti, one of UCINET's developers, and after having read UCINET guide, in which the inputs required for each quantitative analysis are illustrated. Borgatti's opinion was that raw co-citation matrix could be treated as similarity measure and therefore no meaningful difference would emerge from this choice.

Therefore, raw co-citation matrix was used as input for principal components analysis (PCA, see below), from which the correlation matrix was retrieved, while Pearson correlation served as input for multidimensional scaling and cluster analysis; this approach was coherent with more recent co-citation studies (e.g. Nerur et al., 2008; Özçinar, 2015).

2.7 Multivariate analyses for clustering

In bibliometric studies, three approaches to multivariate analysis are used to outline and map the relationships between papers and authors starting from the similarity matrix: these are factor analysis based on principal components analysis (PCA), multidimensional scaling (MDS) and cluster analysis (McCain, 1990). These are called *interdependence techniques* since the variables studied cannot be classified as dependent or independent; they are in fact analysed simultaneously to find the underlying structure of the set of objects. They are also called *data reduction techniques*, as they allow the dimensionality reduction and the extraction of only meaningful information (Hair et al., 2010).

While these techniques may apparently seem redundant to extract information from data, each one of them is characterized by its specificity. PCA, when applied for co-citation purposes, entails the extraction of the subdomains in the research field and the identification of the main entities (in our case articles) in research specialties (Nerur et al., 2008): extracting the number of key conceptual themes is an important issue in co-citation analyses. A deeper understanding of knowledge structure is accomplished through multidimensional scaling, a technique by which data are plotted on a two-dimensional plane (Ramos-Rodríguez and Ruíz-Navarro, 2004), thus providing visual representations of co-citation information. Finally, MDS is usually used together with cluster analysis, whose goal is to group entities according to some shared attributes (McCain, 1990). The following sections will describe how these three interdependent techniques work and how they have been used in our study.

2.7.1 Principal Components Analysis

Factor analysis encompasses methods used to find unobservable but interpretable variables, which are usually based on principal components analysis (PCA) (Azzalini and Scarpa, 2012). Several citation and co-citation studies on different research topics relied on PCA (e.g. Nerur et al., 2008; Hsiao and Yang, 2011; Özçinar, 2015; Ferreira et al., 2017) to identify subfields of research. PCA is in fact the most frequently used set of techniques to derive a smaller number of variables by linear combination of the original variables that can explain most part of the variability of these latter (Azzalini and Scarpa, 2012). The idea behind PCA is the following: given *n* observations measured by *p* features $X_1, X_2, \ldots, X_p$, when *p* becomes larger, a method to find a lower dimensional representation of the data is desirable. Each of the *n* observations lives in a *p*-dimensional space but not all the dimensions explain a large part of the variability: PCA allows the identification of a smaller number of dimensions that are as interesting as possible and these dimensions are called *principal components*. The first principal component of a set of characteristics $X_1, X_2, \ldots, X_p$ is the linear combination of those features $Z_1=\varphi_{11}X_1$ +

$\varphi_{21}X_2 + \dots + \varphi_{p1}X_p$ that has the largest variance; such linear combination is normalized, thus the sum of the squares of the *loadings* is equal to one, where loadings are those elements φ_{11}, φ_{21}, …, φ_{p1} for the first principal component; together, the loadings form the principal component loading vector, which defines the direction in the feature space, where data have the highest variance. The second principal component Z_2 is instead the linear combination X_1, X_2, …, X_p with the maximal variance with respect to all the other linear combinations that are uncorrelated with Z_1. The first principal component is uncorrelated with the second one when the second component loading vector is orthogonal to the first component loading vector (James et al., 2013). When the percentage of variability is largely explained by the first *k* components, the remaining components can be eliminated and take the first *k* components as new independent variables (Azzalini and Scarpa, 2012).

In co-citation analysis, a *component* is interpreted as the subset of elements which load on it, that is substantially contributing to its construction (McCain, 1990). Since co-citation analysis uses the raw co-citation matrix as input, PCA allows the identification of subfields of research from co-citation frequencies. Subfields are represented by the components extracted by PCA and each subfield is defined by works (in our case articles) with higher loads on that component or subfield; the loading represents in fact the degree to which an element belongs to a specific component (Nerur et al., 2008) and therefore the conceptual ideas represented by that component (Lee and Chen, 2011). Thus, by reading articles contributing to the same specific component, the theoretical basis of that subfield of research can be derived. Consequently, documents with high co-citation frequencies are likely to contribute to the same component, as they are perceived similar from the pool of citers (Hsiao and Yang, 2011). In co-citation analysis, every element contributes to each extracted component, but the definition of the latter relies on the elements with high loads, i.e. that contribute more to the identification of such component. A high loading and therefore a strong contribution is given by values equal to or higher than |0.7|; values equal to |0.4|- |0.5| are quite significant, whereas values lower than |0.4| are negligible for that component (McCain, 1990). While most documents have a high load on only one component, PCA exhibits the breadth of contribution of documents with significant loads on more than one component and thus it has an advantage over other techniques -cluster analysis and mapping techniques- in which a single document does not appear in more than one subgroup or cluster (He and Hui, 2002). Articles with a deep influence on the research field are likely to appear in more components (Nerur et al., 2008), as they contribute to the development of the whole research field.

Components were extracted through an orthogonal rotation of the components, the *Varimax rotation*: in accordance with the theory, this results in uncorrelated components with most documents having high loads on only one of them (McCain, 1990) and, according to UCINET guide, this maximizes the purity of the components themselves.

In principle, the number of principal components extracted is the same as the number of original variables input as dataset, therefore, a stopping rule is required (McCain, 1990): the more components extracted, the smaller is the variance that the last components can explain. In our study, the stopping rule was the establishment of the minimum eigenvalue equal to 1 as in other studies (e.g. McCain, 1990; Nerur et al., 2008; Özçinar, 2015); eigenvalues indicate the amount of variance accounted for by each principal component (Hair et al., 2010): for this, choosing eigenvalues lower than 1 would be meaningless. Finally, the number of components to be extracted was set equal 10, which is a quite reasonable number to explain the variability of the dataset. In general, however, we wish to use the smallest number of components to get a good understanding of the data, even though the question on how many components to be considered is inherently undefined. Some techniques are used to define the most appropriate number of components, such as looking at the cumulative variance explained by each component (James et al., 2013). This analysis was complemented by the output provided by the software SPSS statistics 22.0.

2.7.2 Cluster analysis

Cluster analysis is one of the multivariate tools to provide insights into the relationships between documents that form the knowledge domain of a specific topic. This analysis encompasses several methods to group elements based on shared characteristics; when applied to co-citation studies, it allows researchers to understand the intellectual field by grouping documents characterised by higher similarity. Cluster analysis is a visual technique that allows an easier interpretation of the data but its output, in co-citation studies, should be jointly interpreted in light of results coming from PCA and MDS (McCain, 1990). Cluster analysis classifies items on a set of characteristics; the resulting clusters should exhibit high homogeneity within each group and high between-group heterogeneity (Hair et al., 2010). Thereby, when observations of a dataset are clustered, distinct groups are created and while observations within a group are quite similar, observations of different groups are different from each other. Both PCA and cluster analysis look for the identification of a smaller number of determinants but their logic is quite different: PCA tries to find a lower dimensional representation of the data explaining a quite high percentage of the total variance and clustering aims to find homogeneous groups among the set of observations.

Several clustering approaches have been developed but the most famous are *K-means* and *hierarchical clustering*. While K-means seeks group observations, given a prespecified number of groups, in hierarchical clustering the number of clusters to come up with is not known in advance and it provides a representation of observations similar to a tree called *dendrogram* (James et al., 2013). In this latter clustering method, the data are hierarchically structured and organized into groups; this is done by associating the items with a binary structure in which the leaves are the units and the nodes are the subgroups of points; the structure of this clustering technique introduces a hierarchy in the subgroups. The most famous hierarchical method follows a bottom-up approach and it is the *hierarchical agglomerative clustering*. The agglomerative method starts with the identification of as many subgroups as the number of the units of analysis (Azzalini and Scarpa, 2012), then, a simple rule is applied: combine the two most similar units that are not already in the same cluster. This repeated procedure goes forth until all observations belong to a single cluster. This method is called *hierarchical* because it follows a stepwise approach until a range of cluster solutions are created and it is also *agglomerative* because it is based on a peer combination of existing clusters (Hair et al., 2010). In the dendrogram, this is quite clearly represented: each leaf of the tree corresponds to the units of the dataset and, as we move up to the trunk, some leaves become to be fused into branches; the sooner this fusion occurs, the more similar the observations are; on the other hand, when subgroups fuse near to the top part of the tree, the observations are deemed quite different. Therefore, in order to appreciate the similarity of two observations -in our analysis, two papers- it is sufficient to look at the point in which the branches containing both the observations are fused: the lower the distance between that point and the top of the tree, the lower the similarity. In co-citation analyses, hierarchical agglomerative clustering is the most used technique; the algorithm begins to consider each paper as a different cluster, then, inter-papers similarity is used to combine the different articles till the point in which all the articles belong to the same cluster (He and Hui, 2002).

While the concept of similarity can be clear when dealing with pairs of observations, it can raise some issues when similarity is considered between two groups of observations. To address this problem, the concept of *linkage* is introduced and it represents the degree of similarity between subgroups. Different types of linkage exist but statisticians prefer to apply the *complete* and the *average* ones, because of their more balanced outputs; a complete linkage relies on the maximal inter-cluster dissimilarity and average linkage is based on mean inter-cluster dissimilarity (James et al., 2013).

Pearson correlation matrix, which is deemed a fit similarity measure for intellectual proximity of articles, has been used as input for cluster analysis on entrepreneurial spinoffs topic.

Further, the software UCINET was employed to perform hierarchical agglomerative cluster analysis; this kind of clustering was chosen not only because of its broad application in co-citation studies, rather because of the underlying logic: since the focus of this thesis is to depict the intellectual structure of a pretty novel topic, predetermining the number of clusters or subfields by applying K-means clustering would be unreasonable, also because thematic subgroups may not emerge.

UCINET provides also different linkages options: single, weighted, complete and simple mean; the *complete* one was chosen, as it computes the dissimilarity between members to define the distance between two clusters; moreover, a further endeavour was performed by using the weighted average link -the default option of UCINET- which failed to provide significantly different results from complete link.

Expecting to find the right number of subgroups in hierarchical clustering is a wild-goose chase. As a matter of fact, there are no preferable objective rules for it, but a usual approach suggests drawing a line cutting the dendrogram at the level where vertical branches are longer: the number of intersections resulting from branches and the line should be the number of clusters (Azzalini and Scarpa, 2012). Furthermore, the purpose of co-citation studies is to provide an insight into an intellectual field and to inform a more general discussion (McCain, 1990); for this reason, attempting to find the number of clusters of spinoffs literature should be conceived more as a simplistic mean to map the intellectual field, without claiming to be complete and exact.

2.7.3 Multidimensional Scaling

A recommended step in bibliometric analyses is the graphical representation of the subgroups of papers, to make information more effective and intuitive (He and Hui, 2002). MDS is a procedure that enables the researcher to define the relative image of a set of elements. The objective of multidimensional scaling (or perceptual mapping) is to transform similarity into distances represented in a multidimensional space; for instance, if objects A and B are considered the most similar compared with all the other pairs, techniques of perceptual mapping will place these two objects so that the distance between them in the multidimensional space is lower than the distance of any other couple of objects. The basis for the relative positioning of the objects is that any object has many dimensions representing its attributes or features. Attempts to represent the relative positioning of the objects can be performed by using a similarity scale and fit all the elements on it: that is a one-dimensional portrayal in which the similarities between the objects are represented by their distances on the line and only one dimension has been identified to classify them. Even though this approach can work with 3-4 objects, it becomes quite complex when the number of elements increase; often, in fact, at least two dimensions (or scales) are used to represent relative distances among objects (Hair et al., 2010). The purpose underlying MDS can be thereby identified in gathering the maximum amount of information from the data in two or three dimensions, reducing in this way the spatial dimension. Nonetheless, this procedure results in a strong simplification that distorts the original information -the original distances among the data- disregarding some of the variance explained in the original similarity matrix (Ramos-Rodríguez and Ruíz-Navarro, 2004). If a high distortion happens, a decision should be taken: either MDS should be abandoned (e.g. Culnan, 1986) or the number of dimensions representing the data should be increased. The latter option however is linked to some difficulties: three or more dimensions are quite difficult to display on paper and managing more than two dimensions can make MDS useless as a procedure to simplify data. The implicit trade-off in MDS is therefore the loss of information when the optimal configuration of points is in a two-dimensional (or low-dimensional) space, balanced with an extremely wieldy representation of a complex set of relationships that can be understood at a glance (Borgatti, 1997[2]). When the distortion occurs, it is reflected into a statistic called *stress*. The stress measure is a criterion to determine the "best fit" between the original distances of the similarity matrix and the estimated (or derived) distances in the low-dimensional map (McCain, 1990). The best solution of MDS is thereby the best fitting configuration of the mapped elements, that is the solution with the lowest stress (Kruskal, 1964). In other words, the *stress* measures the proportion of the variance explained by the *disparities* that the chosen MDS model does not take into consideration, where *disparities* are meant as the differences of the distances between elements in the visual map and the similarity of the matrix (Hair et al., 2010). From a mathematical point of view, a stress value different from 0 will occur when there is insufficient dimensionality, that is it is impossible to represent the dataset in low-dimensional space (Borgatti, 1997). The stress value in fact depends on the number of units in the dataset: if the number of units increases, the stress value will rise as well and thereby the poorer will be the goodness of fit when more items will be plotted on the map (Ramos-Rodríguez and Ruíz-Navarro, 2004); the logic behind stress value is similar to the one mentioned earlier, that is mapping more elements in a low-dimensional space is a tougher task. Nevertheless, a non-zero stress value is not mandatory for MDS to be successful and some amount of distortion is allowed. Kruskal (1964) proposes a rule of thumb to assess the acceptable values of the *stress* -that is a positive number as it has been computed as sum of squares-: when it assumes values below 0.1 the goodness of fit is deemed excellent, when it assumes values between 0.1 and 0.2 is deemed fair, when values are over 0.2 they are considered poor and therefore should not be accepted.

[2] Source: http://www.analytictech.com/borgatti/mds.htm

In co-citation analysis, MDS is applied quite frequently (e.g. McCain, 1990; Ramos-Rodríguez and Ruíz-Navarro, 2004; Nerur et al., 2008; Hsiao and Yang, 2011). MDS output from co-citation similarity matrix shows the relative position of papers (as in our analysis, which is based on document co-citation frequencies) or authors (in author co-citation analysis); the mapping principle is that the more similar documents are, the shorter the distance between them when they are plotted (Leydesdorff and Vaughan, 2006). This means that when two papers are heavily co-cited, and therefore their correlation coefficient in the Pearson correlation matrix is near to +1, the two papers are placed quite close. Thus, by looking at the perceptual map, it is possible to infer how the community of scholars dealing with entrepreneurial spinoffs conceive those papers that they cite in their works: papers deemed thematically similar are more likely to be grouped and to be located far apart groups of articles deemed dissimilar. By means of a simple visual representation, different strands of research belonging an intellectual field may be thereby identified.

MDS requires a matrix of similarity as input and, as mentioned earlier, the Pearson correlation matrix was chosen for this analysis. The software UCINET was used for this data reduction procedure as well. The first choice when using UCINET was the definition of either a metric MDS or non-metric MDS. Non-metric MDS does not assume any specific type of relationship between distance in the map and similarity and the input is typically formed by rank-ordered objects. Metric MDS, instead, assumes that both input and output are metric and this strengthens the relationship between the dimensionality of the final output and the input; in this case, not only the rank-ordered relationships of the data are maintained, but the ratio and interval qualities as well; however, no significant differences in the results are found between the two MDS (Hair et al., 2010). In a nutshell, this means that while in metric MDS there is a 1-to-1 (inverse) correspondence between the distances of the points on the map and the input correlations, in non-metric MDS only the rank-orders are preserved (i.e. the closest pair of papers in the map is the pair with the highest correlation, the second-closest pair in the map has the second-highest correlation, and so forth). The non-metric MDS was chosen as option in UCINET for a couple of reasons: first, Borgatti's opinion was that non-metric MDS typically does a better job in finding the underlying structure; second, several co-citation studies employ non-metric MDS; third, the ordered relationships among pairs of items are the real focus of our analysis, irrespective of interval and ratio measures. UCINET default program for non-metric MDS is TORSCA, that has been applied in other co-citation analyses (McCain, 1990) and that was chosen in our analysis as well.

Then, UCINET requires to set the number of dimensions for the dimensional representation.

Some attempts were made to find the number of dimensions to be input for the final configuration: first the software was run with one-dimension as option, then with a two-dimensions as option and so forth until a six-dimensions choice (the decision to put “6” as last dimension is purely arbitrary). This was done in compliance with the method suggested by Kruskal (1964) and Hair et al. (2010) for identifying the appropriate dimensionality of MDS: given stress values associated with different dimensionality and plotting stress against dimensions, good data shows a clear elbow in the curve correspondent to the most appropriate number of dimensions to be used in the analysis. The elbow indicates the point where the goodness-of-fit substantially improved (Hair et al., 2010). Therefore, the point indicating the better dimensionality was used as input for UCINET.

Care should be exercised when a non-zero stress map is interpreted, as all the data are, to some extent, distorted from the original one; in general, however, longer distances enable the visibility of main patterns since they are more accurate than shorter ones (Borgatti, 1997).

2.7.4 Labelling subfields

Labelling subfields of research is as useful as complicated when a bibliometric analysis is performed. Since PCA is a recognised method to derive the number of subfields, labelling activity should start from the classification obtained by PCA and thereafter compared to partition of cluster analysis and MDS. Two main methods can be used for this purpose: the first procedure is a *word frequency analysis* performed on titles and abstracts of the papers considered, whereas the second one is based on a *personal assessment* of the contents after having read the articles (Culnan, 1986).

Word frequency analysis was performed by using VOSViewer, an easy-to-use software tool concerned on visualization of bibliometric networks and representation of keywords of networks. The first step was the retrieval of information on titles and abstracts of all the articles considered in co-citation, since titles and abstract represent quite well the contents of the articles; then, a map based on text data was created by VOSViewer, considering five as the minimum co-occurrence of keywords from titles and abstracts and a binary method of counting: this latter option means that the number of times a noun is cited in the abstract or title of each article does not matter (Van Eck and Waltman, 2014); finally, a manual selection was performed for all the keywords identified aimed at excluding words not relevant for the analysis.

Besides, several sub-analyses on keywords were performed for each principal component. The logic behind was the following: if some keywords had been likely to occur a lot within articles contributing to a specific principal component, such words would have provided a thematic

indication of the contents of the papers belonging to that specific subfield of research. Therefore, the same steps of keywords occurrence analysis were followed for the subset of articles of each principal component, even though the co-occurrence threshold was set equal to two (instead of five) because of the lower number of items analysed (i.e. articles).
Despite these automatic results, we should also reckon with the fact that keywords from titles and abstract may lead to some biases and only an accurate reading of all the articles can provide an understanding of subfields of entrepreneurial spinoffs topic. Automatic procedures are thereby only the preliminary step to get acquainted with the different papers' themes.

2.8 Conclusion

Several quantitative tools help researchers and scholars provide a systematization on the knowledge accumulated up to a certain moment. Such methods are based on the concepts of *citation* and *co-citation* which the following quantitative and statistical steps stem from. In our case, we tried to employ them for systematizing spinoffs' literature after Klepper's 2009 review, as several strands are expected to have been investigated after that seminal paper. Thereby, we relied on Scopus database to retrieve relevant articles, we compiled a citation matrix, a raw co-citation matrix, we computed Pearson matrix and finally we performed principal components analysis, cluster analysis and MDS, which are expected to give similar results for literature systemization.
The results of these analyses are reported in chapter 3, while Fig 2.3 sums the procedure followed, from documents retrieval to quantitative analyses for clustering[3].

[3] These steps include simplified examples to provide a schematic presentation of the procedure followed

Fig. 2. 3: Steps followed for literature systematization. Source: personal elaboration

Scopus database

- Articles on entrepreneurial spinoffs (2009-2017)
- Keywords: spin-off, spinoff, spin-out and spinout, both plural and singular form
- Fields: Business, management and accounting, Economics, econometrics and finance and Social sciences

Paper #1	Paper #2	Paper #3
Citing Paper 1	Citing Paper 2	Citing Paper 2
Citing Paper 2	Citing Paper 3	Citing Paper 4
Citing Paper 3	Citing Paper 4	Citing Paper 5

From the set of papers on entrepreneurial spinoffs the list of citing paper (CP) was retrieved as well

	Citation matrix		
	Paper #1	Paper #2	Paper #3
CP 1	1	0	0
CP 2	1	1	1
CP 3	1	1	0
CP 4	0	1	1
CP 5	0	0	1

Papers on spinoffs are in columns while citing papers are in rows. When a citing paper cites the corresponding column paper the cell assumes value 1, otherwise 0.

	Submatrix 1		**Submatrix 2**
	#1 #2	#1 #3	#2 #3
CP 1	0	0	0
CP 2	1	1	1
CP 3	1	0	0
CP 4	0	0	1
CP 5	0	0	0

Submatrices containing pairs of articles are created. When each citing paper cites jointly the pair of articles in column the cell assumes value 1, otherwise 0.

The number of times each couple of papers is cited jointly by citing papers is computed as co-citation count

Pairs	**Co-citation count**
#1 #2	2
#1 #3	1
#2 #3	2

	Raw co-citation matrix		
	#1	#2	#3
#1	D_1	2	1
#2	2	D_2	2
#3	1	2	D_3

The raw co-citation matrix reports the co-citation count for all the pairs of cited articles.
The diagonal values are calculated by using an artificial value

	Pearson matrix		
	#1	#2	#3
#1	1	r_{12}	r_{13}
#2	r_{21}	1	r_{23}
#3	r_{31}	r_{32}	1

The correlation matrix is retrieved from co-citation matrix

Raw co-citation matrix and Pearson matrix are input for:

- Principal components Analysis
- Cluster Analysis
- Multidimensional Scaling

CHAPTER 3: RESULTS OF LITERATURE SYSTEMATIZATION

3.1 Introduction

This section reports the results of the analysis described in the previous chapter.

First, the retrieval of the documents is presented in a table containing titles, authors, year of disclosure and the identification numbers of the papers that are used in all the following procedures. Then, citation analysis results are provided, by identifying the main works in entrepreneurial spinoffs field from 2009 and by depicting the network of relationships among such works. Co-citation analysis outputs include the raw co-citation matrix and examples of articles pairs to assess the strength of the co-citation relationship among them. Finally, results of principal components analysis, cluster analysis and multidimensional scaling obtained by using the software UCINET and SPSS are presented jointly with techniques to detect the number of groups. The last part contains the discussion on subfields of research which stems from quantitative analyses and word frequency analysis for obtaining group labels; furthermore, avenues for future research are identified and results of a sensitivity analysis and limits of the procedure followed are presented.

3.2 Articles retrieval

The Boolean search in Scopus resulted in 744 articles by using as keywords “spinoff” and “spinout”, both with and without hyphen, both plural and singular form. After that, the documents were screened by first reading titles and abstracts; when this preliminary reading was not sufficient to decide whether to retain or to remove the documents from the analysis, introduction and conclusion of the papers were read as well. Finally, as the documents published in 2009 were disclosed before Klepper’s 2009 seminal paper, only documents from 2010, *post* Klepper’s review, were considered.

The establishment of a citation threshold is a common practice in bibliometric studies (McCain, 1990), that is the document is retained only if it has been cited at least as many times as the amount indicated by the threshold. Nonetheless, for this study no threshold has been fixed: on one hand because entrepreneurial spinoffs is a quite novel topic and in order bibliometric studies to be performed a certain number of papers should be considered; on the other hand less cited papers can contribute to understand the extant knowledge and the future dynamics. In sum, 61 source documents were retained and are reported in Table 3.1; they represent the core of the entrepreneurial spinoffs field from 2010 to 2017 and the starting point for the following quantitative analyses.

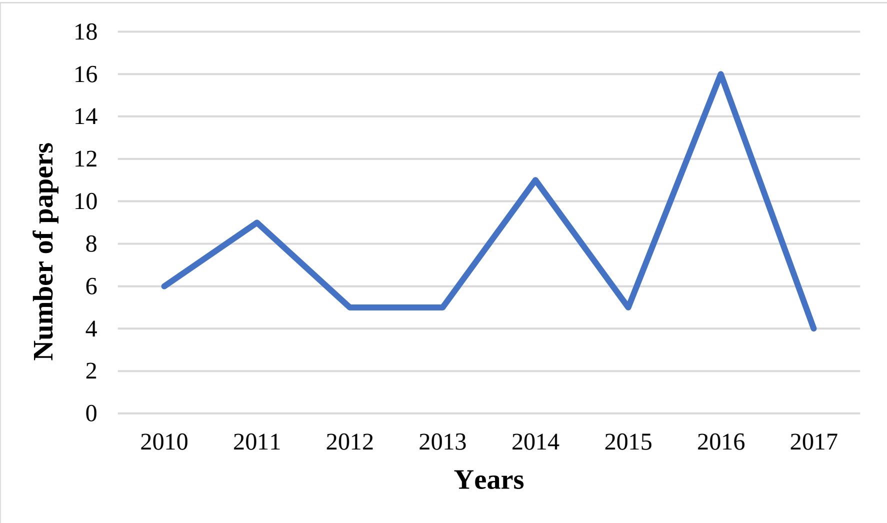
Table 3. 1: Source documents on entrepreneurial spinoffs retrieved from Scopus database. Source: personal elaboration

Paper	Year	Title	Authors
#1	2010	Regional corporate spawning and the role of homegrown companies	Avnimelech G., Feldman M.
#2	2010	Why does entry cluster geographically? Evidence from the US tire industry	Buenstorf G., Klepper S.
#3	2010	Policy principles for the creation and success of corporate and academic spin-offs	Gilsing V., van Burg E., Romme A. G. L.
#4	2010	Disagreements and intra-industry spinoffs	Klepper S., Thompson P.
#5	2010	The origin and growth of industry clusters: The making of Silicon Valley and Detroit	Klepper S.
#6	2010	Do R&D spinoffs have higher R&D productivity? Evidence from Taiwanese electronics firms	Yang C. H., Lin, H. L., Li H. Y.
#7	2011	Entrepreneurial Origin, Technological Knowledge, and the Growth of Spin-Off Companies	Clarysse B., Wright M., Van de Velde E.
#8	2011	Italian industrial districts as cognitive systems: Are they still reproducible?	Camuffo A., Grandinetti R.
#9	2011	"Cluster" creation by reconfiguring communities of practice	Karlsen A.
#10	2011	Performing in Dutch book publishing 1880-2008: The importance of entrepreneurial experience and the Amsterdam cluster	Heebels B., Boschma R.
#11	2011	Nano-economics, spinoffs, and the wealth of regions	Klepper S.
#12	2011	Locational conditions, cooperation, and innovativeness: Evidence from research and company spin-offs	Lejpras A., Stephan A.
#13	2011	Disagreements, employee spinoffs and the choice of technology	Thompson P., Chen J.
#14	2011	Comparative advantages of spinoff firms: An evolutionary perspective	Uzunca B.
#15	2011	The effectiveness of university knowledge spillovers: Performance differences between university spinoffs and corporate spinoffs	Wennberg K., Wiklund J., Wright M.
#16	2012	R&D strategies and entrepreneurial spawning	Andersson M., Baltzopoulos A., Lööf H.
#17	2012	Spinoffs and the market for ideas	Chatterjee S., Rossi-Hansberg E.
#18	2012	Home sweet home: Entrepreneurs' location choices and the performance of their ventures	Dahl M.S., Sorenson O.
#19	2012	Continuity and change in a spin-off venture: the process of reimprinting	Ferriani, S., Garnsey, E., Lorenzoni, G.
#20	2012	Employee spinoffs and other entrants: Stylized facts from Brazil	Muendler M.A., Rauch J.E., Tocoian, O.
#21	2013	Characteristics and performance of new firms and spinoffs in Sweden	Andersson M., Klepper S.
#22	2013	Impact of the Type of Corporate Spin-Off on Growth	Bruneel J., Van de Velde E., Clarysse B.
#23	2013	The 'Right' Knowledge and Spin-off Processes: An Empirical Analysis on Knowledge Transfer	Del Giudice M., Della Peruta M.R., Maggioni V.
#24	2013	Entrepreneurship in a Hub-and-Spoke Industrial District: Firm Survey Evidence from Seattle's Technology Industry	Mayer H.
#25	2013	Entrepreneurial spawning and firm characteristics	Habib M.A., Hege U., Mella-Barral P.
#26	2014	A corporation's culture as an impetus for spinoffs and a driving force of industry evolution	Cordes C., Richerson P.J., Schwesinger G.
#27	2014	Spin-off and clustering: A return to the Marshallian district	Cusmano L., Morrison A., Pandolfo E.
#28	2014	The who, why, and how of spinoffs	Dahl M.S., Sorenson O.
#29	2014	Job machine, think tank, or both: What makes corporate spin-offs different?	Fryges H., Müller B., Niefert M.
#30	2014	The origin of spin-offs: A typology of corporate and academic spin-offs	Fryges H., Wright M.
#31	2014	Spin-off performance in the start-up phase – A conceptual framework	Furlan A., Grandinetti R.
#32	2014	When do spinouts enhance parent firm performance? Evidence from the U.S. automobile industry, 1890-1986	Ioannou I.
#33	2014	How innovative are spin-offs at later stages of development? Comparing innovativeness of established research spin-offs and otherwise created firms	Lejpras A.
#34	2014	Survival, productivity and growth of new ventures across locations	Lööf H., Nabavi P.
#35	2014	Developing new ideas: Spin-outs, spinoffs, or internal divisions	Nikolowa R.
#36	2014	Parent hostility and spin-out performance	Walter S.G., Heinrichs S., Walter A.
#37	2015	Do spinoff dynamics or agglomeration externalities drive industry clustering? A reappraisal of Steven Klepper's work	Boschma R.
#38	2015	Spinoffs and the mobility of U.S. merchant semiconductor inventors	Cheyre C., Klepper S., Veloso F.
#39	2015	Spinoffs and the ascension of Silicon Valley	Cheyre C., Kowalski J., Veloso F.M.
#40	2015	What explains the survival gap of pushed and pulled corporate spin-offs?	Rocha V., Carneiro A., Varum C.
#41	2015	Organisational synergies, dissonance and spinoffs	Shrivastava M., Rao T.V.S.R.
#42	2016	What do I take with me? the mediating effect of spin-out team size and tenure on the founder-firm performance relationship	Agarwal R., Campbell B.A., Franco A.M., Ganco M.
#43	2016	Knowledge inheritance, vertical integration, and entrant survival in the early U.S. Auto industry	Argyres N., Mostafa R.
#44	2016	Spin-offs: Why geography matters	Baltzopoulos A., Braunerhjelm P., Tikoudis I.
#45	2016	Non-compete clauses, employee effort and spin-off entrepreneurship: A laboratory experiment	Buenstorf G., Engel C., Fischer S., Gueth W.
#46	2016	Schumpeterian incumbents and industry evolution	Buenstorf G.
#47	2016	Inherited competence and spin-off performance	Curran D., van Egeraat C., O'Gorman C.
#48	2016	"Opportunistic" spin-offs in the aftermath of an adverse corporate event	Curran D., O'Gorman C., van Egeraat C.
#49	2016	Spinoffs in Germany: characteristics, survival, and the role of their parents	Fackler D., Schnabel C., Schmucker A.

Paper	Year	Title	Authors
#50	2016	Pre-entry experience, technological complementarities, and the survival of de-novo entrants. Evidence from the US telecommunications industry	Fontana R., Malerba F., Marinoni A.
#51	2016	Spinoffs and their endowments: beyond knowledge inheritance theory	Furlan A., Grandinetti R.
#52	2016	Who lives longer? Startups vs spinoffs founded as proprietorships	Furlan A.
#53	2016	Spinoffs and clustering	Golman R., Klepper S.
#54	2016	Spinoff dynamics beyond clusters: pre-entry experience and firm survival in peripheral regions	Habersetzer A.
#55	2016	Dynastic entrepreneurship, entry, and non-compete enforcement	Rauch J.E.
#56	2016	Early stage cluster development: a manufacturers-led approach in the aircraft industry	Steenhuis H. J., Kiefer D.
#57	2016	Where do spinouts come from? The role of technology relatedness and institutional context	Yeganegi S., Laplume A.O., Dass P., Huynh C. L.
#58	2017	Bridging Knowledge Resources: The Location Choices of Spinouts	Adams P., Fontana R., Malerba F.
#59	2017	Structural and longitudinal analysis of the knowledge base on spin-off research	Ferreira M.P., Reis N.R., Paula R.M., Pinto C.F.
#60	2017	The dynamics of cluster entrepreneurship: Knowledge legacy from parents or agglomeration effects? The case of the Castellon ceramic tile district	Hervas Oliver J.-L., Lleo M., Cervello R.
#61	2017	Customers involvement and firm absorptive capacity in radical innovation: The case of technological spin-offs	Scaringella L., Miles R.E., Truong Y.

The time dimension of the phenomenon is represented in Fig. 3.1. The graph reports the disclosure of articles deemed relevant for the analysis. As it can be observed, the trend is quite increasing over time, confirming those studies that highlight the growing importance of the topic (e.g. Klepper, 2009; Ferreira et al., 2017). The exceptions to the pattern are years 2015 and 2017. While the decrease between 2014 and 2015 is not clear, the number of articles disclosed in 2017 may be lower as only articles disclosed up to July are taken into consideration.

Fig. 3. 1: Number of articles on spinoffs per year. Source: personal elaboration



3.3 Citation analysis: results

As reported in the previous chapter, the output of citation analysis is a matrix whose first row contains the set of 61 articles upon which the following analyses are performed and whose first column contains the set of documents that cite them. This matrix reports value 1 if a citing article reports the corresponding cited one in its references and 0 if not; as previously explained, this step of the procedure was performed by using the software R and by a manual normalization process for more reliable and accurate results. Citation matrix unveils important characteristics of the literature. For example, it can be used to assess which works are cornerstones of a specific literature and this can be inferred by counting the number of citations that they received by scholars. The most cited papers on spinoffs that resulted from the count on citation matrix are the ones displayed in Table 3.2.

Table 3. 2: Papers receiving the highest number of citations from Scopus articles. Source: personal elaboration

Paper	Title	Authors	Year	Citations
#5	The origin and growth of industry clusters: The making of Silicon Valley and Detroit	Klepper S.	2010	116
#15	The effectiveness of university knowledge spillovers: Performance differences between university spinoffs and corporate spinoffs	Wennberg K., Wiklund J., Wright M.	2011	72
#7	Entrepreneurial Origin, Technological Knowledge, and the Growth of Spin-Off Companies	Clarysse B., Wright M., Van de Velde E.	2011	67
#18	Home sweet home: Entrepreneurs' location choices and the performance of their ventures	Dahl M.S., Sorenson O.	2012	65
#4	Disagreements and intra-industry spinoffs	Klepper S., Thompson P.	2011	62

A further consideration on the citation matrix can be made: some of citing papers are also included in the set of cited documents, thus in the set of 61 spinoffs articles. Altogether, these citing articles (that are at the same time cited) are the ones that cite more the set of cited articles, as expected. As a matter of fact, cited articles were selected because they dealt with the topic of spinoffs and they are more likely to be cited by papers concerning the same or similar issues. Articles that cite only 1 or a few of these 61 spinoffs articles are instead less likely to contribute to spinoffs literature.

To delve into this issue, a different citation procedure was then performed: instead of counting the citations from all the citing documents retrieved by Scopus, the frequency of citations within the sample of articles on entrepreneurial spinoffs was counted. In other words, only the citations of papers that are both citing and cited were considered; this procedure entailed a sizeable reduction of the number of citations, as a subsample of citing papers was analysed -61 against 553-. This allowed the observation of relationships among articles, that are established by citation ties. Following this procedure, the most cited papers on spinoffs are the ones in Table 3.3.

Table 3. 3: Papers receiving the highest frequency of citation from the set of cited articles. Source: personal elaboration

Paper	Title	Authors	Year	Citations
#4	Disagreements and intra-industry spinoffs	Klepper S., Thompson P.	2010	23
#5	The origin and growth of industry clusters: The making of Silicon Valley and Detroit	Klepper S.	2010	18
#2	Why does entry cluster geographically? Evidence from the US tire industry	Buenstorf G., Klepper S.	2010	9
#21	Characteristics and performance of new firms and spinoffs in Sweden	Andersson M., Klepper S.	2013	7
#20	Employee spinoffs and other entrants: Stylized facts from Brazil	Muendler M.A., Rauch J.E., Tocoian, O.	2012	6

In the case of both Table 3.2 and Table 3.3, we should consider the time-related bias that makes newer articles less cited: Table 3.2 reports that the newest article with a high citation frequency was disclosed in 2012, whereas Table 3.3 reports an article published in 2013. The differences among the two citation tables concern not only the different hierarchy but also the different papers that have entered the list. The second table highlights the central role of Klepper's works for authors on spinoffs (4 out of the 5 most cited works are authored or co-authored by Klepper), although the influence of this author has been acknowledged also by other researchers, as shown in the first table (2 out of the 5 most cited works are authored or co-authored by Klepper). Furthermore, the second table does not include two papers concerning academic and entrepreneurial spinoffs with respect to the first table; this may highlight that the scientific community's interest on academic spinoffs is not wholly shared by the narrower community of entrepreneurial spinoffs. This latter deems as seminal papers those reported in the Table 3.3 and as main milestone the paper "Disagreements and intra-industry spinoffs", written by Klepper and Thompson in 2010.

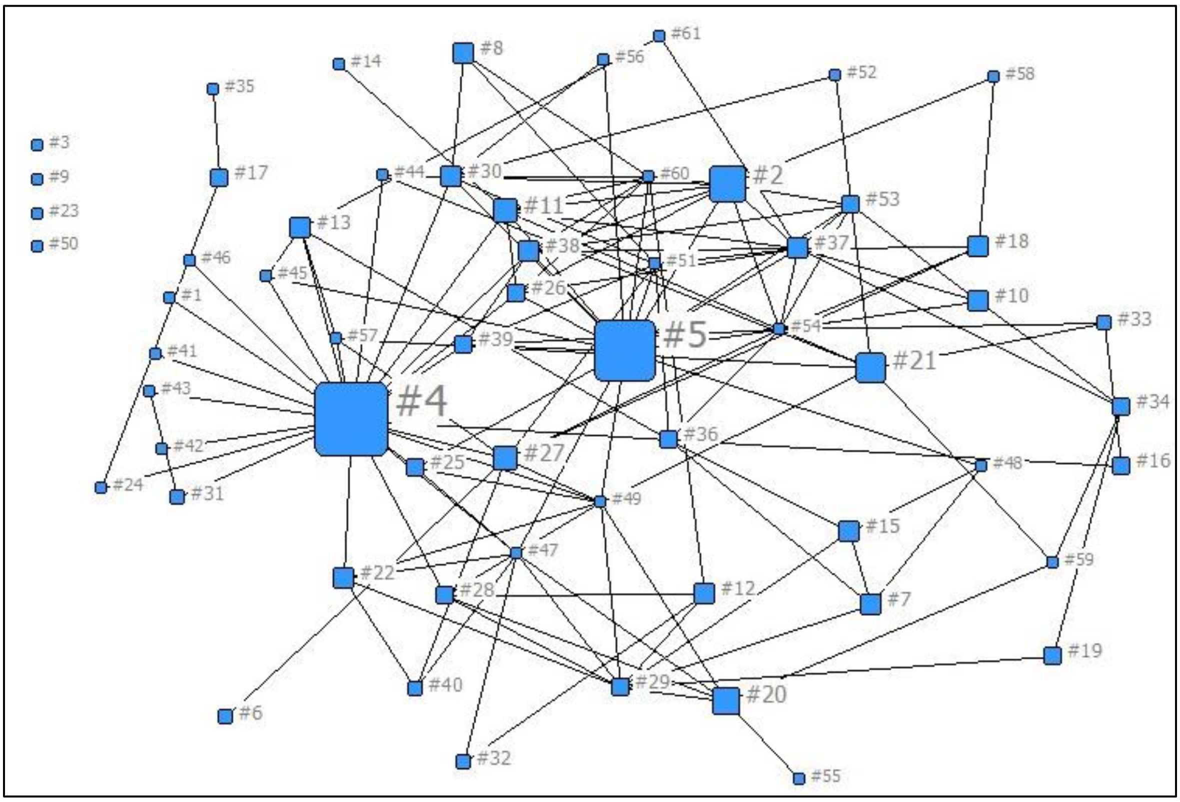
To provide a snapshot of the network of documents on spinoffs and the interlinkages among them, the software UCINET was employed. This software has found application in other bibliometric studies (see Jo et al., 2009; Ferreira et al., 2017) and it is probably the most used software for social network analysis; as a matter of fact, it contains software packages as NetDraw which allow the depiction of networks (Huisman and van Duijn, 2005).

Fig. 3.2 exhibits the network of the 61 source documents retrieved; the labels for each node contain the number of papers as reported in Table 3.1 and this is due to space constraints. Blue nodes represent documents and interlinkages among them represent a citation relationship, that is one document cites the one which it is tied with. For representing the citation relationships, we have considered the frequency with which a source document is cited by other papers that represent the set of cited documents (i.e. 61 papers on entrepreneurial spinoffs) for the considerations mentioned earlier on Table 3.3. In this network, the bigger the size of the nodes and that of the labels, the higher is the citation frequency of papers.

As it can be seen from Fig. 3.2, the biggest nodes represent papers reported in Table 3.3 and many interconnections with such nodes are visible. To reduce clutter, the heads of the arrows signalling the direction of citation relationships have been deselected, that is the edge between two nodes does not clarify which is the citing paper nor the cited.

An analysis which used the arrow heads was performed, albeit not reported in this work. It was possible to observe that the heads of the arrows were mainly oriented towards papers #4, #5, #2, #20 and #21 which, as a matter of fact, are the most cited works, as expected. Furthermore, four documents (#3, #9, #23, #50) are isolated, that is are not linked to the other articles on entrepreneurial spinoffs; this should require further investigation as they may be mistakenly been included in the sample of articles, or simply neither being cited nor citing any of the documents in the sample for other reasons.

Fig. 3. 2: Network of papers on spinoffs. Source: NetDraw (UCINET)



3.4 Co-citation analysis: results

Starting from 61 articles on spinoffs, 1891 pairs of documents were created by using n-1 submatrices (n= 61). The logic behind this procedure was to have each submatrix representative of each article, displaying as rows the citing papers and as columns the pairs of papers. Then, the sum of each column was reported in the cells of the raw co-citation matrix, a square symmetric matrix with identically ordered cited papers both in the first row and first column. Thereby, each cell of this matrix measures the interaction, i.e. the co-citation strength, between a couple of papers, as given by the sum of the corresponding submatrix that includes that pair. To understand this procedure, Fig. 3.3 reports a small part of submatrix 2 and the co-citation count of Paper #2 with Papers #4 and #10, as extracted from the results obtained. Citing Paper 22 cites jointly both Papers #2 and #4 and Papers#2 and #10 and this explains why the cells associated with Citing Paper 22 are filled with 1 for the two pairs considered. Co-citation count then, is the input to fill the cells of co-citation matrix, respectively equal to 6 for pair Papers #2 and #4 and equal to 3 for pair Papers #2 and #10. In this simplified case, co-citation count highlights a higher strength between papers of the first pair and thus a more thematic homogeneity between these ones with respect to the second pair.

In raw co-citation matrix, diagonal values were filled with AV, i.e. artificial value, as computed following the procedure reported in chapter 2.

Fig. 3. 3: Retrieval of co-citation matrix example from submatrix of papers. Source: personal elaboration

Sub-matrix #2	Papers #2 #4	Papers #2 #10
Citing Paper 1	0	0
…	0	0
Citing Paper 21	0	1
Citing Paper 22	1	1
…	0	0
Citing Paper 30	1	0
Citing Paper 31	1	0
Citing Paper 32	1	0
Citing Paper 33	0	1
…	0	0
Citing Paper 38	1	0
….	0	0
Citing Paper 42	1	0
…	0	0
Co-citation count	**6**	**3**

Co-citation matrix

	Paper 1	Paper 2	Paper …	Paper 4	Paper …	Paper 10	Paper …
Paper 1	AV	…	…	…	…	…	…
Paper 2	…	AV	…	**6**	…	**3**	…
Paper …	…	…	AV	…	…	…	…
Paper 4	…	**6**	…	AV	…	…	…
Paper …	…	…	…	…	AV	…	…
Paper 10	…	**3**	…	…	…	AV	…
Paper …	…	…	…	…	…	…	AV

In raw co-citation matrix[4], papers were labelled by using the numbers of Table 3.1 to identify such articles. The diagonal values were computed following the artificial value approach mentioned earlier, that is summing the highest free frequencies for a specific paper and dividing this number by two; space constraints forced to report diagonal values up to the first digit after the comma.

Pairs of frequently co-cited documents were, in decreasing order: Paper #7-Paper #15 (19 as co-citation value); Paper #2- Paper #5 (14); Paper #4-Paper #5 (13); Paper #5- Paper #18 (8). As stated before, these high frequencies of co-citation could reveal similarity of contents between elements of the pair. Raw co-citation matrix can also reveal, at first glance, which documents have been co-cited more with other articles. Paper #5, for instance, have been co-cited nearly with all the other papers, whereas Paper #3 has been co-cited only with other 3 articles: the difference between these two works cannot be ascribed to a different publication date, as both were disclosed in 2010. Rather, it may be inferred that either Paper #5 deals with farther-reaching topics than Paper #3 or articles included in this sample are more content-similar to Paper #5. In any case, this cannot be assessed by looking at frequency numbers on the matrix; further investigation is required, also by means of other analyses.

As stated in the previous chapter, only articles whose co-citation threshold was higher than or equal to 3 were retained and this explains why raw co-citation matrix is a 32 by 32 matrix instead of 61 by 61. Based on this raw co-citation matrix, Pearson correlation matrix was computed and used as input for the following analyses, namely cluster analysis and

[4] Raw co-citation matrix is available at Appendix A

multidimensional scaling, while for Principal components analysis the input used was the raw co-citation matrix.

3.5 Quantitative analyses

For systematizing spinoffs' literature starting from raw co-citation matrix and Pearson correlation matrix[5] PCA, cluster analysis and MDS were performed by using the software UCINET and SPSS.

3.5.1 Principal components analysis: results

In our study, PCA was used for detecting the subfields of research on entrepreneurial spinoffs topic and which articles identify these specialties. The software UCINET was used for this analysis and some procedures were followed.

First, principal components analysis (default option provided by UCINET when conducting a Factor Analysis) was chosen as method. Table 3.4 depicts the output of principal components extracted through a Varimax rotation and with a minimum eigenvalue equal to 1 by using UCINET. Choosing as stopping rule the criterion of eigenvalues higher than or equal to 1 works better when the number of variables is between 20 and 50 (Hair et al., 2010); as in our analyses there were 32 variables, we expected that this criterion would work quite well.

Table 3. 4: Principal components. Source: UCINET and personal elaboration

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Paper #20	*Paper #2*	*Paper #28*	*Paper #3*	Paper #4	*Paper #12*	*Paper #8*	*Paper #42*
Paper #22	*Paper #5*	*Paper #34*	*Paper #7*	*Paper #13*	*Paper #19*	*Paper #31*	
Paper #26	*Paper #10*	*Paper #36*	*Paper #15*	Paper #16	*Paper #29*		
Paper #30	Paper #11	*Paper #37*		Paper #21			
Paper #33	Paper #18	*Paper #39*					
Paper #40	*Paper #24*						
	Paper #27						
	Paper #38						

Each column of PCA output represents a specific rotated component and lists the set of papers with a significant loading (> |0.4|, see paragraph 2.7.1), coherently with other studies (e.g. McCain, 1990; Nerur et al., 2008). Even though the number of components to be extracted was equal to 10, the output resulted in a lower number (8).

Loadings which contribute very significantly to the components (higher than 0.7) are reported in italics. Most of the articles considered (27 out of 32) had a significant loading to a specific

[5] Pearson correlation matrix is available at Appendix B

component and this makes us infer that such papers are relevant to understand the paradigm of research which they belong to.

For the sake of simplicity, we reported papers' contribution only to the specific component where the highest loading is observed (e.g. Ferreira et al., 2017), despite some papers reported a loading higher than |0.4| in more than one component. We deemed more useful to use this approach instead of reporting all the significant contribution of articles to the different components -or subfields of research- for a couple of reasons: first, it provided wieldy solutions that could be easily matched with other analyses' outputs; second, articles whose contribution was higher for a specific component were more likely to belong to a specific cluster that component is referred to.

A further consideration can be made concerning the number of articles that constitute each component. Nearly 70% of articles on spinoffs constitute the first four components and more than 81% of them constitute the first five. This can be a helpful indicator for considering the most relevant subfields of the literature.

In PCA the choice on the number of components to be retained is dependent on the amount of variance that each of them can explain (see chapter 2). Thus, the amount of variance is a helpful indicator for choosing how many components on spinoffs should be retained. The software SPSS provides such information. By performing PCA through SPSS, the outputs concerning principal components loading vectors were the same as the ones found by using UCINET: each principal component loading vector is in fact unique, although two different software packages can provide different signs for the values of such vectors (James et al., 2013). The advantage when using SPSS 22.0 is the correspondence between each component and the amount of the variance that it can explain.

Table 3.5 reports the total variance explained by the 8 components extracted as output of the analysis performed by SPSS 22.0. The total variance explained by the 8 components is equal to 87.28%. As we can see, the first 3 components can explain half of the variance and the first 5 components can explain almost 70%, that is a usual threshold in these studies. Since the variance explained by each component highlights the contribution of that specific component to the development of spinoffs topic, we can gather that the very first components are quite remarkable to explain this field.

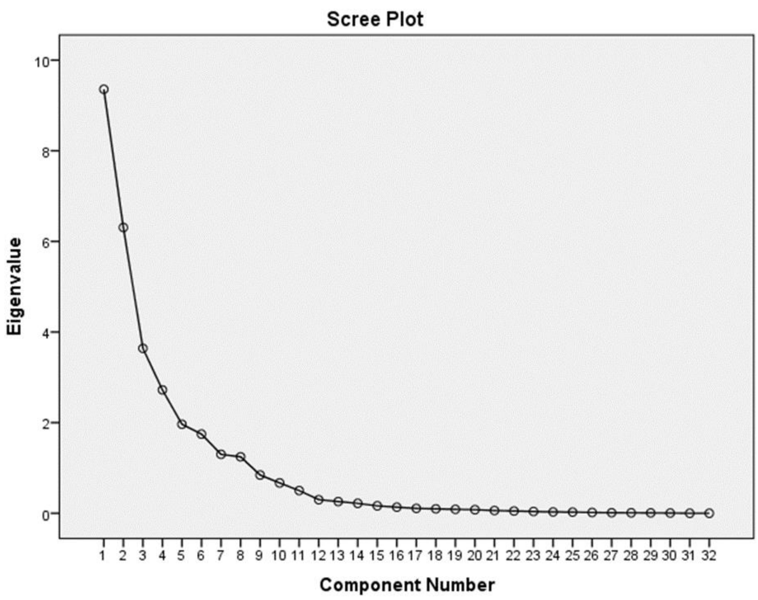
Table 3. 5: Variance explained by principal components. Source: personal elaboration

Components	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
% Variance	17.728	17.108	14.648	9.969	8.542	7.912	7.905	4.572
% Cumulative Variance	17.728	34.836	49.484	59.453	67.994	75.907	83.812	88.384

As stated before, principal components in co-citation studies embody subfields of research topics and, as a matter of fact, our aim is to identify clusters or subgroups which the literature on entrepreneurial spinoffs is composed by.

The decision to establish an eigenvalue higher than 1 or to look at the cumulative variance are two of the most used criteria for settling on the number of components. A further method is the *scree test criterion* which instead relies on the eyeball. This criterion is based on the interpretation of the curve in a graph plotting the eigenvalue against the component number (Hair et al., 2010). Very often in this kind of graph there is a break in the curve and this represents the point from which the slope of the curve becomes close to 0, that is the point from which the scree -the flat portion- starts. The number of components to be extracted is the one before the scree, thus the one before the elbow, albeit several breaks can result. When this situation occurs, the number of components corresponding to the first break should be considered. Nonetheless, this has some limitations, such as the fact that usually researchers prefer to retain at least 3 components -and the break may appear before the third component- and that the scree can appear after the second or third break (Norman and Streiner, 2008).

Fig. 3. 4: Scree plot- eigenvalues and components. Source: SPSS 22.0



In our case, the scree plot resulting from SPSS 22.0 is the one depicted by Fig. 3.4.

The scree plot signals different inflection points, as sometimes happens in this kind of analysis. However, the point in which the downward trend of the curve becomes smoother corresponds to the sixth principal component. After the eighth component then, eigenvalues are lower than 1 and further components cannot be retained according to the latent root (eigenvalue= 1) criterion; we may conclude that 6 principal components, that is subfields of entrepreneurial spinoffs topic, may be identified.

As expected, a purely visual inspection of the graph cannot produce clear results and the identification of thematic specialties requires further investigations.

3.5.2 Cluster analysis: results

To graphically depict groups, a hierarchical agglomerative cluster analysis by using UCINET was performed.

The employment of complete and weighted average linkages of UCINET described in chapter 2 did not provide different results, except for some papers, whose membership was different, depending on the method. Those papers may be barely embedded in a specific group and this is the reason why their membership spans different clusters.

UCINET visual outputs are by default an *icicle plot* and a *dendrogram*. Fig 3.5 reports the dendrogram where the lower part of the tree is located on the right and the leaves are on the left; the number of the corresponding paper are located next to each leaf. As mentioned earlier, finding the right number of clusters following a hierarchical clustering approach is not trivial; *a fortiori* in this case, it would be quite difficult to apply the method of the longer vertical branches to extract the number of clusters.

Fig. 3. 5: Dendrogram and identification of clusters. Source: UCINET and personal elaboration

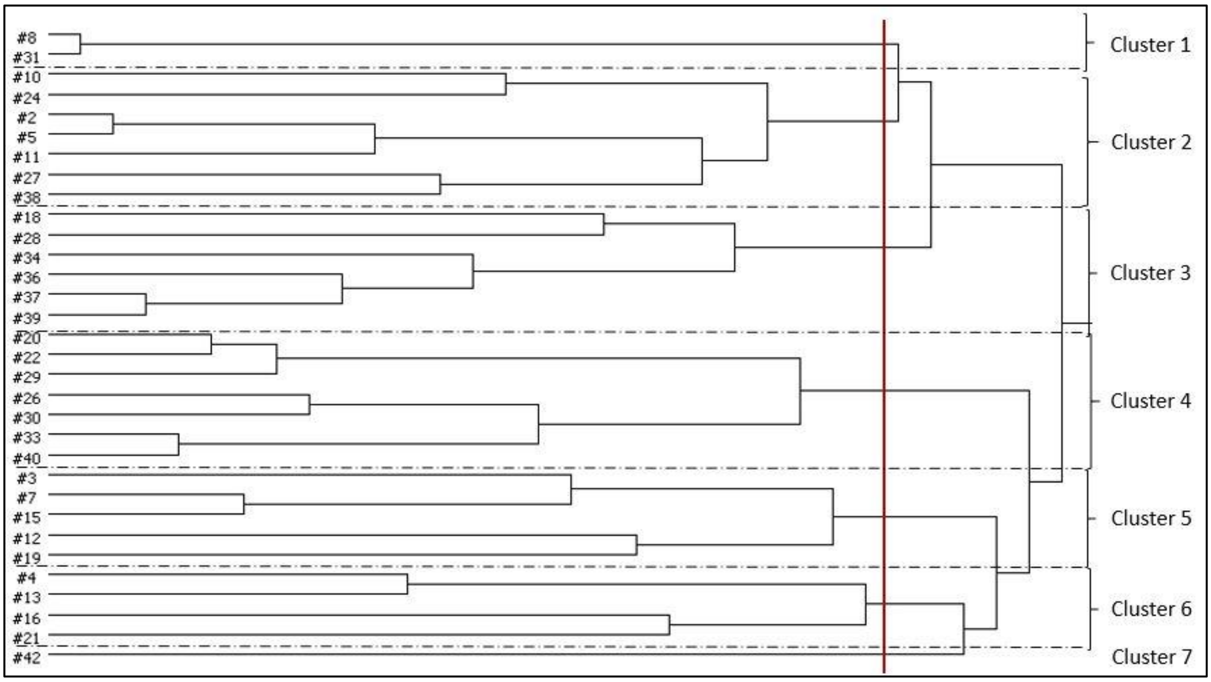


Fig. 3.5 shows the dendrogram and the groups obtained: the thicker red line signals the level where the number of clusters can be computed. The line is reasonably put on the right area of the map since identifying the cut-off point in the in the middle or in the left part of the dendrogram would mean finding nearly as many clusters as the number of observations that is quite meaningless for our purpose.

Looking at the right part of Fig. 3.5, instead, it is quite evident an immediate partition in two broad subgroups which are suddenly parted and their branches are in turn parted again quite soon. At first glance, this process stops at the level in which the red line is put: on a general assessment base, the branches intersected by the line take longer to be parted again and this is reasonably the point where the line can be put.

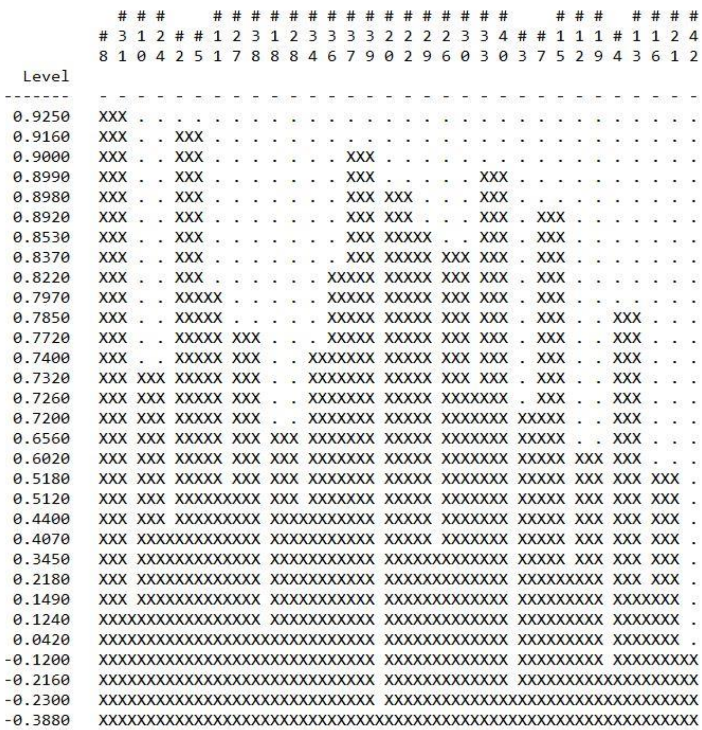
The red line has identified 7 clusters, and while some of them count several units, others like clusters 1 and 6 count a very few papers or even only one as in the case of cluster 7. The different number of units inside each cluster may indicate which thematic groups on entrepreneurial spinoffs have been more deeply investigated, as clusters 2, 3 and 4.

It is also worth noting that the level of similarity between papers in a pair is represented by the proximity of the node to the leaves. Papers #8 and #31 show the highest couple similarity, whereas Papers #16 and #21 the lowest one. By looking at the length of the branches we can infer also the level of heterogeneity inside each cluster: for example, Cluster 4 seems to be formed by two thematic subgroups, one including papers #20, #22 and #29 and another including #26, #30, #33 and #40 and this reasoning can be applied to other groups.

The second output of UCINET is the *icicle plot* (Fig. 3.6): the upper part of this graph reports the number of the papers and the symbol *X* represents a paper that is being embedded within an existing cluster which provides an indication of relative similarity.

For instance, Paper #8 and Paper #31 are the first to be joined in a cluster thereby confirming dendrogram output; Paper #2 and #5 are the second couple to form the second cluster; this two-papers clustering continues until the seventh step, in which Paper #29 is embedded within the cluster previously formed by Papers #20 and #22. The icicle plot can be inversely read: the lower level of the picture is the situation in which all the papers are grouped within a single cluster, while the next to last level shows the first partition in two big subgroups. Furthermore, UCINET provides an additional explanation to the icicle plot, through

Fig. 3. 6: Icicle plot and level of similarity. Source: UCINET



the values reported on the left column “Level”: the levels indicate the degrees of association, in terms of similarity, among the elements within the clusters, for each grouping sub-process, that is at every stage of the clusters formation. In our case, by using similarity data and a complete linkage among clusters, each level value represents the extent to which an item in a cluster is at least similar to other items in the cluster, measured in terms of units.

Neither dendrogram nor icicle plot can help to identify the exact number of clusters, but when jointly read they provide the similarity between papers and an indication of the underlying clustering process among them. For the sake of simplicity, two examples from the two graphs are considered: a) couple #8 and #31; b) couple #4 and #13. Papers in case a) are the very first to join and therefore the similarity between them is very high. Papers in case b) instead are the first to join in a cluster different from the ones created earlier by other papers -6 subgroups are created before Paper #4 and #13 join-; even though they are the first couple to join in the subgroup which they finally belong to, the similarity between them is lower than case a) and lower than other couples of papers which joined to form other clusters or fuse to an existing cluster (as in the case of Paper #29 which joined the subgroup on the middle of the icicle plot). This further proves that the different clusters identified can embed items with different similarity that signal the similarity of the cluster itself.

As mentioned earlier, deciding the height of intersection in the dendrogram and the level of clustering in icicle plot implicitly means establishing the level of thematic specificity of the clusters; sometimes, an approach based on heterogeneity among the items of a cluster is applied to identify the different subgroups to back up analysts' discretion. This procedure relies on the definition of a measure of heterogeneity and postulates that cut-off points correspond to the levels where the highest increases in the heterogeneity among clusters' items occur (Hair et al., 2010). In our analysis, an example of heterogeneity is provided by the level of association in the icicle plot next to each grouping sub-process. The maximum decrease in the level of similarity, that is the maximum increase in the level of heterogeneity, is recorded between the fifth to last (value equal to 0.042) and the fourth to last (-0.12) sub-processes. Other significant decreases occur between the second to last and last process and between sixth to last and fifth to last. Dissimilar clusters are joined at these stages and they represent opportunities to select the appropriate number of clusters; more specifically, the third significant decrease that we have identified provides the same results highlighted by the red line in the dendrogram and therefore 7 groups -even though the seventh cluster embeds only one unit and thus it can barely be defined as a cluster-. Moreover, the decrease in the level of similarity provided by the icicle plot and the location of papers in the dendrogram can contribute to signal which clusters are more similar. By looking at Fig. 3.5 for example, we can infer that cluster 1 is more similar to cluster 2 than cluster 6. Therefore, cluster analysis is helpful insofar it allows units to be grouped because of shared characteristics but it also permits the detection of inter-cluster similarity relationships.

If the classification of clusters is successful, objects belonging to the same cluster are expected to group together when plotted graphically. Still, mapping the intellectual field allows more interpretations related to analysts' judgement of papers' contents; the cut-off level therefore is context-based depending on analysts' discretion, which can decide to provide broad clusters or more specific ones. A general overview of the field is however achieved when results from all the quantitative analyses performed are available. Thus, our hypothesis on a 6 or 7-cluster-based partition should be confirmed and complemented by other analyses, that can yield interesting and helpful insights.

3.5.3 MDS: results

Multidimensional scaling has an utmost importance to detect groups at a glance. Following the procedure outlined in chapter 2, several attempts were made to identify the fittest dimensionality of the plot. The stress values obtained by the different attempts formed the curve in Fig. 3.7 and are reported in Table 3.6.

Fig. 3. 7: Stress value per number of dimension.
Source: personal elaboration

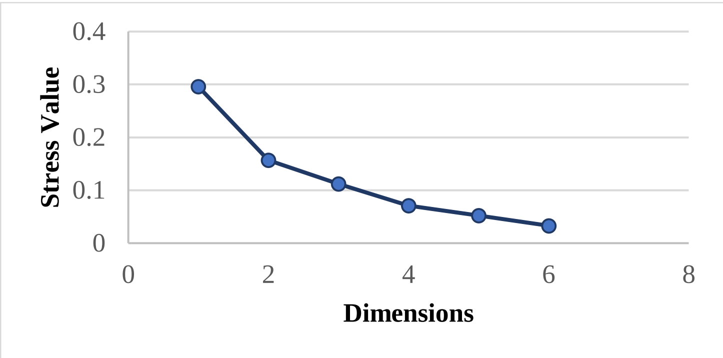


Table 3. 6: Stress values.
Source: personal elaboration

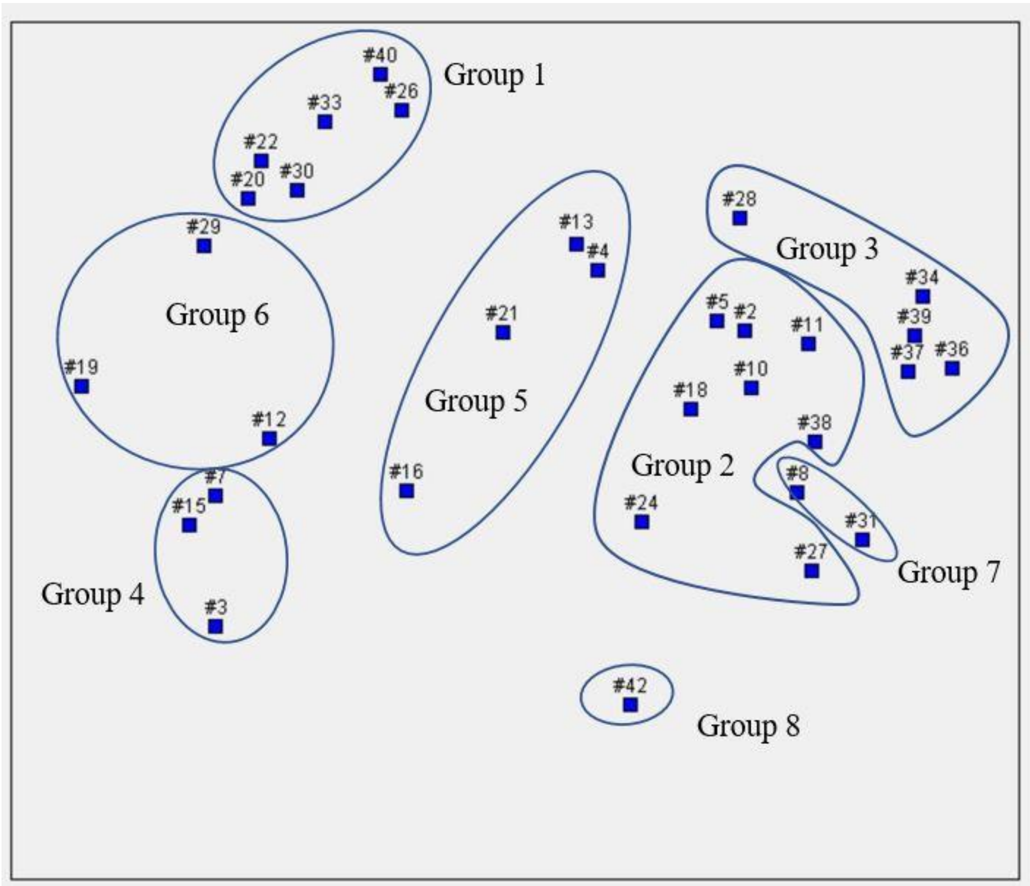
Dimensions	Stress
1	0.296
2	**0.157**
3	0.112
4	0.071
5	0.052
6	0.033

In our case, the improvement in the goodness-of-fit has been reached quite soon, namely it corresponded to a two-dimensional choice, in which the stress value is deemed acceptable. The decrease in the stress value is in fact equal to 0.14 when passing from a one dimension (in which the stress value is unacceptable) to two dimensions; the curve of the stress value becomes flatter with increasing values of dimensionality, as expected: more dimensions reduce the stress. The decision to select two dimensions for the final MDS output was *a fortiori* supported by the spread employment of a two-dimensional solution in several co-citation studies (two dimensions enable the explanation of 85% or more of the variance) and by the fact that the explanatory power increased by an added dimension is offset by an extreme complexity in interpreting that added dimension (McCain, 1990). In our analysis, a stress value equal to 0.157 obtained in 17 iterations was achieved with two dimensions and, as mentioned earlier, it is considered acceptable according to Kruskal (1964) rule of thumb.

McCain (1990) asserts that stress values in co-citation data tend to be higher due to the inherent noisy nature of co-cited data; when stress value is quite high (but lower than 0.2) it is an acceptable trade-off for a two or three-dimensional configuration. Two outputs from UCINET were obtained: the first file was a scatter plot (Fig. 3.8) in which the points representing the papers were plotted on a two-dimensional space; the second file reported the coordinates of the points on the map and the measure of the stress achieved after *n* iterations.

Fig. 3.8 maps the structure of the knowledge on entrepreneurial spinoffs after 2009, by considering the correlation matrix as input. Correlation's employment was emphasized by the relative location of the items: papers that exhibited a higher correlation in the Pearson matrix were closer (such as Papers #2 and #5, whose correlation coefficient is equal to 0.92) while articles with a lower correlation were farther (such as Papers #3 and #28, whose correlation is equal to -0.31).

Fig. 3. 8: MDS and groups partition. Source: UCINET and personal elaboration



Documents' groups were identified with thicker manually drawn lines, depending on their local proximities and the subfields partition provided by PCA. It is quite clear that the partition is similar to the one found by cluster analysis (Fig. 3.5): Group #1 includes most of the documents loading on principal component #1 and belonging to cluster #4; Group #2 embeds most of the documents contributing to principal component #2 and cluster #2; Group #3 is formed mainly by documents loading on principal component #3 and cluster #3; Group #4 includes documents of principal component #4 and of cluster #5; Group #5 is made of documents of cluster #7 and of principal component #5; Group #6 is formed by documents loading on principal component #6 and some belonging to cluster #5; Group #7 is formed by documents contributing to principal component #7 and cluster #1; Group #8 is composed by a single document that is Paper #42, which forms both principal component #8 and cluster #7.

By observing Fig. 3.8, some interesting conclusions may be drawn:

- Behaviour shown by Paper #42, which fails to be grouped within any sub-partition, usually occurs when a document is co-cited infrequently with other ones (Culnan, 1986). This may be due to its later disclosure (in our case 2016) but also to different thematic focus compared to the other papers analysed.
- Some documents are placed in a relatively central position: this is the case of Papers #4, #13 and #21. Their location on the map indicates their far-reaching influence over other strands of research on entrepreneurial spinoffs and therefore these papers may deal with theories, evidence or models shared by different thematic groups; the inverse logic may be applied to the most peripheric areas: the most peripheric papers are likely to deal with topics unrelated to the other issues.
- Some articles are close to the boundaries of the groups, such as papers #7, #12, #20 and #29; it may be inferred that these works deal with issues featuring more than one group and thus are barely concerning with the focal group theme. It is not by chance that article #29 is included with articles of cluster #4 even though it loads with other papers to principal component #6 and therefore its thematic specificity may not be clear.
- Papers location on the map is uneven: for instance, the middle right side of the map swarms with many papers, whereas the lower side is scarcely populated.
- Papers location within the same group is uneven: some articles tend to be closer depending on a higher correlation, while some others farther: it stands to reason to envision that themes of closer articles are more similar. This pattern is well-rendered by Group #2, characterized by a relatively large distance between the subgroup formed by articles #24 and #27 and the one formed by all the other articles. Another example is provided by Group #6, whose papers' positioning is anything but concentrated: articles #12 and #19 are closer to Group #4 than to paper #29, although this latter belongs to the same group, according to PCA.
- The map reveals several empty spaces, as in the case of the lower left area: empty areas can highlight themes neither studied nor investigated yet which can contribute to the knowledge development (Di Guardo and Harrigan, 2012).

- Paper #3 was included within a group of articles, albeit citation analysis did not support any relationship with other papers (see paragraph 3.2). Since paper #3 was disclosed in 2010, it stands to reason to assume that it lacks in citing the other 34 papers. At the same time, its proximity to other articles within the cluster may suggest that other citing documents have acknowledged a thematic similarity among articles of group #4. Therefore, citation analysis *per se* may lead to erroneous interpretations and it should necessarily be complemented by further analyses.
- Despite its simplicity, the aforementioned method based on the univocal membership of one article to a single principal component may trigger incomplete results. As a matter of fact, some articles contribute to more than one principal component and therefore can belong to more subfields. This is the case of articles #16 (contributing also to component #4), #20 (loading also on component #6), #21 (loading also on component #3) and #29 (contributing to component #1).
- Article #4 deserves an in-depth scrutiny: as mentioned earlier, its central position emphasizes its broad influence on the literature. This conclusion is further supported by the results of a complete PCA: paper #4, indeed, contributes quite significantly to principal components #1, #2 and #5. Its pivotal role is also gauged by the results provided by citation analysis, as it resulted to be the most cited paper. At the same, it fails to contribute very significantly to a single component (its highest load is lower than |0.7|) and thus to be a cornerstone for a single subfield.

3.5.4 Labelling subfields: results

PCA allows the retrieval of a specific number of components that signal the number of subfields of a specific topic. A word frequency analysis on keywords of titles and abstracts of all the 32 papers on spinoffs was conducted by using the software VOSViewer, whose output is shown in Fig. 3.9.

Fig. 3. 9: Keywords from word frequency analysis. Source: VOSViewer

The keywords resulting from the automatic procedure seem to represent the different broad topics of entrepreneurial spinoffs literature: “corporate spin-off”, “university” “research and “type” may refer to the different types of spinoffs and their implications; “growth”, “survival” and “performance” may concern with economic results of spinoffs; “cluster” “region” and “agglomeration economy” may be related to the location choice of these new ventures; lastly, “knowledge” and “prior experience” may be related to the background of spinoffs formation. VOSViewer maps of keywords are created so that words from the same articles tend to be nearer, while farther words indicate different sources from which these terms are taken (Van Eck and Waltman, 2014). The map therefore may imply that articles concerning the typologies of different spinoffs deal also with spinoffs performance and prior experience (left area of the map); at the same time, geographic locations are treated jointly with spinoffs performance and previous experience (see the right area of the map); finally, geographic locations seem not to be related with the types of spinoffs, thus articles concerning location choices of spinoffs do not deal with different types of spinoffs at the same time.

It stands to reason also to apply the word frequency analysis to each specific component settled on by PCA. In this case, the word frequency analysis applied to each component stands out the specific theme dealt with by articles forming a specific component. Thus, by analysing subsets of articles referred to different components, the words that occur the most in the subset may communicate the thematic areas that distinguish a specific principal component from other ones.

Table 3.7 reports some of the keywords occurring the most for each subset of articles, by following the same procedure applied for the whole set of articles using VOSViewer.

Table 3. 7: Keywords with highest frequency per principal component. Source: personal elaboration

Principal Component	Keywords	Label
1	“type” “spinoff” “performance”	Type of Spinoffs
2	“cluster” “region” “agglomeration economy” “performance” “spinoff”	Cluster and Agglomeration economy
3	“prior experience” “knowledge” “performance” “growth” “spinoff”	Spinoffs Characteristics and Performance
4	“knowledge” “university” “firm”	Academic and Entrepreneurial spinoffs
5	“spinoff formation” “disagreement” “parent” “performance” “employee”	Motives for Spinoffs Formation
6	“firm” “innovation” “change”	Innovation
7	“network” “spinoff” “social capital” “knowledge transfer”	Network and Social Capital
8	“founder” “team” “performance”	Founding Team

The third column of the table reports a label inferred from the keywords analysis and from a preliminary reading of articles according to their contribution to principal components. What appears is that some keywords like "performance" and "knowledge" tend to be used in different subfields of research, while some others like "cluster" and "university" are peculiar of a specific subset of articles. Apparently, we may infer that there are central themes shared by different subtopics and at the same time there are different issues worthy to be independently studied in entrepreneurial spinoffs literature.

Despite these automatic results, we should take into account that keywords from titles and abstract may lead to some biases and therefore the themes of different subfields can be detected only through an accurate reading of all the papers.

3.6 Discussion

3.6.1 Subfields identification

PCA resulted in 8 components, thus 8 subfields, with decreasing explanatory power of the data and fewer articles contributing to each successive subdomain. Our opinion is in fact that 8 is a quite large number for representing the subfields of entrepreneurial spinoffs topic given its novelty and perhaps these 8 subdomains can be merged in broader subfields. Nonetheless, labels were created for each principal component, according to the content of each article and word frequency analysis; as a matter of fact, a component is given by a descriptive theme name that is founded on the evaluation of the areas created by the whole set of papers (concepts) loading on that component (Lee and Chen, 2011).

- Articles contributing to principal component #1- and therefore belonging to the first subfield of research- deal with the comparison among *different types of spinoffs*. Employee spinoffs have proved to survive longer than other business entrants -but with the same frequency of diversification divestitures of existing businesses- and to be larger at entry; firms' productivity and riskiness is inherited by employees and thus transferrable (Muendler et al., 2012). Further, other studies (e.g. Bruneel et al., 2013) distinguish among incumbent-backed spinoffs -thus spinoffs created by parent companies-, opportunity spinoffs -resulting from the exploitation of an opportunity discovered when employees work at the parent companies- and necessity spinoffs -resulting from and adverse situation occurred at the parent company. Opportunity spinoffs are outperformers with respect to other spinoffs categories since the motivational background seem to be a powerful weapon for the success, even though Rocha et al. (2015) demonstrate that pushed (read: necessity) spinoffs can be superior performers than pulled (read: opportunity) spinoffs when an appropriate human capital is endowed in the start-up phase. The different spinoffs origins are investigated through a model provided by Cordes et al. (2014) who posit that business environment, business culture and intra-organisational process learning define different types of spinoffs and are mechanisms to trigger the market performance: a dynamic business environment, a cooperation culture and the entrepreneur's influence in workers' socialization are conducive to outperformer spinoffs. The different origin of spinoffs does not *per se* trigger a higher performance in terms of innovation: the local proximity to research facilities and a frequent collaboration with them mediate the impact of spinoff origin on performance (Lejpras, 2014). Finally, Fryges and Wright (2014) provide a framework to identify all the types of spinoffs, according to their commercial or university context origin and if they form a new firm or are derivative start-ups from an existing activity; they find that both external (environmental) factors and internal (i.e. parental) ones contribute to spinoffs' different performance.

Thus, the motivational aspect related to the formation of the spinoff and the parent background are the main aspects which this group of papers deals with. It is worth pointing out that this partition was obtained also by cluster analysis and MDS; nonetheless, from dendrogram it is quite clear that a modest heterogeneity exists among these papers (the length of the branches is quite long): some of these papers investigate the role of human capital and innovation activities, while others deal with knowledge transfer and the birth of new firms.

- Principal component #2 articles concern with *cluster* and the role of *agglomeration economies* on spinoffs. Agglomeration economies (or Marshallian externalities) are defined as economies of scale external to the firm and related to the local system favouring the agglomeration of entities and consist of lower production costs due to physical proximity and regional knowledge spillovers (Cusmano et al., 2014). New ventures perform better when entrepreneurs locate in the area where they have lived longer, as they could rely on the appropriate capital and personnel; new ventures thrive when founders have prior industry experience and locate in known areas (Dahl and Sorenson, 2012). Discerning the effect of agglomeration economies and of previous industry experience on performance is instead a tougher task, even though the evidence is prone to the superiority of the latter determinant. Questions have arisen also concerning the offspring of leading companies in Detroit area and Silicon Valley, thereby inferring the role of specific locations on spinoffs growth rate; however, organizational reproduction and heredity seem to be more responsible than agglomeration externalities for the development of clusters (Buenstorf and Klepper, 2010; Klepper, 2010; Mayer, 2013) and this explains why new founders tend to establish near to flourishing parents (Klepper, 2011). The agglomeration of spinoffs explains also the empirical pattern of higher mobility of workers within Silicon Valley area even before semiconductor industry clustered there (Cheyre et al., 2015a). By studying Amsterdam cluster, Heebels and Boschma (2011) find that this area was an incubator for many new firms that relocated then to other regions.

These articles inspect the dyad location-performance of spinoffs and they are mainly focussed on this category of new entrants as opposed to new ventures founded by non-experienced entrepreneurs. From the dendrogram it may be inferred that in this cluster some heterogeneity exists as well: some articles deal more with performance, some others with network, some others with employee mobility issue. It is also worth noting the proximity in MDS map of this group of articles to the third and seventh groups of papers.

- Principal component #3 articles aim at presenting *characteristics* of *spinoffs* and their superior *performance* with respect to other entrants. While studies have always focussed on spinoffs performance given benevolent or neutral parents, Walter et al. (2014) examine whether different parental attitudes can affect spinoffs performance whereby attitudes range from friendly to hostile. The results show that spinoffs will exploit industry-specific experience advantages if parents are friendly or neutral, whereas if parents establish sanctions and therefore a conflictual relationship with the progeny, spinoffs will take longer to break-even. Nevertheless, spinoffs can soothe these negative consequences by creating a network of new ties with customers, competitors and suppliers different from parental ones. Moreover, the empirical evidence shows that hostile attitude is not rare: the first IBM spinoffs have been called "the dirty dozen" (McKendrick et al., 2009).

  Knowledge lacks also on other characteristics that distinguish spinoffs from other categories of new entrants. By employing data from Denmark, Dahl and Sorenson (2014) show that spinoffs entrepreneurs have less managerial experience but more experience both in the same industry and in related ones; their former workplaces are younger, smaller and more profitable firms; they tend to hire more former colleagues and less family members; they are less likely to be previously unemployed before founding a new business. Moreover, the industry which spinoffs and their parents belong to plays a critical role in determining the new businesses performance: service spinoffs gain a superior value added with respect to all the other new entrants irrespective of the location, whereas manufacturing spinoffs perform better in metro cities; in general, spinoffs show higher value added than other entrants in all the locations (Lööf and Nabavi, 2014). An analysis on the long-term performance of new firms in Silicon Valley reveals that the bulk of spinoffs that become top leaders come from top incumbents and their location choices do not affect the probability to become outperformers (Cheyre et al., 2015b). Finally, spinoffs performance in manufacturing industries seem to be less affected by local buzz and networks typical of creative and project-based sectors (Boschma, 2015).

Papers belonging to this group are mostly empirically oriented and analyse patterns, especially in clustered areas, to infer the dynamics of spinoffs' success. This is not surprising: in MDS, in fact, papers belonging to principal components #2 and #3 are very close; some articles contribute significantly to both the components and for this reason we may posit a fair interdependence among these two subfields.

- Articles contributing to principal component #4 concern with the logics behind *entrepreneurial* and *academic spinoffs* and their *performance*. Both academic and entrepreneurial spinoffs are vehicles to exploit the knowledge resulting from new discoveries taking place at the parent companies. Nevertheless, academic and entrepreneurial spinoffs need different technological endowments in the start-up phase: while university spinoffs can benefit from a broader technology that allows a wider array of applications as the knowledge of the market can be quite scarce, entrepreneurial spinoffs can rely on a narrower scope of the technology as the market knowledge is deeper. Furthermore, university spinoffs risk to achieve an excessive level of new technologies as result of their exploration activities and this translate in lower economic performances that can be upended by technology transfer offices (Clarysse et al., 2011). While determinants for the creation and success of entrepreneurial spinoffs are the establishment of sectors with high technological opportunities and endowments of science parks and incubators for new business ideas, universities should attract and retain eminent scholars, provide leading Phd programs and drift towards a more entrepreneurial climate (Gilsing et al., 2010). Universities have an appropriate role in knowledge spillovers both for entrepreneurial founders and for academic founders; entrepreneurial spinoffs seem to outperform academic ones and founders of university spinoffs can fill this gap by hiring personnel with industry experience; lastly, parent companies are more likely to shape entrepreneurial spinoffs economic performance than academic one (Wennberg et al., 2011).

  In MDS map, this group of papers is located near to group #6 which deals with innovation topic, as Figure 3.8 shows.
- Articles contributing to subfield #5 cover the *motives* behind *spinoffs formation*, also by accounting for some parent firms' characteristics that can affect spawning rate. As mentioned in chapter 1, Klepper and Thompson (2010) posit the "theory of disagreements" to explain the formation of spinoffs: these new ventures emerge when a misalignment among leading decision makers concerning basic ideas of strategy and technology causes dissidents' leaving and foundation of their new firms. This model contributes also to explain why

organizational changes are more likely to be followed by a higher spawning rate and it postulates the limitation in evaluating novel ideas by the parent firms. Thompson and Chen (2011) instead, assert the existence of two types of spinoffs: the first type is created when an employee deems worthwhile the adoption of a new technology but the parent firm refuses to implement it, whereas the second type is established when the firm considers the adoption of a novel technology but some dissidents prefer to carry on with the traditional technology path. Moreover, when the parent firm engages persistently in R&D activities and pays higher wages and has higher sales and exports, it is less likely to spawn: this can be explained by the opportunity cost that employees suffer if they start a new business but also to the fact that accumulated knowledge is embedded within physical capital (Andersson et al., 2012) and therefore previous studies are confirmed (see paragraph 1.5.3). Furthermore, spinoffs in Sweden are less likely to be spawned by larger firms (in particular MNEs) and breed when economic environment thrives and this is in line with evidence in other countries; fiscal and labour market policies tend also to affect the formation of spinoffs and this might explain why a significant number of these *de novo* firms have been established since 1991, in response to the major tax reform in Sweden; finally, tenure seems to dampen employee mobility and new firms' foundation (Andersson and Klepper, 2013).

Although this group of articles deals with elements that impact on spinoffs' formation, the first two papers are more theoretical while the latter two are more empirical oriented. This difference can be grasped in the dendrogram and icicle plot, since the two couples do not show very high similarity. In MDS map, subfield #5 is located in the middle because of the far-reaching influence of the topic over other research subfields: article on spinoffs in Sweden (#21) concerns also with different types of spinoffs (i.e. pushed and pulled) and with their characteristics (topics of principal components #1 and #2), article on the impact of R&D activities of the parent firm (#16) deals also with innovation (topic of principal component #7).

- Subfield #6 is formed by articles which deal with *innovation* in spinoffs. Ferriani et al. (2012) introduce the concept of *re-imprinting*, a process that indicates both the continuity of parent organizational routines and culture and the change towards market needs leading to innovation and superior performance; by focussing on the dyad parent-spinoffs in fact, the origin of some innovations can be grasped but only when spinoff can adapt to the environmental selection forces such innovation can be fruitful. Furthermore, some locational conditions, such as closeness to research institutes and support entities affect cooperation activities with other players that in turn improve market performance in terms of innovation; moreover, non-local cooperation ties are more conducive to innovation (Lejpras and

Stephan, 2011). Finally, ex-employees that detected business opportunities essential to found their new businesses are more likely to implement post-entry innovation activities than other entrants (Fryges et al., 2014).

In MDS map and cluster analysis these articles form a separate group that, to some extent, can be embedded within the group of articles dealing with academic and entrepreneurial spinoffs or, as in the case of the latter article, included within group #1: while the first two articles compare innovation both in academic and entrepreneurial spinoffs, the last one deals with innovation of pulled spinoffs. This is well-rendered by the position of the papers in the map and the cluster formation in the dendrogram; we may also infer that innovation subfield is approached by several perspectives.

- Articles with high loading on principal component #7 investigate the importance of *network* and *social capital* on emergence and performance of spinoffs: spinoffs whose founders rely on a wider array of relationships -intra-organizational, inter-organizational and social or informal- have higher chances to survive and thrive; the network of these relationships is built during the incubation phase -during which the employee is working but has not recognized a business opportunity yet- and during the emergence phase -during which the future entrepreneur has collected the resources and has identified a business opportunity- and it forms the endowment of the new venture (Furlan and Grandinetti, 2014). Camuffo and Grandinetti (2011) claim the role of spinoff as vehicle for transferring knowledge in industrial districts based on the network of relationships that can contribute to new venture's performance and innovation. This group of papers acknowledges the importance of both social capital and inherited capital for spinoffs performance, with a focus on industrial clusters.

  In MDS and cluster analysis it is quite clear the proximity of these papers to articles of group #2 (principal component #2): for this reason and for the similarity of the contents, it may be justified to deem these two papers as members of this group.
- The only article contributing to principal component #8 acknowledges the linkage between founder characteristics and initial team composition of spinoffs, asserting that the successful performance of spinoffs is triggered by wise entrepreneurs who hire employees to fill human capital gaps (Agarwal et al., 2016). Nonetheless, a single article cannot be considered a group, by definition and we will not consider it henceforth.

From PCA, MDS, cluster analysis and articles' content interpretation some conclusions on entrepreneurial spinoffs subfields of research can be drawn. It is worthwhile reminding that this analysis does not demand to find the exact number of clusters within the literature, rather it may be useful to understand how the literature is evolving and which are the subjects under investigation.

As mentioned earlier, some techniques in PCA and cluster analysis may inform how many components and clusters to retain. Table 3.8 synthetizes the desired number of groups according to the different criteria used and provided that a final evaluation should be based on a collective assessment.

Table 3. 8: Criteria and number of groups. Source: personal elaboration

Analysis	Criterion	Notes	Number of groups
PCA	Cumulative variance	75% as common threshold	6
	Scree test	Inflection point	6
	Eigenvalues	To identify the acceptable number of components	8
Cluster Analysis	Level of similarity	Maximum decrease in similarity	5 or 6
	Length of branches	Red line intersection	7 (or 6, excluding paper #42)

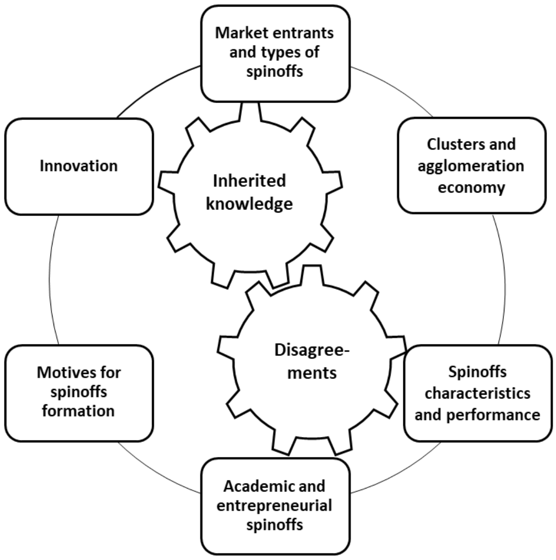
The three analyses and their techniques provide different interpretations starting from a similar grouping: principal components analysis allows the detection of different subfields of research and its major advantage is to indicate those papers contributing to more than one strand, MDS allows the immediate detection of thematic groups at first sight and the possibility to visually mapping an intellectual field, where cluster analysis permits the identification of different groups on the base of shared characteristics and how the similarity between papers affects groups' formation. The three procedures therefore, should be viewed jointly and a collective assessment to identify thematic groups should encompass the results of all these techniques. Some conclusions can be drawn based on these criteria, on all the analyses performed and on the personal interpretation deriving from the 8 components that was reported earlier.

First, in compliance with the criteria of the quantitative analyses, 6 subfields of research should be considered. By using a very short description for each subfield, we may state that the first one concerns with the different types of spinoffs and their performance with respect to other market entrants, the second subfield regards clustering process and agglomeration economy, the third deals with characteristics and performance of spinoffs, the fourth focusses on the comparison between academic and entrepreneurial spinoffs, the fifth investigates factors conducive to spinoffs formation and the sixth examines the topic of innovation in spinoffs. According to eigenvalue criterion of principal components analysis, another group has been detected and it regards the impact of network and social capital on spinoffs, whereas the eighth component is formed by only one article and therefore should not be deemed an independent subfield. As mentioned earlier, articles contributing to the seventh component may be joined with articles of subfield #2, as they delve into the topic within the cluster context. Therefore, 6

main subfields can be extracted and PCA, cluster analysis, MDS configurations have proved to work well for subfields detection.
A second consideration may regard the thematic distinction among different subfields. Sometimes, in fact, a single article can concern with different themes and this posits the possibility to treat jointly topics appearing as different; as mentioned earlier, in the case of groups #2 and #3, parent characteristics and inherited knowledge underpin cluster development and articles dealing with location choice tend to discuss characteristics of performance and characteristics of spinoffs as well. Likewise, articles that contribute to principal component #6 contribute also to subfield #4, since they deal with academic and entrepreneurial spinoffs. This is particularly evident in PCA, in which the loads of some articles are significant to more than a single component. Papers position in MDS map is also coherent with articles contents, as papers at the boundaries of a specific group can deal with different subfields (as in the case of paper #16 included in the subfield of characteristics behind spinoffs formation which also investigates the amount of innovative activity and thus it is close to innovation group).
Third, most of the papers have investigated spinoffs performance and the evidence has shown that spinoffs are outperformers; despite this, the economic environment can benefit more from some types of spinoffs -such as *opportunity spinoffs* and spinoffs with the appropriate endowments of technology and relationships, (see above)-. This entails recommendations to policy makers who should foster entrepreneurship by means of some categories of spinoffs.
Fourth, Klepper and Thompson (2010) article may be deemed a seminal paper: its contribution to three subfields, as results from PCA, is an indicator of its perceived importance in the broad field; as a matter of fact, the *theory of disagreements* is deemed the engine of spinoffs formation and, with the *inherited knowledge theory*, one of the pillars of entrepreneurial spinoffs literature. Most of the papers in fact, stems from consolidated Klepper's theory of the knowledge inheritance and Klepper and Thompson's model to address different research questions on spinoffs. As paper #4, it is almost impossible to bound some articles to a specific group, rather some papers contribute to more than one subfield. Paper #21 is notable as well, since it highlights under which circumstances spinoffs are more likely to occur, but it also compares different types of spinoffs and other market entrants; thus, its pervasive influence on the broad spinoff literature is evident and it confirms the results of citation analysis (see Table 3.3).
Fifth, the automatic keywords frequency analysis illustrated before has proved to be a quite good indicator for a taxonomy purpose; some of the words occurring the most in each subfield were reasonably used as labels to identify the different topics.
Fig. 3.10 depicts the 6 subfields that compose entrepreneurial spinoffs literature as identified by quantitative analyses and our personal assessment from 2009. The two central engines

represent instead the core of the knowledge, namely the two theories which the knowledge on spinoffs stems from.

Fig. 3. 10: Six subfields of spinoffs literature and main engines. Source: personal elaboration



A final consideration should be done: high interdependence has been detected among different subfields, both in terms of common articles and in terms of the wide array of topics which articles can deal with. Some groups of articles may be easily separated, as the cluster corresponding to the comparison between academic and corporate spinoffs, but in other cases the distinction has not been immediate. Topics as knowledge transfer, spinoffs formation and performance recur quite frequently in most articles; perhaps these shared themes are symptoms of the novelty of the topic and that a clear partition between different thematic groups may be reached as the topic matures; at the same time, the distinction in different themes was possible as spinoffs literature has received greater interest by scholars whose interest spans different topics identified in this analysis. Finally, among all the subfields the innovation one is the least clear and the papers which belong to it delve into different issues, which range from innovation in academic and entrepreneurial spinoffs, to which kind of knowledge a spinoff should be endowed with to be innovative.

3.6.2 Articles excluded from co-citation analysis

Following the methodology outlined in previous chapters, some articles on spinoffs have been excluded from the analysis because of threshold requirements. This is the case of articles disclosed in 2015, 2016 and 2017 that, given their novelty, are characterized by very low citations and co-citations frequencies. Therefore, we delved into these articles' contents to gauge their contribution to strands of research identified in the previous section or rather to new research frontiers.

One group of articles scrutinises *factors* triggering *spinoffs' formation* (subfield #5). Employees related to the firm technology focus are less likely to found new ventures; this effect is however mediated by the presence of institutional factors such that strong intellectual property rights decrease further the likelihood to found new ventures for technology-related employees, whereas a higher venture capital increases more the probability of spinoffs for technology-related employees (Yeganegi et al., 2016). Then, the effects of non-compete agreements on spinoffs founders and spawning has been studied. Non-compete clauses are less likely to jeopardise employees' efforts but their enforcement dampens the competition (Buenstorf et al., 2016). However, enforcement of non-compete agreements increases the rate of spinoffs and the social welfare provided that employees can buy out these clauses: governments should therefore support future entrepreneurs by providing them with loans for buying out (Rauch, 2016).

Other articles deal with *characteristics of spinoffs* and *different types* of these new entrants (subfields #1 and #3). In Germany, spinoffs coming from smaller and better performing parents are also more likely to survive and this addresses one of the issues raised by Klepper in his review (see chapter 1). Furthermore, spinoffs tend to be larger and to hire more paid and more skilled employees; pulled spinoffs' likelihood to exit is lower than pushed spinoffs one (Fackler et al., 2016), thereby previous studies (e.g. Eriksson and Kuhn, 2006; Andersson and Klepper, 2013) are confirmed. Another article, instead, posits the existence of another type of new venture along with opportunity and necessity spinoffs, namely *opportunistic spinoffs*: when an adverse event occurs and employees can detect opportunities from such event, this category of spinoffs is created; they are called *opportunistic* since opportunities are discovered in the aftermath of an event (Curran et al., 2016b). While the core of these studies relies on inheritance knowledge theory, the amount of industry-specific experience is investigated with the aim to assess how much the industry-specific experience can affect spinoffs' performance. It is found that in new proprietorships, too much industry-related experience can be harmful for spinoffs and thus a medium level of experience is desirable (Furlan, 2016a). Knowledge inheritance theory, vertical integration and strategic positioning have been jointly studied to predict new firms' performance: by means of knowledge inheritance mechanisms, spinoffs are found to exploit a key value chain activity that their parents previously integrated and to establish a strategic positioning that allows them to survive more than other entrants (Argyres and Mostafa, 2016).

It should be also mentioned that a recent paper by Furlan and Grandinetti (2016) puts forth a new theoretical approach underpinning spinoffs' competitive advantage: the new framework integrates the inheritance theory with the role of *social capital*, positing that intellectual capital of spinoffs stems from parental inherited knowledge but also from valuable relationships, that is social capital, built both within and outside the parent firm. Thus, this paper may contribute to the development of a seventh subfield of research, related to network and social capital in spinoffs (as papers #8 and #31 identified before).

Another group of papers, instead, concerns with *clusters and agglomeration economy* (subfield #2). By examining a cluster in the aircraft industry in the Spokane region, Steenhuis and Kiefer

(2016) find that the initial phase of cluster formation encompasses two stages, the first characterized by conditions and assets endowments that trigger new firms (such as favourable opinions and attentions by policy makers) and a second one characterized by the emergence of new firms and especially spinoffs. Spinoffs were also investigated in peripheral areas with respect to core ones and the results highlight that inheritance theory still holds in the former; furthermore, parent hostility in peripheral areas is more emphasised because of specific geographic conditions occurring in these areas and this can curb spinoffs performance (Habersetzer, 2016). Golman and Klepper (2016) posit a framework to explain clustering without agglomeration economies: clustering may be driven by spinoffs dynamics and new discoveries that form the base for innovation. Finally, the copious social capital in clusters triggers cooperation amongst firms and competitors and new firms' foundation: social capital, agglomeration economies and knowledge legacy are all intertwined and lead to cluster development (Hervas-Oliver et al., 2017).

One of the remaining papers deals with the comparison between spinoffs from *public and private sector* (subfield #4) in biotech industry, claiming the superiority of the latter in terms of attracting venture capital because of deeper innovation competence (Curran et al., 2016a).

Finally, *innovation* topic is dealt with, albeit from different viewpoints. Buenstorf (2016), for example, investigates the extent to which incumbents contribute to industry evolution with respect to new entrants and concludes that innovation of an industry should be conceived more as a systematic phenomenon which blends contributions from both existing and new players rather than a challenge between them to assess who innovates more. A successful innovation occurring at spinoff's level demands instead a mix of capabilities coming from market and technical knowledge, but also from the participation of customers who can pull innovation along the path of needs fulfilment (Scaringella et al., 2017). Finally, innovation steers entry decisions for spinoffs in focal (i.e. founded by ex-employees in the same industry) and downstream industries, as innovative spinoffs are less likely to be influenced by location characteristics (i.e. industry conditions) when they enter a market (Adams et al., 2017).

This array of newer studies has emphasized some intriguing patterns: first, knowledge inheritance model is the litmus test of spinoffs competitive advantage but it should be integrated with some other aspects such as social capital to have a broader view of the spinoffs phenomenon; second, these studies have delved into existing strands of research and have tried to answer to Klepper's 2009 review research questions; third, the interest on innovation is much more widespread and it spans different viewpoints; fourth, newer studies appear to be less engaged in the comparison between academic and entrepreneurial spinoffs. These conclusions are on the same wavelength as the ones drawn by Ferreira et al. (2017) concerning the entire literature on spinoffs: scholars are much more oriented towards innovation, knowledge and social capital.

Therefore, we can conclude that:

> *Spinoffs' literature after Klepper's 2009 review encompasses different thematic groups, that is different subfields of research, which are evolving over time, even though some topics and assumptions are shared in all of them.*

3.6.3 Avenues for future research

The analysis of spinoffs literature *post* Klepper's 2009 review highlights several avenues for future research which can either nurture existing strands or create totally new thematic subfields.

First, social capital and network have proved to be determinants for spinoffs formation and performance even though studies have focussed on cluster environment in Italy; for this reason, they should be investigated, both at the incubation and start-up phase, also in other contexts. In this respect, the topic of interlocking directorates and relationships with competitors should be examined since network and social capital are found to provide competitive edge. Furthermore, they have challenged the absolute validity of knowledge inheritance theory by positing other elements to account for which affect spinoffs in general and in our opinion, they can be deemed a further engine of spinoffs literature.

Second, reasons behind spinoffs formation deserve more attention since previous studies have always delved into the spawning event when parent firm is neutral even though the empirical evidence has shown several examples of hostile parents (e.g. Walter et al., 2014); more specifically, the theory of disagreements may not be enough to explain why entrepreneurial spinoffs are established. In this regard, the consequences related to non-compete agreements should be further analysed, both from a social welfare perspective and from parent and spinoff one and therefore policies should be designed to encourage some parental attitudes for better performer spinoffs. This point may contribute to the development of the subfield related to factors behind spinoffs formation.

Third, the role of spinoffs on employment rate is not clear since it is still unknown to what extent spinoffs -and different types of spinoffs- contribute to employment compared to other market entrants; this kind of analysis may be included into the thematic cluster on the different types of new entrants which has focussed on economic performance so far.

Then, other analyses should focus on clusters and agglomeration economy in order to understand whether clustered areas form because of positive externalities, because of spinoffs events or rather to a combination between the two which exploits also the social capital role; besides, the interrelations among these factors may offer actual explanation for spinoffs outperformance in clustered areas. Thus, this type of investigation falls within the subfield related to cluster and agglomeration economy.

Moreover, academic and entrepreneurial spinoffs comparison should account for many characteristics and mediating factors which can influence the performance of these two new ventures: entrepreneurial spinoffs may perform better than academic ones under certain conditions and worse under other ones. These analyses fall within the cluster of academic and entrepreneurial spinoffs.

To the same extent, further investigations on characteristics and performance of spinoffs are wished for. Given the fact that spinoffs represent the bulk of new entrants in the economy, it is worth understanding what characteristics make them different from others in terms of founding team composition and its influence on performance but also by inspecting their contribution to evolution of industries with respect to other start-ups. Moreover, it should be clarified whether spinoffs are rapacious plunderers of parents' innovations and discoveries or not.

Finally, it is worth pointing out that studies on innovation examined so far have taken several perspectives, even though they have failed to clarify whether spinoffs are innovators or not, also with respect to other new market entrants. In order to evaluate to what extent industry-specific experience affects innovation, an empirical analysis has been conducted and it is presented in the next chapter.

3.7 Sensitivity analysis

To check the robustness of the results, that is to understand whether the identification of different thematic groups occurred because of a specific method, the same analyses have been repeated, albeit with an exception. As mentioned in chapter 2, two schools of thought debate on which values should be put on the diagonal cells of raw co-citation matrix while McCain (1990) acknowledges insignificant differences between the two approaches. Differently from the procedure reported heretofore, we treated diagonal values as missing ones and conducted again principal components analysis, cluster analysis and MDS *ceteris paribus*.

The results[6] are the following:

- PCA results in 6 principal components, thus 6 subfields of research, according to eigenvalues criterion. The main difference with respect to PCA reported before is that Paper #42, which contributed alone to a single principal component, is joined with other articles. Furthermore, articles #4, #19 and #21 contribute significantly to more than one component. Differently

[6] Results of multivariate analyses performed by UCINET are available at Appendix C

from previous subfields identification, the group of papers dealing with motives behind spinoffs formation disappears. According to the scree test criterion, only the first 4 components should be retained. Some papers' loadings on components have changed but the differences are negligible.

- Cluster analysis results in 8 clusters of papers, despite more interpretations are allowed, as in the case of the former cluster analysis. Moreover, the composition of these groups does not fully overlap to PCA ones (e.g. Paper #42 in this cluster analysis forms a different group while in PCA based on missing values it is joined with other documents). In cluster analysis, the group of papers dealing with motives behind spinoffs formation appears.
- Papers in MDS are plotted in a lookalike fashion as in the former MDS map (the stress value is equal to 0.135 and thus acceptable). A very few relative distances among papers are stretched and some others are shortened, but papers altogether do not exhibit significant differences from the analysis reported before.

Overall, these three analyses offer similar results in terms of number and composition of research subfields, despite minimal differences regarding a few papers. Such differences, as happened in the former analysis can be ascribed to articles whose contents are not suitable to be univocally clustered. For example, articles #12 and #19 in cluster analysis form a separate cluster while in PCA contribute to the subfield of academic versus entrepreneurial spinoffs; as mentioned earlier, this is due to the broad content of these articles. Another main difference is that articles dealing with spinoffs formation are not recognised by PCA even though they form an independent cluster in cluster analysis. This group of papers was however recognised also by the former PCA; the different output may be explained by the broad thematic scope of articles dealing with factors which steer spinoffs formation: they regard the origin of different types of spinoffs, characteristics and performance of spinoffs and innovation as well. Therefore, we may state that this group of articles has a cross influence among all the different subfields and its central location on MDS map is not due by chance; sensitivity analysis' s outputs have thus provided useful and deeper insights that complement former findings.

Nonetheless, papers seem to form almost the same groups previously identified: differences of performance of different types of spinoffs, spinoffs and agglomeration economy, performance of spinoffs and their characteristics, academic and entrepreneurial spinoffs, motives behind spinoffs formation. Less clear is however the role of innovation, whose limits have already been presented in previous paragraphs. Notwithstanding these diversities, subfields resulting from sensitivity analysis embed almost the same articles as the ones of previous analysis. Similar to McCain (1990), treating diagonal values in raw co-citation matrix as missing values does not offer different results compared to the ones offered by artificial numbers in the main diagonal in terms of subfields interpretation. We may state, therefore, that literature partition is not significantly affected by the two different methodologies that we have compared in this section.

3.8 Limits of co-citation analysis

This thesis aims to provide a systematization of the literature of spinoffs, which is considered a quite novel topic. Most of the studies concern analysis with a broader time horizon (e.g. Ramos-Rodríguez and Ruíz-Navarro, 2004; Ferreira et al., 2017) and in which the partition into subperiods of analysis can help researchers track the development of a field. In our analysis, instead, the decision to delimit the time span was due to the disclosure of Klepper's review in 2009, which took stock of entrepreneurial spinoffs situation, by synthetizing theoretical and empirical patterns. Our endeavour was therefore to go beyond this analysis and understand what has emerged in this field after Klepper's synthesis. Said differently, we wanted to gauge the status of research on spinoffs and how it is evolving.

We relied on citation and co-citation which are not error-free analyses and which have some limits. One issue is related to why citations occur: it might be misleading to include papers whose citations are irrelevant for the aim of the research, i.e. citations made for methodological purposes should be neglected if the aim of the analysis is the thematical clustering of the papers (Sorkun and Furlan, 2016). Some other citations can occur for paying homage, for criticizing (Ferreira et al., 2017), as self-citations to manipulate citation rates (Garfield, 1979) or to abide by a target quota set by the journal, for example when citing articles previously disclosed in that journal (Backhaus et al., 2011). On the contrary, some citations may be omitted as they could be taken for granted by the researchers and therefore become meaningless; furthermore, a time-related bias occurs, as more recent articles have not been evaluated by the scientific community as older disclosures and thereby they can be underestimated in terms of citations. Co-citation as well may hinder a correct assessment, as the number of co-cited documents is inherently restricted and the interpretation of the results output by quantitative analyses is extremely subjective (Ramos-Rodríguez and Ruíz-Navarro, 2004). Other limitations of this study may concern with the research design: the decision to consider a predefined number of years of analysis (from 2009 to 2017) does not allow a broader judgement of a phenomenon whose roots were taken before 2009. Nonetheless, as mentioned earlier, our attempt was to identify the most recent patterns of the literature on spinoffs and therefore it can be the starting point for a wider assessment on the entire spinoffs literature. A rigorous selection was performed as well on the documents to be retrieved and on the threshold for co-citation analysis, but this was made to ensure coherent works and to limit the probability of biases.

3.9 Conclusion

This chapter has highlighted that after Klepper's 2009 review, spinoffs literature has witnessed different contributions by scholars and by means of quantitative analyses and personal evaluations, subfields of this literature have been identified. As we hypothesised, the literature of spinoffs is motley and the groups detected have tried to shed light on the performance and characteristics of spinoffs with respect to other commercial entrants, on the comparison among different types of entrepreneurial spinoffs, on the relationship between spinoffs and clustering process, on similarities and differences between academic and entrepreneurial spinoffs, on theoretical model for understanding spinoffs' formation and on innovation.

Several issues require further investigations and they emerge from both Klepper's 2009 review and this literature systematization. For example, it is still unclear the role that spinoffs play in the employment rate of economy, the extent to which spinoffs dynamics are intertwined with industries' evolution, the effects of non-compete agreements for social welfare, if spinoffs are plunderers of parents' innovations or are inherently innovative and if spinoffs' outperformance is caused by a screening process which selects the most talented workers or rather to inherited knowledge. Finally, the relationship between spinoffs and innovation is an elephant in the room and, specifically, it is not self-evident the extent to which industry-specific experience can affect the degree of innovation implemented by a new firm that is entering a market. Since this is one of the most debated issues in spinoffs' literature *pre* and *post* Klepper's latest review, we tried to provide an answer by means of an empirical study that is reported in the following chapter.

CHAPTER 4: SPINOFFS AND INNOVATION

4.1 Introduction

As pinpointed in previous chapters, researchers debate on the extent to which spinoffs innovate more than other new ventures. Several attempts were performed to address this issue, but they seldom adopted an empirical viewpoint. Our endeavour is to shed light on the relationship between spinoffs and innovation by studying a sample of new firms and thereby adopting an empirical viewpoint with a two-fold aim: filling this literature gap does not only bootstrap entrepreneurship literature in general but can also orient management and policy makers' actions. Therefore, we investigate whether spinoffs are more likely to innovate with respect to other market entrants during the start-up phase.

4.2 Literature background on spinoffs and innovation

While spinoffs play a crucial role in entrepreneurship process, less understood is their contribution to innovation. Innovation is deemed vital for competitiveness, but also for the entire economy: according to the latest Eurostat reports[7] on the triennium 2012-2014, the bulk of innovative firms witnessed an increase in their turnover and provided environmental benefits, both within and outside them. Indeed, it is not by chance that European Commission has increasingly boosted innovation by means of innovation-oriented policies, such as Horizon 2020 programme.

Innovation is not only associated with a totally new product or technology but it is based on the discovery and exploitation of something new which can be sold in the market at a price higher than the cost to produce it (Shane and Venkataraman, 2000). Stemming from this definition, Furlan (2016b) posits an innovation continuum along which firms can be located and identifies four main points which correspond to different innovation degrees. The first point characterises the so-called *replicative* firms and occurs "when the new firm merely reiterates the products, processes and business model of the incubator firm" (Furlan and Grandinetti, 2014, p.539); founders of these firms do not introduce any innovation. The second degree of innovativeness is based on *marginal* innovation with respect to existing products or services of the industry and therefore leads to the creation of submarkets to fulfil customers' needs. The third degree is typical of firms which introduce *radically* new technology or products in a specific industry whereas the fourth degree of innovativeness characterises those firms which *shape* industry's boundaries and competitive environment by means of totally new technology or product or business models.

[7] Source: https://ec.europa.eu/growth/industry/innovation_en

Literature on spinoffs witnesses two divergent opinions on innovativeness of these market entrants with respect to other new firms. Some aforementioned models (see chapter 1) posit the birth of a spinoff in response to a refusal from the parent firm to undertake innovative projects; a spinoff so originated is, thereby, inherently innovative. Industry experience of founders in hard-disk-drive industry was a determinant in new products that have been deemed innovative (e.g. Christensen, 1993; Agarwal et al., 2004). Moreover, spinoffs tend to show higher innovative potential when they are established than other *de novo* ventures (Lejpras, 2014). Some supporters of this position elicit the innovative nature of spinoffs by recalling previous studies whose focus is mainly on high-tech industries, where innovations are more likely to breed for inherently industry characteristics (Furlan and Grandinetti, 2016). Furthermore, spinoffs are "expected to be innovative, because they can build on experience and relationships established during their previous job at parent firms that other start-ups lack" (Boschma and Weterings, 2005, p.571).

Nevertheless, this latter consideration can be a double-edged sword for supporters of spinoffs as font of innovation. If inheritance knowledge theory holds, spinoffs will be expected to implement practices and routines and exploit knowledge learnt at the former organizations and thus the intellectual capital between parent firm and offspring is likely to be partially overlapping. We use the term "partially" since from a theoretical viewpoint every organization is different from the others both because of internal and milieu characteristics. Given this intellectual capital endowment, spinoffs are expected to implement a not radical type of innovation. As a matter of fact, skills accumulated during prior experience in related industries are more useful to better run a business rather than to trigger superior innovation activities (Fryges et al., 2014). In laser industry, spinoffs initially produced the same lasers of their parents and entered submarkets and niches to meet specific customers' demand (Klepper and Sleeper, 2005) according to a partially differentiating strategy (Chatterji, 2009). Spinoffs are not more likely to implement radical or disruptive innovations since they tend to produce outputs similar -albeit not equal- to their parents' ones (Klepper, 2001).

We may conclude that spinoffs are a peculiar typology of entrants, since they are created by employees who have seized a market opportunity inside the industry and tend to address a market niche's needs better than their parents; at the same time, the accumulated industry experience value is exploited through the implementation of successful practices inherited from the parents. Entrants with no industry knowledge, instead, rely on a different type of know-how and while this increases the likelihood to perish with respect to spinoffs, it lowers the probability to replicate existing products or to create slightly improved products that usually occurs when industry knowledge is inherited. From this consideration, spinoffs seem to implement a not radical type of innovation, which corresponds to the first and second degree of innovation in the framework proposed by Furlan (2016b). Given the different opinions of scholars regarding innovation strategy implemented by spinoffs at the start-up phase, our endeavour is to fill this literature gap. By means of an empirical analysis, we want to test *whether spinoffs are less likely to implement a radical type of innovation when they are established than other entrants or, said differently, whether industry-specific knowledge is less likely to be associated with radical innovation than not industry-specific knowledge.*

4.3 Methodology and data

To conduct our analysis, we used a sample of new ventures founded by Alumni of University of Padova who graduated between 2000 and 2010. The sample was developed from an administered survey already created also for different research purposes[8]; an Alumni survey was deemed appropriate for our research because it allowed cross-industry investigations among comparable individuals from a well-defined population; furthermore, its frequent application in recent years is due to the advantage to survey a population which has ties with university, leading to a higher trust and response rate (Eesley, 2011).

An initial database was created by merging information from statistical office of the University of Padova and from Infocamere S.c.p.A, resulting in 20,338 firms founded by Alumni graduated between 2000 and 2010. However, several restrictions have been made: for firms reporting branches and headquarters only headquarters observations were kept; all entrepreneurs below 18 years old were removed and likewise natural persons, consortia, partnerships and alumni who did not appear as entrepreneurs or owners of their firms but only as managers. These restrictions resulted in a final database of 4,172 firms. Then, a stratified sample was built starting from the legal nature of the firm (individual companies or limited liability companies), the course attended and when the firm was created (before graduation, during university studies, within 5 years after the graduation or post 5 years after the graduation) to obtain as much accurate information as possible (Montana, 2017).

The survey (see Appendix D) was formed by 35 questions grouped in different sections according to the information investigated: the first part collected personal information of entrepreneurs such as personal data and previous working experience; the second part pertained business information concerning markets entered, number of workers, financial performance; the third part regarded environmental considerations such as the presence of organizations and associations and their effects on firms' location choice; the fourth part concerned the founding

[8] We would like to thank Prof.ssa Sedita, Dr.ssa Apa and Dr. Montana of University of Padova for having helped us develop the entire research which will be reported henceforth

team and the relationship ties when multiple founders created the business; the fifth part investigated to what extent university experience affected entrepreneurial choices.

Out of 800 entrepreneurs contacted, a total of 235 answers were collected between March and May 2017 by means of SurveyMonkey platform and Computer-Assisted-Telephone-Interviewing technique[9], obtaining a response rate equal to 29%.

Among these 235 answers, 5 were removed as interviewed people declared to not have founded a firm. Since our research focus is on *de novo* firms, 35 respondents who established a not *de novo* firm were excluded from this analysis. Not *de novo* firms encompassed both voluntary spinoffs (i.e. *de alio* firms) and franchising activities. Furthermore, some respondents omitted some answers and the questionnaires resulted incomplete (in particular concerning financial information of the firm). Notwithstanding this, they were included in the analysis as they contained all the relevant information for this research.

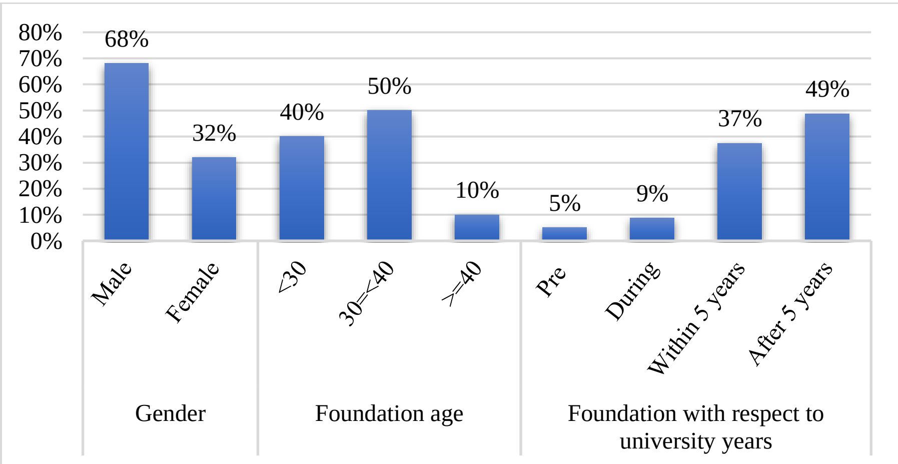
A final sample of 195 *de novo* firms was the final set of observations that was analysed and it encompassed new firms whose founders' working background was completely different.

4.3 Descriptive statistics

Before analysing innovation implementation of new firms, it was worth understanding the composition of the sample according to different characteristics. In the following sections, this information is represented following a specific order: first, information on founders is shown, then information on the new venture at the foundation date and on the perceived importance of ecosystem players is exposed; finally, founders' former working experience and the type of innovation implemented are exhibited, which represent the focus of this chapter.

4.3.1 Entrepreneurs' characteristics

Fig. 4. 1: Descriptive statistics: entrepreneurs' characteristics. Source: personal elaboration



[9] We would like to thank also Dr. Montana and Dr. Cortese of University of Padova for the aid to collect the answers

Entrepreneurs' characteristics analysed regard gender, age at the foundation date and when the new venture was founded with respect to university experience of founders. Percentage values of such features are represented in Fig. 4.1. In our sample, 68% of entrepreneurs are males and only 32% are females while when considering the entire sample (230 answers, see previous sections) this gap shrinks and the percentage values are 64% for males and 36% for females. While such data may appear astonishing, it perfectly reflects the European trends on female entrepreneurship, as only 30% of start-up entrepreneurs in the European Union are women; this disparity is ascribed to some problems which women must cope with, such as information and finance-related issues but also reconciliation of business and family. Since female entrepreneurship represents a source of growth to be exploited, the European Commission is working to foster it through several initiatives[10] (European Commission, 2017).

The sample shows also that the bulk of entrepreneurs (60%) founded their ventures after 30 years old and half of respondents started up their businesses between 30 and 40 years old. These results confirm that the *garage belief* has been unduly celebrated as new entrepreneurs are not very young and they may have accrued some working experience before their start-up act, as several studies report (e.g. Bettiol and Furlan, 2014). Similar percentages are found when the entire number of respondents is considered.

Likewise, by looking at when new start-ups are created relatively to university experience of their founders, 86% of new firms are founded after the graduation and among these nearly half of them are established after 5 years from the graduation date. Although some information on entrepreneurs' experience between graduation date and new firms' establishment lacks, we may hypothesize that entrepreneurs of the sample have likely accumulated some working experience during this time span. Based on these descriptive statistics, the myth of college dropout who starts a new venture which underpins the garage belief is more a myth than an empirical evidence.

4.3.2 New ventures characteristics

Descriptive analysis was performed concerning some firms' characteristics such as location choice, operating sector, nature of the firm, initial number of employees and nature of funds used whose percentage values are reported in Fig. 4.2.

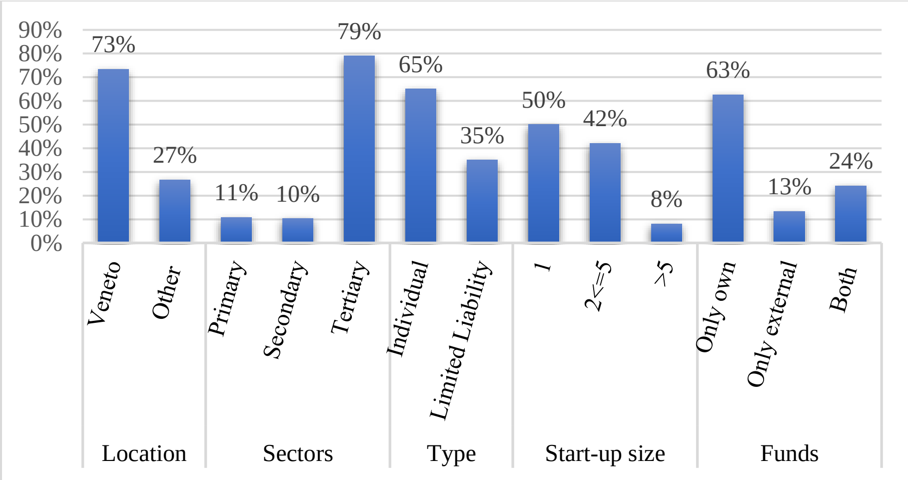
First, the bulk of *de novo* firms are founded in Veneto (73%) and the remaining in different Italian regions. This data may be influenced by the origin of entrepreneurs but also by the network relationships established by founders during their years at the University of Padova. Second, most of the firms are established in the tertiary sector and only a few firms are created

---

[10] Source: https://ec.europa.eu/growth/smes/promoting-entrepreneurship/we-work-for/women_en

in agriculture and manufacturing industries and this is coherent with the recent European pattern highlighted by Istat (2015). Third, individual firms are predominant with respect to limited liability firms as the former reach 65% and the latter only 35% in our sample.

Fig. 4. 2: Descriptive statistics: new ventures' characteristics. Source: personal elaboration



Then, start-up size at the foundation has been analysed, measured in terms of workers number, encompassing both managers and employees; according to the answers collected, half of new ventures counted 1 worker who was the entrepreneur herself, 92% of new firms counted at most 5 workers, whereas only 3% of new firms counted more than 10 workers. This trend reflects what Eurostat pinpointed concerning the size analysis of new firms related to year 2012, where sole-entrepreneur's firms represented 47% of newly formed businesses in 2012 in the European Union; moreover, Eurostat reported that in 2014 the share of micro enterprises (i.e. firms with less than 10 workers) in EU-28 was 93%[11].

Finally, the financing dimension was investigated and three categories were arbitrarily created depending on the different funds which entrepreneurs relied on when they established a new firm: 63% of entrepreneurs relied only on personal funds and/or funds provided by family or friends; 13% of entrepreneurs used only external funds such as loans, private venture, seed and business angels' capitals, public funds or capitals provided by other firms; 24% used both own and external capital. Access to finance is not deemed by new entrepreneurs the biggest challenge when establishing a new venture[12] since problems related to customer base, competition and availability of skilled personnel are more difficult to overcome. Nevertheless, European Commission acknowledges that SMEs' growth in Europe is curbed by loans' refusal and lack of funds and for these reasons it has drafted the programme COSME (for the

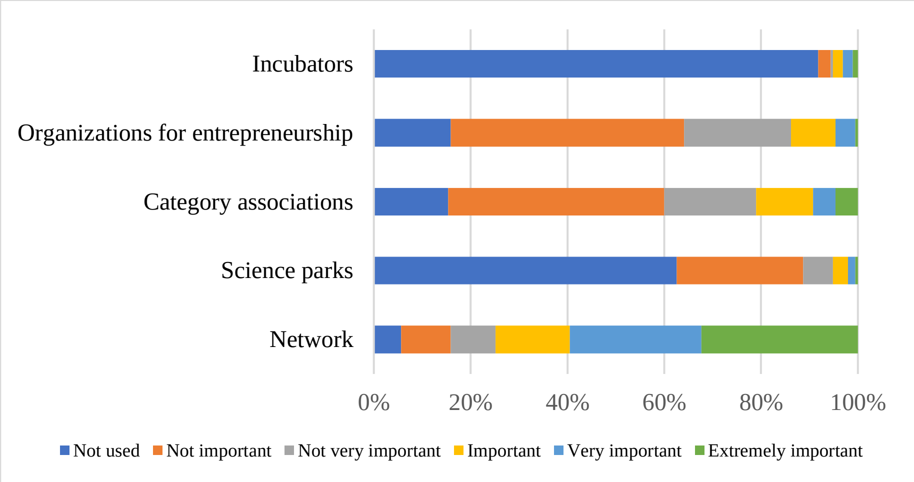
[11] Source: http://ec.europa.eu/eurostat/statistics-explained/index.php/Structural_business_statistics_overview#Size_class_analysis

[12] Source: Survey on the access to finance of enterprises, 2016, European Commission: https://ec.europa.eu/growth/access-to-finance_en

Competitiveness of Enterprises and Small and Medium-Sized Enterprises), which would facilitate SMEs' access to loans and equity finance.

4.3.3 Ecosystem players

Fig. 4. 3: Descriptive statistics on ecosystem players. Source: personal elaboration



As mentioned earlier, one of survey's sections regarded the interaction between the new business and some ecosystem actors in the decision to create a new venture. Fig. 4.3. synthetizes entrepreneurs' personal evaluations on to what extent the presence of key ecosystem players affected such choice. These evaluations were recorded by using a Likert scale which assumed values from 1 to 5, where 1 was "not important at all" and 5 was "extremely important"; when entrepreneurs did not rely on these ecosystem players they answered "not used" and a 0 value was recorded. Fig. 4.3 highlights some interesting trends: first, incubators and science parks were not used respectively by 92% and 63% of entrepreneurs for new ventures' creation choice; then, the presence of public and private organizations to foster entrepreneurship and category associations was deemed not so important respectively in 70% and 64% of the cases; on the contrary, personal relational network was considered important by 74% of respondents. Results point out the dominance of network relationships with respect to other environmental players, whose roles have been deemed crucial for growth and entrepreneurship by the European Commission (2017) but which fail to impact on new ventures creation in our sample.

4.3.4 Experience accumulated: spinoffs or start-ups?

The focus of this chapter is to find to what extent the type of experience that founders accumulated before the creation of the new firm affects the innovation implemented during the start-up phase. More specifically, we want to understand whether the fact to be a *spinoff* compared to be a *start-up* influences the orientation towards a radical type of innovation which, in broader terms, means to find a relationship between the type of new venture and the entity of innovation.

Therefore, an insight concerning the previous experience of entrepreneurs in the sample was yielded and it is exhibited in Fig. 4.4 and Fig. 4.5.

Fig. 4. 4: Previous experience. Source: personal elaboration

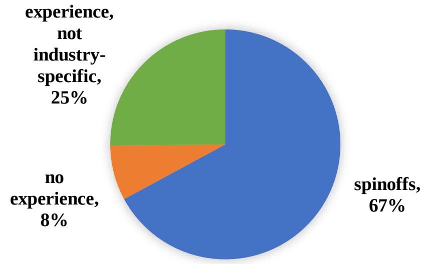


Fig. 4. 5: No industry-specific experienced respondents. Source: personal elaboration

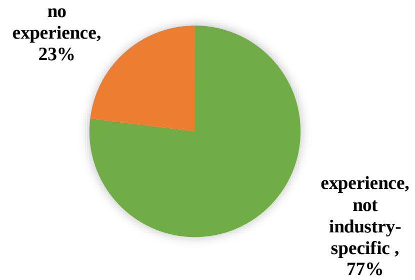
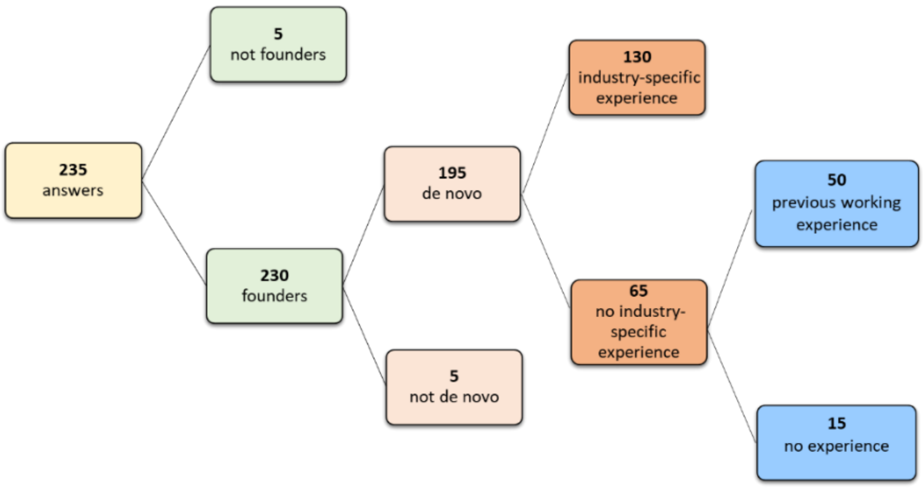


Fig. 4.4 shows the sample composition depending on whether the entrepreneur had previous working experience and which type of experience she accumulated. If respondents answered that they accumulated previous working experience, they were asked whether they accumulated it in their family's businesses or other firms, operating in either the same or in a different sector, but also if they worked for consulting firms or if they had previously founded other ventures, operating in either the same or in a different sector. New ventures result to be spinoffs in 67% of the cases, that is 67% of entrepreneurs accumulated industry-specific experience; in our research, a new venture is deemed a *spinoff* if its founder worked either in her family's business operating in the same industry or in other firms belonging to the same industry, but also if entrepreneur has founded previous businesses in the same industry. Then, 25% of *de novo* firms are founded by entrepreneurs with previous working experience, albeit in industries other than the one where they were operating, that is they worked for either their family's businesses operating in different industries or in other firms belonging to different industries or for a consulting firm or they have founded a new firm in other industries. Finally, only 8% of entrepreneurs are inexperienced, that is they founded a new firm without previous working experience of any kind. These descriptive statistics show that the bulk of entrepreneurs had previous working experience which is mainly industry-specific, thereby confirming the trend highlighted by previous studies (e.g. Klepper, 2001; Audia and Rider, 2005; Furlan and Bettiol, 2014; Furlan, 2016b).

Fig. 4.5 instead reports the composition of entrepreneurs who lack industry-specific experience (and therefore they are not spinoffs). Among these, 77% of entrepreneurs accumulated previous working experience in other industries and only 23% of industry-specific inexperienced entrepreneurs did not accumulate any kind of experience; this further points out the importance of working experience in the creation of a new firm, coherently with previous research.

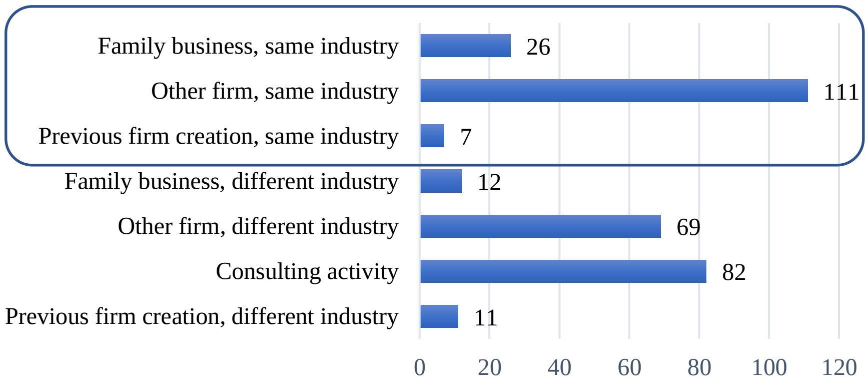
Fig. 4.6 wraps up the steps followed and the dimensions used to classify entrepreneurs starting from 235 answers collected, according to their founder status, the type of firms they created and the type of former working experience they accumulated.

Fig. 4. 6: Entrepreneurial classification summary. Source: personal elaboration



While our focus is on industry-specific experience, we think that analysing where entrepreneurs previously worked is likewise important. Fig. 4.7 shows where entrepreneurs accumulated previous working experience: a large share of respondents worked for other firms in the same industry (111), then many entrepreneurs worked as consultants and in other firms operating in different industries. As only 7 respondents created former firms in the same industry, we may affirm that serial entrepreneurs are very few in our sample. Furthermore, the sum of the answers does not result in 195 (number of respondents in our *de novo* firms sample) since several respondents have accumulated both industry-specific and not industry-specific experience. Furthermore, the blue circle highlights where spinoffs' founders accumulated previous industry-specific experience and since the total answers are higher than 130 (i.e. the number of spinoffs in our sample), we may conclude that some spinoffs' entrepreneurs accumulated experience in more than one industry-specific context.

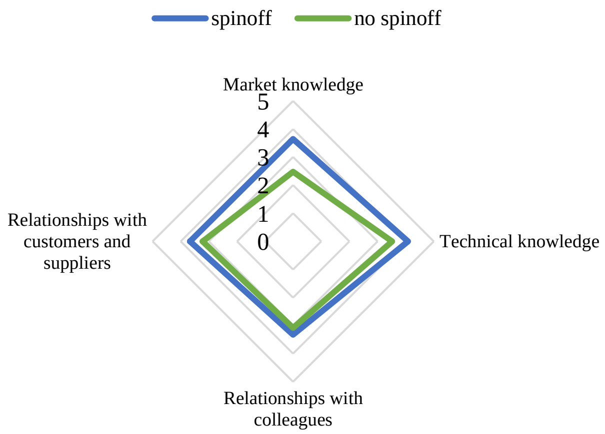
Fig. 4. 7: Type of working experience accumulated. Source: personal elaboration



The impact of former working experience on new ventures' performance and survival is one of the hottest topics in entrepreneurship literature whose focus is on both spinoffs and other market entrants. Previous working experience is expected to influence entrepreneurial choice and development from different viewpoints as the new entrepreneur may have inherited technical and market knowledge to be exploited when running her new business (e.g. Klepper, 2001; Agarwal et al., 2004). Indeed, previous working experience is characterized by different elements which are not expected to equally influence entrepreneurs' choices, even though we may hypothesize that altogether these elements are grist to spinoffs' mill and therefore they are more relevant for this category of market entrants than for other *de novo* firms.

Consequently, we examined how new entrepreneurs evaluated some of these working knowledge facets, depending on the type of new venture and the results are reported in Fig. 4.8.

Fig. 4. 8: Knowledge elements accumulated during previous experience. Source: personal elaboration



Likert scale was used to evaluate market knowledge, technical knowledge, relationships with customers and suppliers and relationships with colleagues which represented some of the knowledge elements gained by previous working experience; values ranged from 1 to 5, where 1 was "not important at all" and 5 "extremely important". As Fig. 4.8 shows, all these elements were rated with higher average scores by spinoffs' founders, thereby confirming what we expected. Furthermore, while the average rank for the element "relationships with colleagues" is almost equal between the two groups of new firms, market and technical knowledge report the largest average scores difference. This could be explained by the fact that since market and technical knowledge are the most industry-related characteristics, spinoffs' founders are more likely to exploit them when they run a new venture in the same industry in which such knowledge has been accumulated. Other *de novo* founders instead are more likely to deem unexploitable market and technical knowledge they gained in other industries; as a matter of fact, market knowledge received an average score of 2.48 by not spinoffs' founders, which was the lowest score among all the four elements analysed. On the contrary, spinoffs' founders rated technical knowledge with the highest score (score: 4.08), followed by market knowledge (3.65); still, the importance of these types of know-how with respect to the other elements confirms some empirical analyses conducted on spinoffs who rely on market and technical knowledge as distinctive parameters of spinoff firms (e.g. Agarwal et al., 2004).

4.3.5 New firms and type of innovation

While recent years have witnessed an increasing attention for entrepreneurship phenomenon, likewise innovation has elicited growing interest by several scholars in different research fields. Entrepreneurship and innovation are, in fact, two intertwined phenomena, as put forth by Schumpeter (1934). According to this author, an entrepreneur is someone who introduces an innovation to the market, that is something capable to increase the usefulness of existing products. Schumpeterian innovation is thereby a change, even small, which can revert an economic business cycle and it is responsible for hectic fluctuations of economic history.

Innovation has also become one of the priorities in European Commission's agenda as firms which develop at least one innovation increase their turnovers significantly.

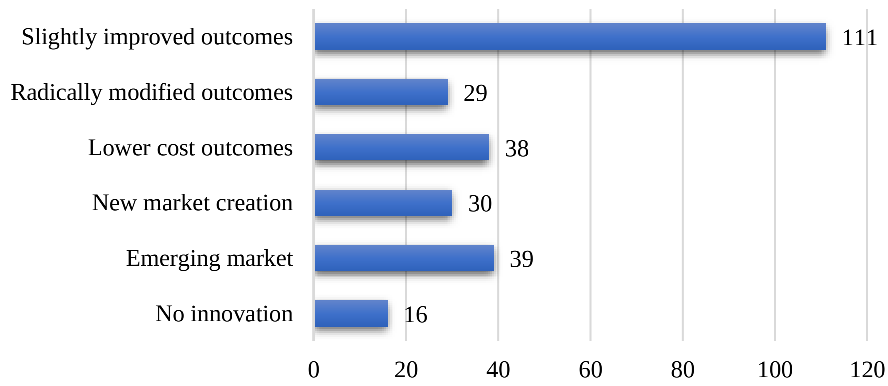
It is worth reminding that innovation is a complex and difficultly measurable phenomenon. A study conducted by Apa et al. (2016) on Veneto firms highlights that ventures located in this region do not frequently earmark funds for R&D activities, albeit they show significant innovations. Consequently, measuring innovation only by considering input variables such as R&D investments can be misleading. Innovation can neither be deemed a bipartite phenomenon (incremental innovation versus radical innovation), rather it can be measured by a continuum along which different types of innovation can be detected and along which firms are located (see above).

In the section of the questionnaire related to firms' characteristics, new entrepreneurs were asked whether they established their ventures in an existing market by producing slightly improved products or services, or by providing radically modified outcomes, or by providing products and services at a lower cost, or they realised a totally new product/service which created a new market, or they entered an emerging market. By considering Furlan's framework for innovation (see above), a *de novo* firm was deemed to implement *radical* innovation if it created a new market or/and if it produced radically modified products or services; when respondents instead answered one or more of the other alternatives they were labelled as *incremental* innovators; moreover, a few respondents answered that no innovation at all was implemented by their firms and for the sake of simplicity, we deemed them as incremental innovators. This crude classification entails some limits but since we are interested in evaluating to what extent spinoffs are more likely to implement a radical innovation with respect to other start-ups, naming *incremental* innovation a situation in which no innovation occurred is a matter of labelling and we believe that it will not compromise our findings.

Since the focus of this section is to shed light on the relationship between the type of new venture, that is the type of previous working experience, and the type of innovation, descriptive statistics were conducted on the sample by analysing also this latter dimension.

Fig. 4.9 reports the number of respondents according to the different types of innovation implemented. As respondents could choose more than one answer and some entrepreneurs did not provide only one answer, the sum of answers of the graph does not result in 195. Fig. 4.9 shows also entrepreneurial inclination towards incremental innovation and in particular towards slightly improved products and services in existing markets (111 on 195 total respondents), whereas 16 acknowledged to not have implemented any innovation. A few of them introduced radically modified outcomes or created totally new markets (respectively 29 and 30).

Fig. 4. 9: Type of innovation of de novo *firms in the sample. Source: personal elaboration*



De novo firms were also analysed on a more aggregated basis, that is by considering whether they were incremental or radical innovators. Results are provided in Fig. 4.10. The bulk of *de novo* firms implemented incremental innovation (73% improved existing products or services or provided them at a lower cost or entered an existent emerging market or implemented no innovation at all) and less than one third of them implemented radical innovation.

Fig. 4. 10: Innovation implemented by de novo *firms. Source: personal elaboration*

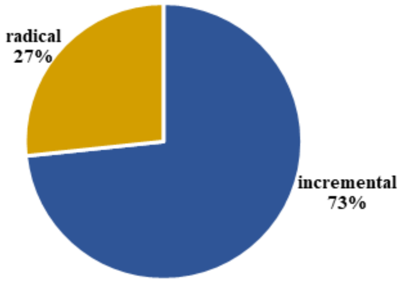


Fig. 4. 11: Type of innovation per de novo *firm. Source: personal elaboration*

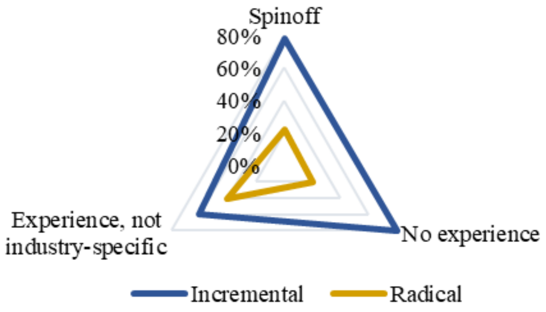


Fig. 4.11 instead shows the relationship between type of innovation implemented and type of new firm that appears from a sample analysis. The three types of firms analysed are spinoff firms, firms whose entrepreneurs accumulated experience albeit not industry-specific and finally firms founded by inexperienced entrepreneurs. In general, all the firms were more inclined towards incremental than radical innovation; nevertheless, this latter was implemented more by firms founded by entrepreneurs with previous experience not industry-specific (40%) rather than by spinoffs (22%) or other start-ups (20%). These preliminary descriptive results appear to suggest that firms with previous industry-specific experience (spinoffs) are less likely to pursue radical innovation than firms which lack such industry-specificity despite having accumulated working experience.

4.4 Type of experience and innovation: beyond descriptive statistics

Although descriptive statistics is a required step to frame the phenomenon under investigation, the existence of a relationship between type of innovation and type of experience has not been demonstrated so far. In this section, results coming from inferential statistics and based on sample information are provided.

4.4.1 Chi-square tests

A joining link between such descriptive statistics and the presentation of a model to draw conclusions about the population from sample information is represented by a Chi-square test. This test examines whether two variables take their values in an independent fashion one from the other or not and therefore it is based on a test for independence whose null hypothesis claims the independence of the variables.

Chi-square test is computed by means of *observed* and *expected* frequencies, that is respectively the frequency with which a certain event occurs and the frequency expected for each category if the null hypothesis is true. When H_0 is true, observed and expected frequencies are not significantly different (Field, 2009).

The starting point for this is a contingency table which sums the observed frequency per category and the estimates of expected frequencies.

By using the software SPSS 22.0 with sample information as input, Table 4.1 results.

Table 4.1 sums the observed ("count") and expected ("expected count") frequencies of three groups (firms with industry-specific experience, firms with no experience at all and firms with previous experience albeit not industry-specific: variable *experience*) according to the type of innovation (incremental and radical). As it can be noticed, some differences between observed and expected frequencies occur but we need to understand whether such difference is due to chance or to a significant relationship between the variables.

Table 4. 1: Table with expected and observed frequencies. Source: SPSS 22.0 and personal elaboration

			experience			
			ind.specific	*no.exper*	*no.ind.specific*	Total
innovation	*incremental*	Count	101	12	30	143
		Expected Count	95.3	11.0	36.7	143.0
		% within experience	77.7%	80.0%	60.0%	73.3%
	radical	Count	29	3	20	52
		Expected Count	34.7	4.0	13.3	52.0
		% within experience	22.3%	20.0%	40.0%	26.7%
Total		Count	130	15	50	195
		Expected Count	130.0	15.0	50.0	195.0
		% within experience	100.0%	100.0%	100.0%	100.0%

The χ^2 test was performed by using SPSS 22.0 and the statistic χ^2 (computed as the sum of the observed frequencies minus the expected, squared, and divided by the expected frequencies) resulted 6.149 and the asymptotic significance equal to 0.046[13]. If the significance level is set equal to 5%, as normally occurs, we reject the null hypothesis of independence between the two variables as the observed pvalue is lower than 0.05 (0.046<0.05); furthermore, the χ^2 statistic (6.149) falls in the rejection region since it is higher than the critical value of a $\chi^2_{(2,0.05)}$= 5.991. Consequently, we may affirm with a confidence level of 95% that the difference between expected and observed frequencies in our case is not due only to chance, rather there is a real difference in the type of innovation implemented by different categories of new firms and thereby an association between the background of the entrepreneur and innovation implemented by the firm. Furthermore, since a rule of thumb for the use of the χ^2 requires that expected values should not be less than 5 in more than 20% of the cells (Field, 2009) and since in our case the percentage was equal to 16.7%, no assumption was violated.

Another χ^2 test was performed, this time by merging those *de novo* ventures which lack industry-specific experience, that is those with no experience at all and those with previous working experience in different sectors. Thereby, the two groups on which the test was performed were represented by spinoffs (130 firms) and firms with no industry-specific experience (65, that is 15 founded by totally inexperienced entrepreneurs and 50 created by experienced entrepreneurs). The χ^2 statistic resulted equal to 3.8 and the associated pvalue equal to 0.052[14] which compared respectively to a $\chi^2_{(1, 0.05)}$= 3.84 and a pvalue equal to 0.05 make us accept the null hypothesis that the variables innovation and experience are not associated, with a confidence level of 95%. Nonetheless, with a confidence level of 90% the two variables result associated and no assumption was violated since the percentage of minimum expected frequencies with values lower than 5 was 0%.

A further χ^2 test was conducted by excluding those de novo ventures founded by entrepreneurs

[13] Results of χ^2 are available at appendix E
[14] Results of χ^2 are available at appendix E

with no previous working experience and thus the analysis was conducted on 180 (130 spinoffs and 50 no spinoffs) *de novo* ventures. This has a two-fold aim: on one hand, it provides a more trustworthy scenario since the bulk of firms are usually founded by experienced entrepreneurs as shown both in this work and in other studies; on the other hand, our interest is more on the extent to which industry-specificity affects innovation choices than on having an experienced background or not. Therefore, the contingency table, expected and observed frequencies were retrieved and a χ^2 statistic equal to 5.71 and an asymptotic significance equal to 0.017 were obtained[15]. Since the χ^2 statistic was higher than a $\chi^2_{(1,\,0.05)}$=3.841 and the observed pvalue lower than 0.05, with a confidence level of 95% we can affirm that a significant association between industry-specificity experience and type of innovation implemented exists. In this case neither the assumption on the percentage of cells with minimum expected frequencies was higher than 20% (the percentage was equal to 0%).

The three χ^2 tests performed provide an indication on the existence of an association between the innovation implemented and the nature of *de novo* entrants. However, the entity and the sign of such relationship deserves a deeper attention through other analyses.

In the following sections, the nature of this association is analysed by using different logit models which almost reflect the path followed by χ^2 tests, that is by classifying the firms differently according to their attributes: first an analysis will be performed by distinguishing firms with previous industry-specific experience and those without such experience, thereby two categories will be analysed; then, firms will be distinguished into 3 categories, which correspond to the ones analysed in the first χ^2 test and different baseline categories will be employed to appreciate the innovation phenomenon; finally, only firms with previous working experience, either industry-specific or not will be analysed and therefore only two categories will form the two groups under investigation.

Before moving forward with these models, a description of the variables employed is provided.

4.4.2 Variables: dependent, independent and controls

To build the model conducive to an explanation of the relationship between spinoffs and innovation and to understand which variables should have been used, questionnaire's sections were scrutinized. As re-marked several times in this chapter, our focus is on previous working experience of the founder and its impact on the degree of innovation implemented by her firm at a start-up phase and thus the impact of being spinoff rather than other entrants on innovation. We employed as dependent variable the *type of innovation* implemented by a firm, which could be *incremental* or *radical,* according to the market choice made during the start-up phase.

[15] Results of χ^2 are available at appendix E

Respondents were asked if they created a new market or if they delivered radically modified products or services and whenever they performed at least one of these two alternatives the firms were deemed radical innovators, otherwise incremental ones. We relied on such definition of radical innovation as Furlan (2016b) reports that firms implement a radical innovation when they radically change technology or products or when they "can initiate the development of new markets" (Keuning, 2007, p.621). On the other side, incremental innovation involves small cost or feature improvements in existing products and services in existing markets (Pham, 2011). We used the same approach employed by Boschma and Weterings (2015) to classify the type of innovation strategy implemented, that is by asking to respondents the kind of innovation at the start-up phase and consistently with previous definitions of incremental and radical innovation, we classified entrepreneurs into these two categories of innovation. Boschma and Weterings (2015) employ a dummy variable which assumes value 1 when innovation is radical and 0 when it is not; while the authors consider radical innovation only the situation in which new firms introduce new products and services in the market, we employed a dummy variable as well which assumed value 1 when innovation strategy was radical and 0 when incremental, but we relied on a broader definition of radical innovation, as mentioned earlier.

The independent variable of interest was instead the *type of experience* that founders accumulated before founding their new business; such experience was distinguished into 3 categories, that is no experience, experience but not industry-specific and industry-specific experience which are observed respectively in 15, 50 and 130 *de novo* firms. As Furlan (2016a) reports, the definition of *spinoff* is also debated in the literature; while some studies deem *spinoffs* those new firms founded by entrepreneurs who have been employed in the industry only in the previous period, in this study we consider all the firms whose entrepreneurs have accumulated industry-specific experience at any moment in the past, coherently with other analyses (e.g. Furlan, 2016a).

Finally, a set of control variables was used, at firm and industry-level, to isolate the effect of experience and which are capable to affect the innovation implemented by firms:

- *start-up size* of the firm, that is the number of workers at the foundation which included both founders and employees
- *location* of the new venture, namely if it is located in Veneto or not
- use of *ecosystem players*, that is if the entrepreneur relied on incubators, organizations for entrepreneurship, category associations, science parks or network for firm's foundation
- presence of a *partner* to start the new business
- type of *funds* used to create the new venture, that is if the entrepreneur relied on only external funds or not

- *industry* in which the firm started to operate, that is if it belongs to primary, secondary or tertiary sector, according to the ATECO code reported.

Table 4.2 synthetizes all the relevant information for the variables used, that is the type of variable, the labels used in the software SPSS 22.0 through which the analyses were performed and a brief description of each variable.

Table 4. 2: Variables employed in the model. Source: personal elaboration

Variable	Description	Label
Dependent: **Innovation implemented**	Dummy: value 1 if radical innovation and 0 if incremental	innovation
Independent: **Entrepreneurial experience background**	1st model: 2 categories, 1 dummy: value 1 if industry-specific experience was accumulated and 0 if not	spinoff
	2nd model: 3 categories, two dummies. Two models with different baseline: A) spinoff as reference category ▪ 1st dummy: value 1 if no experience at all was accumulated and 0 if it was ▪ 2nd dummy: value 1 if experience, but not industry-specific and 0 if it was industry-specific B) no experience as reference category ▪ 1st dummy: value 1 if spinoff and 0 if not ▪ 2nd dummy: value 1 if experience but not industry-specific and 0 if not	A) ▪ 1st dummy: no.exper ▪ 2nd dummy: no.ind.spec B) ▪ 1st dummy: spinoff ▪ 2nd dummy: no.ind.spec
	3rd model: 2 categories, 1 dummy: value 1 if entrepreneur had previous industry-specific experience and 0 if not. Inexperienced entrepreneurs excluded.	spinoff
Control: number of workers at the foundation date	Discrete, it ranges from 1 to 30	size
Control: Location, where the firm is established	Dummy: value 1 if in Veneto and 0 if elsewhere	location
Control: Reliance on incubators or science parks	Dummy: value 1 if entrepreneur relied on incubator and/or science park for firm's foundation and 0 if not	incu.scipark
Control: Reliance on organizations or associations	Dummy: value 1 if entrepreneur relied on organizations for entrepreneurship and/or category associations for firm's foundation and 0 if not	org.assoc
Control: Reliance on network	Dummy: value 1 if entrepreneur relied on her network for firm's foundation and 0 if not	network
Control: Presence of a partner	Dummy: value 1 if entrepreneur founded her business with at least one partner and 0 if she founded it alone	partner
Control: External funds usage	Dummy: value 1 if entrepreneur relied on only external funds (venture capitals, loans, other public and private funds) and 0 if not	funds
Control: Industry in which the firm was established	Categorical (3 categories: primary sector, secondary sector and tertiary sector). Two dummies (reference category: tertiary sector): ▪ 1st dummy: value 1 if firm operates in primary sector and 0 if not ▪ 2nd dummy: value 1 if firm operates in secondary sector and 0 if not	▪ 1st dummy: primary ▪ 2nd dummy: secondary

4.4.3 Results of the analysis: logit models

The focus of this work is to delve into the relationship between the type of new venture and the type of innovation implemented. More specifically, we want to understand whether spinoffs are more likely to implement radical innovation with respect to firms which lack industry-specific experience or not.

In order to do that, we distinguish among different logit models which can be useful to shed light on spinoffs' innovative behaviour. We relied on binary logistic regressions for prediction in all the models as in all of them the dependent variable (innovation) could assume either 0 or 1 values, and we conducted the analysis by using SPSS 22.0.

Rather than predicting the response of the dependent variable (Y) directly as it happens in linear and multiple regression, logistic regression models the probability that the dependent variable belongs to a specific category. Logistic regression employs the logistic function

$p(X) = \frac{e^{\beta_0+\beta_1 X_1+..\beta_p X_p}}{1+e^{\beta_0+\beta_1 X_1+..\beta_p X_p}}$ and relies on the concept of *odds*, that is defined as the quantity $\frac{p(X)}{1-p(x)} = e^{\beta_0+\beta_1 X_1+..\beta_p X_p}$. Odds values near 0 signal very low probability that Y=1 while values near ∞ signal a high probability that Y=1 (James et al., 2013). In our study, the odds quantity is the probability of implementing radical innovation given some characteristics of the predictors divided by the probability of implementing an incremental innovation, held equal the same predictors' characteristics.

It is also worth reminding that in logistic regression the estimate of a coefficient measures the change in the log-odds when the corresponding X variable increases by 1 unit and therefore there is not a straight-line relationship between p(X) and X; nonetheless, when β_i (the estimate of the coefficient) is positive and X increases by 1 unit p(X) will increase whereas when β_i is negative and X increases by 1 unit p(X) will ebb (James et al., 2013).

Each logit model will be presented below in a separate section.

4.4.3.1 Logit 1: industry-specific experience and no industry-specific experience

The first logit model studies to what extent having industry-specific experience increases the probability to implement radical innovation than firms whose entrepreneurs lack such industry specificity. In other words, to what extent being a spinoff affects the probability to implement radical innovation than being a non-spinoff?

In this model, the entire sample of 195 *de novo* ventures are considered and are distinguished according to whether they are spinoffs or not.

In this study, the logistic function is therefore the following one:

$$p(X) = \frac{e^{\beta_0+\beta_1 \text{spinoff}+\beta_2 \text{size}+\beta_3 \text{location}+\beta_4 \text{incu.scipark}+\beta_5 \text{org.assoc}+\beta_6 \text{network}+\beta_7 \text{partner}+\beta_8 \text{funds}+\beta_9 \text{primary}+\beta_{10} \text{secondary}}}{1 + e^{\beta_0+\beta_1 \text{spinoff}+\beta_2 \text{size}+\beta_3 \text{location}+\beta_4 \text{incu.scipark}+\beta_5 \text{org.assoc}+\beta_6 \text{network}+\beta_7 \text{partner}+\beta_8 \text{funds}+\beta_9 \text{primary}+\beta_{10} \text{secondary}}}$$

where our variable of interest is *spinoff*, associated with coefficient β_1.

Since one of the fundamental assumption is that independent variables should not exhibit high multicollinearity (Field, 2009), we conducted both a correlation analysis among the variables and we employed the Variance Inflation Factor (VIF)[16]; the predictors showed VIF values significantly lower than 10 and correlation coefficients lower than 0.8, which are the rules of thumb to assess multicollinearity (Field, 2009). Therefore, we could conduct a logistic regression, whose output is recorded in Table 4.3.

Table 4.3 synthetizes information provided by the software SPSS 22 on coefficient, standard error, observed pvalue, odds ratio and confidence interval at 95% for each variable of logistic regression.

[16] Multicollinearity statistics available at Appendix F

Table 4. 3: First logistic regression: difference in innovation between spinoffs and no spinoffs. Source: SPSS and personal elaboration

	Coefficient	Standard error	Sig. Pr(>|z|)	Exp(β): Odds ratio	95% C.I.for EXP(B)	
					Lower	Upper
spinoff	-0.774	0.366	0.034	0.461	0.225	0.944
size	0.036	0.049	0.456	1.037	0.942	1.141
location	-0.102	0.409	0.804	0.903	0.405	2.012
incu.scipark	1.332	0.381	0.000	3.787	1.796	7.986
org.assoc	-0.779	0.577	0.177	0.459	0.148	1.423
network	-0.978	0.748	0.191	0.376	0.087	1.629
partner	-0.259	0.432	0.548	0.772	0.331	1.799
funds	-0.141	0.368	0.702	0.869	0.423	1.786
primary	0.063	0.582	0.914	1.065	0.340	3.335
secondary	1.090	0.557	0.051	2.973	0.997	8.865
Constant	0.448	0.822	0.586	1.565		

The first variable that should be analysed is our variable of interest, namely *spinoff*. The estimate of the coefficient is a negative number (-0.774) and therefore it seems that being a spinoff with respect to not being a spinoff lowers the probability to implement radical innovation. By looking at the observed pvalue (0.034) and the confidence interval at the level 95% associated with such variable, we infer that such estimate is also significant as the pvalue is lower than 0.05 (the common threshold in these studies) and the confidence interval does not include the value 1 for the odds ratio. Therefore, with a confidence level of 95% we may state that passing from not being a spinoff to being a spinoff is significantly different and lowers the probability to implement radical innovation. The extent to which such reduction occurs is provided by the value of the odds ratio, that is 0.461 and which is calculated in this way:

$$e^{\beta_{\text{spinoff}}} = e^{-0.774} = \frac{odds(spinoff)}{odds\ (not\ spinoff)} = \frac{\frac{P(radical\ innovation|X = spinoff)}{P(incremental\ innovation\ |X = spinoff)}}{\frac{P(radical\ innovation|X = not\ spinoff)}{P(incremental\ innovation|X = not\ spinoff)}}$$

The significant odds ratio lower than 1 emphasizes that being a spinoff is associated with a lower probability to implement radical innovation than other *de novo* ventures. Specifically, the odds of implementing radical innovation for a spinoff is 0.46 times the odds for a non-spinoff or, equivalently, non-spinoff odds are almost twice the ones of a spinoff. Thereby, when a new venture is a spinoff, it is less likely to implement a radical innovation strategy with respect to a firm whose entrepreneur does not have industry-specific experience.

While the focus of our analysis is to understand the kind of innovation implemented by a spinoff with respect to other market entrants, it is likewise intriguing to evaluate how other characteristics can influence the probability to implement innovation in *de novo* firms. From Table 4.3 we can infer that only two predictors can impact significantly, albeit with different confidence levels, on the probability of radical innovation. The first element is the reliance on incubators and/or science parks for starting the business which exhibits a strong significance (observed pvalue lower than 0.05 and lower than 0.01) and a positive estimate of the coefficient, with a consequent odds ratio higher than 1 (3.787). These results indicate that relying on incubators and/or science parks for establishing a new venture increases the probability of implementing radical innovation and the odds for entrepreneurs who relied on these ecosystem elements are 2.79 higher than the ones for entrepreneurs who did not rely on them.

The second element is the operativity in the secondary sector of the economy: with a confidence level of 90%, we may state that the probability to implement radical innovation is higher for firms which are active in manufacturing industries rather than for firms belonging to tertiary sector, since the coefficient associated with the variable *secondary* is positive, the corresponding pvalue (0.051) is lower than 0.10 and the odds ratio is higher than 1 (2.973). The inclination of manufacturing industries towards innovation emerges also from the latest Eurostat reports, according to which the 65% of private sector investments in R&D (a common proxy for innovation) comes from manufacturing firms.

For assessing the goodness of the entire model, the Hosmer-Lemeshow test is usually performed and it is based on the null hypothesis that the retrieved model fits well the data (Field, 2009). In this first logistic regression, we obtained a significance value equal to 0.256 which makes us accept the null hypothesis of goodness of the model for the data. With the same aim, some other basic statistics on residuals were obtained such as Cook's distance, standardized residuals and DFbetas which should show respectively a maximum value of 1, maximum 5% of them should lie outside ±1.96 and should be less than 1 (Field, 2009). Residuals statistics in our case are a Cook's distance maximum equal to 0.67, 3% of standardized residuals lie outside ±1.96 and DFbetas are all lower than 1. Therefore, the model fits well the data and no assumption was violated.

Consequently, from the first logistic model it can be affirmed that:

Spinoffs are less likely to implement radical innovation with respect to other firms whose entrepreneurs lack industry-specific experience and which cannot be classified as spinoffs.

4.4.3.2 Logit 2: industry-specific experience, experience in other industries and no experience

This section explores the difference in innovation not only between firms with industry-specific background or no industry-specific background, but also by discerning among firms lacking industry-specific experience those which accumulated no experience at all and those which accumulated experience in other industries. Therefore, the entire sample -195 *de novo* ventures- was used, by distinguishing three categories of *de novo* entrants and by making pair comparisons among them to catch their inclination towards radical innovation.

Since three categories of firms are analysed, only two dummies representing such categories should be included per model. Therefore, two different logistic regressions are performed whose difference relies on the baseline category used for the variable of interest. In the first model (A) the baseline category is the variable *spinoff* and therefore the innovation behaviour of the two market entrants lacking industry-specific experience are compared to the one of spinoffs, while in the second model (B) the variable *no.exper* is the reference category and therefore the difference in innovation implementation of firms with working experience are compared to the one of inexperienced firms.

A) Logit model. Baseline category: *spinoff*

In this model, the logistic function is the following one:

$$p(X) = \frac{e^{\beta_0+\beta_1 \text{no.exper}+\beta_2 \text{no.ind.spec}+\beta_3 \text{size}+\beta_4 \text{location}+\beta_5 \text{incu.scipark}+\beta_6 \text{org.assoc}+\beta_7 \text{network}+\beta_8 \text{partner}+\beta_9 \text{funds}+\beta_{10} \text{primary}+\beta_{11} \text{secondary}}}{1 + e^{\beta_0+\beta_1 \text{no.exper}+\beta_2 \text{no.ind.spec}+\beta_3 \text{size}+\beta_4 \text{location}+\beta_5 \text{incu.scipark}+\beta_6 \text{org.assoc}+\beta_7 \text{network}+\beta_8 \text{partner}+\beta_9 \text{funds}+\beta_{10} \text{primary}+\beta_{11} \text{secondary}}}$$

and the output of the logistic regression is reported in Table 4.4. Moreover, no significant multicollinearity appears, since the correlation coefficients associated with each pair of independent variable is not higher than the threshold 0.8 and each predictor shows VIF lower than 10[17].

Table 4. 4: Second logistic regression (spinoff as baseline): difference in innovation between spinoffs and other entrants. Source: SPSS and personal elaboration

	Coefficient	Standard error	Sig. Pr(>|z|)	Exp(β): Odds ratio	95% C.I.for EXP(B)	
					Lower	Upper
no.exper	0.016	0.754	0.983	1.016	0.232	4.451
no.ind.spec	0.945	0.389	0.015	2.573	1.200	5.518
size	0.038	0.049	0.432	1.039	0.944	1.144
location	-0.019	0.415	0.964	0.981	0.435	2.214
incu.scipark	1.303	0.382	0.001	3.681	1.740	7.785
org.assoc	-0.703	0.583	0.227	0.495	0.158	1.551
network	-1.038	0.753	0.168	0.354	0.081	1.549
partner	-0.298	0.435	0.494	0.742	0.317	1.742
funds	-0.110	0.370	0.767	0.896	0.434	1.851
primary	0.023	0.582	0.969	1.023	0.327	3.203
secondary	1.110	0.557	0.046	3.036	1.020	9.038
Constant	-0.384	0.800	0.631	0.681		

First, we analyse the output referred to the categorical variable *experience*, whose reference category is the variable *spinoff*, and from which the dummies *no.exper* and *no.ind.spec* are built. The estimate of the coefficient associated with the dummy *no.exper* is 0.016 and since it is a positive estimate, it seems that when entrepreneur is inexperienced there is a higher probability to implement radical innovation than when the new venture is a spinoff. However, since the observed significance level related to this variable is very high (0.983) and in particular higher

[17] Multicollinearity statistics available at Appendix F

than 0.05 and the critical odds ratio equal to 1 belongs to the confidence interval at 95%, we may state with a confidence level of 95% that the variable *no.exper* is insignificant and therefore passing from industry-specific experience to no experience at all has no effect on the probability to implement radical innovation. Consequently, there is no significant difference in the type of innovation implemented at the start-up phase between a spinoff and a *de novo* venture whose entrepreneur is totally inexperienced.

A different consideration can be made referring to the variable *no.ind.spec*: the estimate of the coefficient is equal to 0.945 and since it is a positive number, it seems that when an entrepreneur had previous working experience, albeit not industry-specific, she has a higher probability to implement radical innovation than in the case in which the new venture is a spinoff. However, to determine whether the relationship is significant, a look should be taken at observed significance level and confidence interval. As the observed pvalue (0.015) is lower than the critical pvalue 0.05 and as the confidence interval for the odds ratio at 95% does not include the critical value equal to 1, we can set forth that there is a significant difference in the probability to implement radical innovation by entrepreneurs who had previous working experience in other industries with respect to entrepreneurs with industry-specific experience, with a confidence level of 95%.

The extent to which such probability is different is given by the odds ratio, which in this case is calculated as:

$$e^{\beta_{\text{no.ind.spec}}} = e^{0.945} = \frac{odds(no\ industry - specific\ experience)}{odds\ (industry - specific\ experience)} =$$

$$= \frac{\frac{P(radical\ innovation|X = no\ industry - specific\ experience)}{P(incremental\ innovation\ |X = no\ industry - specific\ experience)}}{\frac{P(radical\ innovation|X = spinoff)}{P(incremental\ innovation|X = spinoff)}}$$

Since the odds ratio results significant and equal to 2.573, the odds of implementing radical innovation for *de novo* ventures whose entrepreneurs do not have previous industry-specific experience are 2.573 times the odds of implementing radical innovation for spinoffs, that is the odds of the former are 1.573 times higher than the odds of the latter.

These results can draw an interesting conclusion: held equal all the other variables, a new venture with no previous industry-specific experience is more likely to implement radical innovation with respect to a spinoff. Therefore, the intellectual capital that spinoffs inherit can be a cumbersome endowment for innovation at the start-up phase.

By observing Table 4.4, we can grasp other elements which influence the probability to implement radical innovation in *de novo* firms. As in the first model, the reliance on incubators and/or science parks increases the likelihood to implement radical innovation with respect to firms who do not rely on them, since the odds are 2.68 times higher and such relationship is significant (observed pvalue= 0.001, confidence interval at 95% does not include the value 1). Furthermore, with a confidence level of 95% operating in secondary sector rather than in tertiary one increases the probability to implement radical innovation since the odds are 2.04 times higher.

For what concerns all the other variables included into the logistic regression, we can affirm that such predictors do not influence the probability to implement radical innovation.

Hosmer-Lemeshow test results in a pvalue equal to 0.932, therefore we accept the hypothesis that the model fits very well the data; then, Cook's distance is equal to 0.77, 2% of residuals lie outside ±1.96 and all the Dfbetas are lower than 1. Thus, the model fits the data and no assumption is violated. What instead emerges is that:

Spinoffs are less likely to implement radical innovation at the start-up phase with respect to ventures founded by entrepreneurs with experience, albeit not industry-specific. Nonetheless, spinoffs are not neither more likely nor less likely to implement radical innovation with respect to ventures founded by totally inexperienced entrepreneurs.

B) Logit model. Baseline category: *no.exper*

Although the core of this work is on spinoffs and on their attitude towards innovation with respect to other entrants, it is likewise interesting to briefly look over the degree of innovation of other market entrants with respect to each other. In other words, our model has proved that spinoffs tend to implement a lower degree of innovation than firms with experience not industry-specific while they cannot be deemed more or less innovative than *de novo* firms totally inexperienced. What is still unresolved is therefore: do firms whose entrepreneurs are totally wet behind the ears implement radical innovation with respect to firms founded by entrepreneurs with some previous working experience even though not industry-specific?

To answer to this question we repeated the logistic regression, this time using *no.exper* as baseline category for the variable *experience* instead of the reference category *spinoff* employed in model *A)* and by keeping all the other variables equal.

Therefore, the logistic function employed for this regression is the following one:

$$p(X) = \frac{e^{\beta_0+\beta_1\text{spinoff}+\beta_2\text{no.ind.spec}+\beta_3\text{size}+\beta_4\text{location}+\beta_5\text{incu.scipark}+\beta_6\text{org.assoc}+\beta_7\text{network}+\beta_8\text{partner}+\beta_9\text{funds}+\beta_{10}\text{primary}+\beta_{11}\text{secondary}}}{1+e^{\beta_0+\beta_1\text{spinoff}+\beta_2\text{no.ind.spec}+\beta_3\text{size}+\beta_4\text{location}+\beta_5\text{incu.scipark}+\beta_6\text{org.assoc}+\beta_7\text{network}+\beta_8\text{partner}+\beta_9\text{funds}+\beta_{10}\text{primary}+\beta_{11}\text{secondary}}}$$

Where we can see that the dummy variable associated with β_1 is *spinoff*, while in the previous logistic was *no.exper* and all the other variables are kept equal.

Results from this logistic regression are exhibited in Table 4.5; given correlation coefficients lower than 0.8 and VIF values lower than 10, no evidence of multicollinearity is observed.[18]
As it can be observed in Table 4.5, coefficients, pvalues and confidence interval for odds ratio at the level 95% associated with the controls preserve the same values reported in Table 4.4 and therefore the only significant relationship is for *incu.scipark* and *secondary*. The exceptions are indeed represented by values related to the variables of interest, that is *spinoff* and *no.ind.spec*, since the baseline category has changed.

Table 4. 5: Second logistic regression (no.exper as baseline): difference in innovation among new entrants. Source: SPSS and personal elaboration

	Coefficient	Standard error	Sig. Pr(>|z|)	Exp(β): Odds ratio	95% C.I.for EXP(B)	
					Lower	Upper
spinoff	-0.016	0.754	0.983	0.984	0.225	4.309
no.ind.spec	0.929	0.785	0.237	2.531	0.543	11.799
size	0.038	0.049	0.432	1.039	0.944	1.144
location	-0.019	0.415	0.964	0.981	0.435	2.214
incu.scipark	1.303	0.382	0.001	3.681	1.740	7.785
org.assoc	-0.703	0.583	0.227	0.495	0.158	1.551
network	-1.038	0.753	0.168	0.354	0.081	1.549
partner	-0.298	0.435	0.494	0.742	0.317	1.742
funds	-0.110	0.370	0.767	0.896	0.434	1.851
primary	0.023	0.582	0.969	1.023	0.327	3.203
secondary	1.110	0.557	0.046	3.036	1.020	9.038
Constant	-0.368	1.085	0.735	0.692		

By analysing these two variables, we may erroneously infer from the negative and positive signs of estimates of coefficients that being respectively a spinoff and a *de novo* venture with experienced founder, albeit not industry-specific, lowers and increases the probability to introduce radical innovation than a *de novo* firm with inexperienced founder. However, pvalues and confidence interval make us accept the hypothesis that, with a confidence level of 95%, ventures founded by wet behind the ears entrepreneurs do not innovate neither more nor less than spinoffs and other firms which lack industry-specific experience, even though they accumulated some degree of working experience.
It is also worth noting that the insignificant difference between the reference category used in this regression for the variable *experience* and the variable *spinoff* was proved in the previous logistic regression as well (Table 4.4); this logistic regression is useful insofar it allows the discrimination on the type of innovation implemented by firms lacking industry-specific experience even though in this case no significant difference appears. In this model assumptions on logit are not violated: observed pvalue equal to 0.932 in Hosmer-Lemeshow test suggests

[18] Multicollinearity statistics available at Appendix F

that the model fits well the data; then, Cook's distance is 0.77, 2% of residuals lie outside ±1.96 and no Dfbeta is higher than 1.

Therefore, we can conclude that:

> De novo *ventures with inexperienced entrepreneurs are not neither more likely nor less likely to implement radical innovation at the start-up phase with respect to firms whose entrepreneurs had previous working experience in industries other than the one in which their business is established.*

4.4.3.3 Logit 3: experienced firms, industry-specific or not

The logistic regression developed in this section is based on a sample of 180 *de novo* firms, thereby excluding firms founded by totally inexperienced entrepreneurs. As mentioned earlier (see paragraph 4.4.1), focussing only on experienced firms provides an analysis more coherent with what usually comes about, since the bulk of firms are usually founded by experienced entrepreneurs. However, we expect also from this logistic regression that spinoffs are less likely to implement innovation than firms with experience in other industries.

The logistic function employed in this regression is thereby:

$$p(X) = \frac{e^{\beta_0+\beta_1\text{spinoff}+\beta_2\text{size}+\beta_3\text{location}+\beta_4\text{incu.scipark}+\beta_5\text{org.assoc}+\beta_6\text{network}+\beta_7\text{partner}+\beta_8\text{funds}+\beta_9\text{primary}+\beta_{10}\text{secondary}}}{1+e^{\beta_0+\beta_1\text{spinoff}+\beta_2\text{size}+\beta_3\text{location}+\beta_4\text{incu.scipark}+\beta_5\text{org.assoc}+\beta_6\text{network}+\beta_7\text{partner}+\beta_8\text{funds}+\beta_9\text{primary}+\beta_{10}\text{secondary}}}$$

Results from the logistic regression are reported in Table 4.6. Correlations among independent variable are also lower than the threshold 0.8 and no predictor has a VIF higher than 10[19], therefore no multicollinearity is evident.

Table 4. 6: Third logistic regression: difference in innovation between only experienced new entrants. Source: SPSS and personal elaboration

	Coefficient	**Standard error**	**Sig. Pr(>|z|)**	**Exp(β): Odds ratio**	**95% C.I.for EXP(B)**	
					Lower	**Upper**
spinoff	-0.923	0.385	0.017	0.397	0.187	0.845
size	0.031	0.050	0.540	1.031	0.935	1.137
location	-0.058	0.413	0.888	0.943	0.420	2.121
incu.scipark	1.199	0.388	0.002	3.318	1.551	7.100
org.assoc	-0.670	0.579	0.248	0.512	0.164	1.594
network	-1.146	0.766	0.135	0.318	0.071	1.427
partner	-0.169	0.440	0.702	0.845	0.356	2.003
funds	-0.059	0.380	0.877	0.943	0.448	1.986
primary	-0.029	0.581	0.960	0.971	0.311	3.033
secondary	0.721	0.599	0.228	2.057	0.636	6.653
Constant	0.713	0.849	0.402	2.039		

Table 4.6 confirms the expected results, as the estimate of the coefficient associated with the variable *spinoff* is negative (-0.923) and significant (0.017<0.05). Therefore, with a confidence

[19] Multicollinearity statistics available at Appendix F

level of 95% we can affirm that spinoffs are less likely to implement radical innovation with respect to other experienced firms which lack the industry-specific dimension.

The magnitude is provided by the odds ratio, computed as:

$$e^{\beta_{\text{spinoff}}} = e^{-0.923} = \frac{odds(spinoff)}{odds\ (experience, not\ industry - specific)} =$$

$$= \frac{\frac{P(radical\ innovation|X = spinoff)}{P(incremental\ innovation\ |X = spinoff)}}{\frac{P(radical\ innovation|X = experience, not\ industry - specific)}{P(incremental\ innovation|X = experience, not\ industry - specific)}}$$

which results equal to 0.397, that is the odds of implementing radical innovation being a spinoff are 0.397 times the odds of such implementation being an experienced entrant lacking industry-specificity; equivalently, the odds of implementing radical innovation for firms founded by entrepreneurs with previous experience even though not industry-specific are 1.52 times higher the odds of a spinoff and such amount almost reflects previous findings (see paragraph 4.4.3.2).

For what concerns other predictors, a very strong significance is observed for *incu.scipark* whose magnitude is similar to previous logistic regressions and likewise it translates into a higher likelihood of implementing radical innovation when an entrepreneur relies on science parks and/or incubators for establishing her new business than if she does not employ them. Differently from previous models, instead, the variable *secondary* has no significant effect on the likelihood to implement radical innovation (observed pvalue = 0.281).

Hosmer-Lemeshow test highlights that the model fits well the data (pvalue equal to 0.945); moreover, statistics on residuals are the following: Cook's distance equal to 0.56, residuals in 2.23% of the cases lie outside ±1.96 and no Dfbeta is higher than 1.

Since no assumption has been violated, our conclusion is that:

> *Between two* de novo *ventures founded by experienced entrepreneurs, the one classified as* spinoff *is less likely to implement radical innovation than the other one at the start-up phase.*

4.5 Theoretical, managerial and policy implications

Our study has highlighted that during the start-up phase spinoffs are less likely to implement radical innovation than other start-ups with experienced entrepreneurs; furthermore, *de novo* firms whose entrepreneurs lack any kind of experience in our sample cannot be deemed neither more innovative nor less innovative than spinoffs, a further proof to debunk the innovative aptitude put forward by the "garage belief". These conclusions confirm our expectations and reconcile with the theory on spinoffs as children whose intellectual capital is inherited by parent firms (Klepper, 2001); in so far as the intellectual capital between the parent and the new firm is overlapping, the less innovative will be the latter (Furlan, 2016b). This is particularly true for spinoffs which by definition inherit and implement practices and routines to compete in the same industry of their parents. Thereby, previous working experience is not only the key driver for entrepreneurship phenomenon but also for understanding the innovative behaviour of new firms at entry, whereby industry-specificity is the determinant to explain such behaviour.

When established, spinoffs may innovate less radically than other firms whose entrepreneurs accumulated previous working experience, since the former tend either to partially replicate parents' outputs because of intellectual capital legacy or to implement slight improvements in order to fulfil customers' needs that they seized as new business opportunities related to the industry in which they worked. On the other hand, experienced entrepreneurs with no industry-specific knowledge are less influenced by parental inheritance and, at the same time, they acquainted with different contexts which make them rely on a network of multiple ties that span different industries and markets. This may translate into the impossibility to replicate *pari passu* parents' practices that do not fit the environment, as the industry is totally different; furthermore, multiple ties reflect into multiple knowledge contributions that can foster more innovation (Furlan and Grandinetti, 2016). Entrepreneurs totally inexperienced instead lack both market and technical knowledge which are necessary to the firm to implement some innovation, but they do not suffer any kind of parent legacy which would force them into a less extent of innovation; for this reason, we think that the comparison on innovation between spinoffs and totally inexperienced entrepreneurs is not straightforward and it is expressed by the insignificant difference on innovation implemented by these two latter entrants shown by the logistic regressions.

In order to support innovation at the start-up phase in *de novo* ventures, a founding team with different skills and background is required. The diversity of founders' experiences can indeed lead to innovation in a new business, with respect to a situation in which entrepreneurs have homogeneous experiences or, out of the frying pan into the fire, when there is a single founder. On the whole, we suggest that policies to bootstrap spinoffs can benefit the whole economy as a spinoff creation is not a zero-sum game (Klepper, 2009). Nevertheless, the lower likelihood to implement radical innovation that characterises spinoffs during the start-up phase can curb the innovation pattern of the economy. Policy makers can deter this consequence by fostering infrastructure for innovation such as incubators and science parks which have proved to trigger more radical innovation when entrepreneurs rely on them for establishing a new business, rather than when they are not employed by new founders. Likewise, policies for the development of new firms whose entrepreneurs lack industry-specific experience are desirable since such businesses appeared to be the engine for radical innovation. Finally, since entrepreneurs rely upon social capital when they build a new firm (Audia and Rider, 2005) and relationships in different industries can be conducive to innovation (Furlan, 2016b), policies should nurture those contexts in which different people with different experience background interrelate with each other in order to steer a more radical innovation.

4.6 Limitations

This study fails to consider the different degrees of innovation implemented by *de novo* firms since only two types of implementation are assumed, namely incremental and radical innovation and when no innovation was implemented, the start-up was considered as incremental innovator. Moreover, the distinction between the two types is not measured by objective indicators such as R&D investments or R&D employees which are usually employed for quantifying innovation activities, despite their limits. For classifying the type of innovation implemented by sample firms, we considered founders' entry choices according to the answers provided which can be biased because result from a personal assessment. Future research may consider the percentage on revenues coming from innovative products or R&D investments in order to determine the innovative behaviour. Furthermore, our study focusses only on innovation implemented at the start-up phase and fails to consider subsequent innovations; thus, it may be plausible that spinoffs, that are found to be less radically innovative than other start-ups at the entry phase, can introduce radical innovations in following periods. Finally, the role of social capital is not considered: new firms were classified as *spinoffs* if their founders accumulated previous working experience in the same industry, but entrepreneurs were not discerned according to whether they garnered social ties also in other industries or not; this consideration was taken for granted in the case of entrepreneurs with experience, albeit not industry-specific, since they accumulated knowledge and relationships in other industries, by definition; this may be a starting point for future research to shed light on the extent to which social capital can effectively raise the likelihood of radical innovation and interesting results may derive.

4.7 Conclusion

The analysis reported in this chapter contributes to the literature on spinoffs, by trying to fill the existing gap on the relationship between spinoffs and innovation. The myth of young inexperienced dropouts who found new innovative businesses is dispelled by this analysis, thereby confirming previous empirical studies. What is more, our findings underline that the bulk of new businesses is founded by entrepreneurs with previous industry-specific experience (spinoffs) whereas only a small percentage of *de novo* venture is established by wet behind the ears entrepreneurs. Furthermore, spinoffs are found to be less innovative during the start-up phase with respect to other experienced entrants which lack industry-specific knowledge.

CONCLUSIONS

The goals of this thesis were to provide a systematization of the literature on entrepreneurial spinoffs and to try to solve one of the most debated topics provided by this literature.

Our starting point was the seminal paper of Klepper disclosed in 2009 as it represented the latest review and synthesis on what was dealt with spinoffs up to that date. The aforementioned review sets forth some *stylized facts* on spinoffs, among which their outperformance in terms of survival rate with respect to other market entrants, their higher likelihood to be spawned by industry leading incumbents and their implications for economy. It is worth mentioning that such patterns condense conclusions of different studies conducted on spinoffs, which range from more managerial oriented ones, towards more industrial focussed ones. The heterogeneity of studies has led to a cross-disciplinary literature and insofar these new market entrants represent not only the bulk of new firms in different industries but also an engine for economic growth, topics already investigated and themes not yet discovered need to be clearly identified. In order to do that, we performed a systematization of the literature after 2009, that is *post* Klepper's latest review, which has never been carried out, to the best of our knowledge. By means of bibliometric tools we demonstrated that spinoffs literature is composed by different subfields. Among the techniques used, citation and co-citation analyses were performed since they assume that more related papers, thus papers more co-cited, are more likely to delve into similar topics and to belong to the same thematic cluster. Statistical analyses such as Principal Components Analysis, Cluster Analysis and Multidimensional scaling were employed along with personal assessment to identify the subfields of research, whose labels were assigned by employing the output of a word frequency analysis. Finally, six subdomains of the literature on spinoffs were detected and they concerned with different topics: different types of spinoffs and comparison with other entrants in terms of features and performance, spinoffs and agglomeration economy inside clusters, the performance of spinoffs and their main characteristics also in terms of intellectual capital endowment, factors triggering spawning events both from an empirical and a theoretical viewpoint, the comparison between academic and entrepreneurial spinoffs and finally spinoffs and innovation. This latter subfield is however approached from different perspectives, among which the way through which spinoffs can innovate more, the extent of innovation between academic and entrepreneurial spinoffs, the kind of relationships between spinoffs and other actors which can steer more innovation. The six subfields are underpinned by *knowledge inheritance theory* and the *theory of disagreements*, which respectively assume that spinoffs' legacy is based on parents' intellectual capital and that spinoffs are established by employees after disagreements on strategic or managerial decisions with their employers. Notwithstanding this, these two pillars have been complemented by following studies and analyses which have questioned their absolute validity and have provided a richer knowledge on the topic.

Along with *stylized facts*, Klepper identified some unresolved issues which embodied avenues for future research and our analysis confirms that papers considered for literature systematization did not answer to some of them. Among these outstanding dilemmas, reasons behind spinoffs formation deserve more attention and, more specifically, the spawning event when parent firms behave hostilely: in this regard, the effects of non-compete agreements should be further investigated for designing *ad hoc* policies; then, the role of social capital and network should be more deeply analysed and its impact on cluster dynamics as well. Finally, although a thematic cluster on innovation has been identified, the extent to which industry-specific experience accumulated by spinoffs' founders could affect the degree of innovation implemented during the start-up phase by their new ventures is an elephant in the room: spinoffs literature *pre* and *post* Klepper's latest review has not clarified this relationship. To cope with this latter issue, we conducted an empirical analysis on a sample of 195 *de novo* firms founded by Alumni graduated at the University of Padova between 2000 and 2010. The bulk of new firms were categorized as *spinoffs*, confirming previous empirical studies on the prevalence of spinoffs as new ventures and the innovation that they implemented was compared to the one of other firms which lacked industry-specific experienced founders. The logistic regressions performed showed that spinoffs were more likely to implement an incremental type of innovation, that is either to implement an improvement of products or services or to provide them at a lower cost or to enter emerging market or to implement no innovation at all with respect to other start-ups founded by experienced entrepreneurs who did not accumulate industry-specific knowledge. On the other hand, entrepreneurs with experience not industry-specific were more likely to introduce radically modified outputs and/or provide totally new products or services which created a new market, that is implement a more radical innovation with respect to spinoffs. This may be explained by the fact that spinoffs tend to rely more on intellectual capital legacy that makes them either to replicate outputs or slightly improve them when spinoffs founders seize customers' needs related to the same industry to be fulfilled. On the contrary, entrepreneurs with experience not industry-specific accumulated market and technical knowledge of other markets and cross-industry social relationships that prevent a servile implementation of practices and routines inherited from the parents, also given the fact that such implementation may not be appropriate in a different industry. The lower reliance on parents' inheritance and the wider nature of relationship ties may offer an explanation to the higher likelihood of radical innovation implementation that characterises these market entrants with respect to spinoffs. Furthermore, inexperienced entrepreneurs represented only 8% of new entrepreneurs in our sample and they were not found to be neither more radical innovators nor more incremental innovators than spinoffs.

Moreover, relying on incubators or science parks for creating new firms increased the likelihood to implement radical innovation with respect to not relying on these ecosystem players and this is in line with recent European policies that bootstrap them.

The results then seem to suggest that while spinoffs are acknowledged as superior performers than other market entrants, they are not more innovative at the beginning and this has multiple implications. When the policy maker is more inclined towards the survival of new firms, she should foster spinoffs formation; when instead innovation becomes the priority, she should encourage either the formation of firms with founders who accumulated previous experience even though in different industries or the development of incubators and science parks that trigger more radical innovation.

Nevertheless, this thesis has some limits which can represent opportunities for future research. Systematization of spinoffs literature was performed starting from 2009 but it may be interesting to analyse the development of the field from its very beginning, that is to track the evolution of subfields over time. Moreover, our empirical analysis has tried to clarify the degree of innovation of spinoffs with respect to other entrants but only two types of innovation have been considered and were based on market entry choice; thereby, it may be helpful to consider several types of innovation, also based on objective measures such as sales from innovative products or R&D expenses and establish the degree of innovativeness of spinoffs not only at the start-up phase, but also during their entire life for more focused policies.

REFERENCES

ADAMS, P., FONTANA, R. and MALERBA, F., 2017. *Bridging Knowledge Resources: The Location Choices of Spinouts*. Strategic Entrepreneurship Journal, 11(2), pp. 93-121.

AGARWAL, R., CAMPBELL, B. A., FRANCO, A. M. and GANCO, M., 2016. *What Do I Take With Me? The Mediating Effect of Spin-out Team Size and Tenure on the Founder-Firm Performance Relationship*. Academy of Management Journal, 59(3), pp. 1060-1087.

AGARWAL, R., ECHAMBADI, R., FRANCO, A.M. and SARKAR, M., 2004. *Knowledge transfer through inheritance: spinout generation, development, and survival*. The Academy of Management Journal, 47(4), pp. 501-522.

AHLGREN, P., JARNEVING, B., and ROUSSEAU, R., 2003. *Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient*. Journal of the American Society for Information Science, 54(6), pp. 550–560.

ANDERSSON, M. and KLEPPER, S., 2013. *Characteristics and Performance of New Firms and Spinoffs in Sweden*. Industrial and Corporate Change, 22(1), pp. 245-280.

ANDERSSON, M., BALTZOPOULOS, A. and LÖÖF, H., 2012. *R&D strategies and entrepreneurial spawning*. Research Policy, 41(1), pp. 54-68.

ANTON, J. J. and YAO, D. A., 1995. *Start-ups, Spin-offs, and Internal Projects*. Journal of Law, Economics and Organization, 11(2), pp. 362-378.

APA, R., DE MARCHI, V., GRANDINETTI R. and SEDITA, S. R., 2016. *Oltre la visione tecnocratica dell'innovazione: i risultati di una ricerca sulle piccolo-medie imprese (Beyond a technocratic view of innovation: insights from an investigation on small and medium-sized enterprises)*. Economia e Società Regionale, 3, pp. 101-113.

ARGYRES, N. and MOSTAFA, R., 2016. *Knowledge Inheritance, Vertical Integration, and Entrant Survival in the Early U.S. Auto Industry*. Academy of Management Journal, 59(4), pp. 1474-1492.

AUDIA, P.G. and RIDER, C.I., 2005. *A Garage and an Idea: what more does an entrepreneur need?*. California Management Review, 48(1), pp. 6-28.

AZZALINI, A. and SCARPA, B., 2012. *Data Analysis and Data Mining: An Introduction*. New York: Oxford University Press, pp.213-222.

BACKHAUS, K., LÜGGER, K. and KOCH, M., 2011. *The structure and evolution of business-to-business marketing: A citation and co-citation analysis*. Industrial Marketing Management, 40(6), pp. 940-951.

BETTIOL, M. and FURLAN, A., 2014. Istruiti, esperti, autofinanziati: ritratto degli imprenditori italiani oltre i cliché. *Corriere Innovazione*, 10 ottobre. Available at: http://corriereinnovazione.corriere.it/sistemi/2014/7-ottobre-2014/istruiti-esperti-autofinanziati-startupper-italiani-oltre-cliche-230298503169.shtml?refresh_ce-cp [Date of access: 20/08/2017]

BORGATTI, S. P., EVERETT, M. G. and FREEMAN, L. C. 2002. *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.

BOSCHMA, R. and WETERINGS, A. B. R., 2005. *The effect of regional differences on the performance of software firms in the Netherlands.* Journal of Economic Geography, 5(5), pp. 567-588.

BOSCHMA, R., 2015. *Do spinoff dynamics or agglomeration externalities drive industry clustering? A reappraisal of Steven Klepper's work*. Industrial and Corporate Change, 24(4), pp. 859-873.

BRUNEEL, J., VAN DE VELDE, E. and CLARYSSE, B., 2013. *Impact of the Type of Corporate Spin-Off on Growth*. Entrepreneurship Theory and Practice, 37(4), pp. 943-959.

BUENSTORF, G., 2009. *Opportunity spin-offs and necessity spin-offs*. International Journal of Entrepreneurial Venturing, 1(1), pp. 22-40.

BUENSTORF, G., 2016. *Schumpeterian incumbents and industry evolution*. Journal of Evolutionary Economics, 26(4), pp. 823-836.

BUENSTORF, G., ENGEL, C., FISCHER, S. and GUETH, W., 2016. *Non-compete Clauses, Employee Effort and Spin-Off Entrepreneurship: A Laboratory Experiment*. Research Policy, 45(10), pp. 2113-2124.

BUENSTORF, G. and FORNAHL, D., 2009. *B2C—bubble to cluster: the dot-com boom, spin-off entrepreneurship, and regional agglomeration*. Journal of Evolutionary Economics, 19(3), pp. 349-378.

BUENSTORF, G. and KLEPPER, S., 2009. *Heritage and agglomeration: the Akron tyre cluster revisited*. The Economic Journal, 119(537), pp. 705-733.

BUENSTORF, G. and KLEPPER, S., 2010. *Why does entry cluster geographically? Evidence from the US tire industry*. Journal of Urban Economics, 68(2), pp. 103-114.

CALLAN, B., 2001. *Generating Spin-offs: Evidence from Across the OECD*. STI Review, 26(26 Special Issue), pp.14-54.

CAMUFFO, A. and GRANDINETTI, R., 2011. *Italian industrial districts as cognitive systems: Are they still reproducible?* Entrepreneurship & Regional Development, 23(9-10), pp. 815-852.

CHATTERJI, A. K., 2009. *Spawned with a silver spoon? Entrepreneurial performance and innovation in the medical device industry*. Strategic Management Journal, 30(2), pp. 185-206.

CHEN, C., IBEKWE-SANJUAN, F. and HOU, J., 2010. *The Structure and Dynamics of Co Citation Clusters: A Multiple Perspective Co-Citation Analysis*. Journal of the American Society for Information Science and Technology, 61(7), pp. 1386-1409.

CHEYRE, C., KLEPPER, S. and VELOSO, F. M., 2015a. *Spinoffs and the Mobility of U.S. Merchant Semiconductor Inventors*. Management Science, 61(3), pp. 487-506.

CHEYRE, C., KOWALSKI, J. and VELOSO, F. M., 2015b. *Spinoffs and the Ascension of Silicon Valley*. Industrial and Corporate Change, 24(4), pp. 837-858.

CHRISTENSEN, C. M., 1993. *The rigid disk drive industry: A history of commercial and technological turbulence*. Business History Review, 67(4), pp. 531-588.

CLARYSSE, B., WRIGHT, M. and VAN DE VELDE, E., 2011. *Entrepreneurial Origin, Technological Knowledge, and the Growth of Spin-Off Companies*. Journal of Management Studies, 48(6), pp. 1420-1442.

CORDES, C., RICHERSON, P. J. and SCHWESINGER, G., 2014. *A Corporation's Culture as an Impetus for Spinoffs and a Driving Force of Industry Evolution*. Journal of Evolutionary Economics, 24(3), pp. 689-712.

CULNAN, M. J., 1986. *The Intellectual Development of Management Information Systems, 1972-1982: A Co-Citation Analysis*. Management Science, 32(2), pp. 156-172.

CURRAN, D., O'GORMAN, C. and VAN EGERAAT, C., 2016a. *Inherited competence and spin-off performance*. European Planning Studies, 24(3), pp. 443-462.

CURRAN, D., O'GORMAN, C. and VAN EGERAAT, C., 2016b. "*Opportunistic" spin-offs in the aftermath of an adverse corporate event*. Journal of small business and enterprise development, 23(4), pp. 984-1008.

CUSMANO, L., MORRISON, A. and PANDOLFO, E., 2015. *Spin-Off and Clustering: A Return to the Marshallian District.* Cambridge Journal of Economics, 39(1), pp. 49-66.

DAHL, M. S. and SORENSON, O., 2012. *Home Sweet Home: Entrepreneurs' Location Choices and the Performance of Their Ventures.* Management Science, 58(6), pp. 1059-1071.

DAHL, M. S. and SORENSON, O., 2014. *The Who, Why, and How of Spinoffs*. Industrial and Corporate Change, 23(3), pp. 661-688.

DAHL, M.S. and REICHSTEIN, T., 2007. *Are you experienced? Prior experience and the survival of new organizations*. Industry and Innovation, 14(5), pp. 497-511.

DELGADO, M., PORTER, M. E. and STERN, S., 2014. *Clusters, convergence, and economic performance*. Research Policy, 43(10), pp. 1785–1799.

DELMAR, F. and SHANE, S., 2006. *Does experience matter? The effect of founding team experience on the survival and sales of newly founded ventures*. Strategic Organization, 4(3), pp. 215-247.

DENCKER, J. C., GRUBER, M. and SHAH, S. K., 2009. *Pre-entry knowledge, learning, and the survival of new firms*. Organization Science, 20(3), pp. 516-537.

DI GUARDO, M. and HARRIGAN, K. R., 2012. *Mapping research on strategic alliances and innovation: a co-citation analysis*. The Journal of Technology Transfer, 37(6), pp. 789-811.

DIODATO, V. P., 1994. *Dictionary of Bibliometrics*. New York: Haworth Press, 1994.

EESLEY, C. E., 2011. *Alumni surveys as a data collection methodology*. Working paper. Available at: http://www.stanford.edu/~cee/Papers/Eesley_Alumni_surveys.pdf [Date of access: 04/09/2017)

ERIKSSON, T. and KUHN, J.M., 2006. *Firm spin-offs in Denmark, 1981-2000 – patterns of entry and exit*. International Journal of Industrial Organization, 24(5), pp. 1021-1040.

FACKLER, D., SCHNABEL, C. and SCHMUCKER, A., 2016. *Spinoffs in Germany: characteristics, survival, and the role of their parents*. Small Business Economics, 46(1), pp. 93-114.

FERREIRA, M. P., REIS, N. R., PAULA, R. M. and PINTO, C. F., 2017. *Structural and longitudinal analysis of the knowledge base on spin-off research*. Scientometrics, 112(1), pp. 289-313.

FERRIANI, S., GARNSEY, E. and LORENZONI, G., 2012. *Continuity and change in a spin-off venture: the process of reimprinting*. Industrial and Corporate Change, 21(4), pp. 1011–1048.

FIELD, A., 2009. *Discovering statistics using SPSS*. 3rd edition. London: SAGE Publications Ltd.

FRANCO A. M. and FILSON D., 2000. *Knowledge Diffusion through Employee Mobility*. Federal Reserve Bank of Minneapolis, Staff Report 272. Available at: https://www.econstor.eu/bitstream/10419/94651/1/2000-61.pdf [Date of access: 03/09/2017]

FRYGES, H. and WRIGHT, M., 2014. *The origin of spin-offs: a typology of corporate and academic spin-offs.* Small Business Economics, 43(2), pp. 245-259

FRYGES, H., MÜLLER, B. and NIEFERT, M., 2014. *Job machine, think tank, or both: what makes corporate spin-offs different?* Small Business Economics, 43(2), pp. 369-391.

FURLAN, A. and GRANDINETTI, R., 2014. *Spin-off performance in the start-up phase – a conceptual framework*. Journal of Small Business and Enterprise Development, 21(3), pp. 528-544.

FURLAN, A. and GRANDINETTI, R., 2016. *Spinoffs and their endowments: beyond knowledge inheritance theory*. Journal of Intellectual Capital, 17(3), pp.570-589.

FURLAN, A., 2016a. *Who lives longer? Startups vs spinoffs founded as proprietorships*. International Journal of Entrepreneurial Behavior & Research, 22(3), pp. 416 – 435.

FURLAN, A., 2016b. *Genialità o esperienza? L'innovatività delle nuove imprese (Ingenuity or experience? the innovation of new ventures)*. Economia e Società Regionale, 3, pp. 68-83.

GARFIELD, E., 1979. *Is citation analysis a legitimate evaluation tool?* Scientometrics, 1(4), pp. 359-375.

GARFIELD, E., MALIN, M. V. and SMALL, H., 1978. *Citation data as science indicators*. In Elkana, Y. et al. (Eds), *Toward a metric of science: The Advent of science indicators*. New York: John Wiley and Sons, pp. 179-207.

GARFIELD, E., SHER, I. H., and TORPIE, R. J., 1964. *The use of citation data in writing the history of science.* Philadelphia: ISI Press.

GILSING, V. A., VAN BURG, E. and ROMME, A. G. L., 2010. *Policy principles for the creation and success of corporate and academic spin-offs*. Technovation, 30(1), pp. 12-23.

GOLMAN, R. and KLEPPER, S., 2016. *Spinoffs and clustering*. The RAND Journal of Economics, 47(2), pp. 341-365.

HABERSETZER, A., 2016. *Spinoff dynamics beyond clusters: pre-entry experience and firm survival in peripheral regions*. Entrepreneurship & Regional Development, 28(9-10), pp. 791-812.

HAIR, J. F., BLACK, W. C., BABIN, B. J. and ANDERSON, R. E., 2010. *Multivariate Data Analysis*. 2nd edition. Upper Saddle River, NJ: Pearson Prentice Hall.

HANNAN, M. T. and FREEMAN, J., 1977. *The Population Ecology of Organizations*. The American Journal of Sociology, 82(5), pp. 929-964.

HE, Y. and HUI, S. C., 2002. Mining a Web Citation Database for author co-citation analysis. *Information Processing and Management*, 38(4), pp. 491-508.

HEEBELS, B. and BOSCHMA, R., 2011. *Performing in Dutch book publishing 1880–2008: the importance of entrepreneurial experience and the Amsterdam cluster*. Journal of Economic Geography, 11(6), pp. 1007-1029.

HELFAT, C.E. and LIEBERMAN, M.B., 2002. *The birth of capabilities: market entry and the importance of pre-history*. Industrial and Corporate Change, 11(4), pp. 725-760.

HERVAS-OLIVER, J. L., LLEO, M. and CERVELLO, R., 2017. *The Dynamics of Cluster Entrepreneurship: Knowledge Legacy from Parents or Agglomeration Effects? The Case of the Castellon Ceramic Tile District*. Research Policy, 46(1), pp. 73-92.

HSIAO, C. H. and YANG, C., 2011. *The intellectual development of the technology acceptance model: A co-citation analysis*. International Journal of Information Management, 31(2), pp. 128-136.

HUISMAN, M. and VAN DUIJN, M. A. J., 2005. *Software for Social Network Analysis*. In Carrington, P. J., Scott, J. and Wasserman, S. (Eds.), *Models and Methods in Social Network Analysis*. New York: Cambridge University Press, pp. 270-316.

ITO, K., 1995. *Japanese spinoffs: Unexplored survival strategies*. Strategic Management Journal, 16(6), pp. 431–446.

JAMES, G., WITTEN, D., HASTIE, T. and TIBSHIRANI, R., 2013. *An introduction to statistical learning: With applications in R*. New York: Springer.

JO, S. J., JEUNG, C.-W., PARK, S. and YOON, H. J., 2009. *Who Is Citing Whom: Citation Network Analysis among HRD Publications from 1990 to 2007*. Human Resource Development Quarterly, 20(4), pp. 503-537.

KEUNING, D., 2007. *Management: A European Perspective*. 2nd edition. Groening: Wolters-Noordhoff, pp. 620-623.

KLEPPER, S. and SIMONS, K., 2005. *Industry shakeouts and technological change*. International Journal of Industrial Organization, 23(1-2), pp. 23-43.

KLEPPER, S. and SLEEPER, S., 2005. *Entry by spinoffs*. Management Science, 51(8), pp. 1291-1306.

KLEPPER, S. and THOMPSON P., 2005. *Spinoff entry in high-tech industries: Motives and consequences*. Available at: http://web.mit.edu/iandeseminar/Papers/Spring2006/Klepper.pdf [Date of access: 20/08/2017].

KLEPPER, S. and THOMPSON, P., 2010. *Disagreements and intra-industry spinoffs.* International Journal of Industrial Organization, 28(5), pp. 526-538.

KLEPPER, S., 2001. *Employee startups in high-tech industries*. Industrial and Corporate Change, 10(3), pp. 639-674.

KLEPPER, S., 2002. *The capabilities of new firms and the evolution of the US automobile industry*. Industrial and Corporate Change, 11(4), pp. 645-666.

KLEPPER, S., 2007. *Disagreements, spinoffs, and the evolution of Detroit as the capital of the US automobile industry*. Management Science, 53(4), pp. 616-631.

KLEPPER, S., 2009. *Spinoffs: A review and synthesis*. European Management Review, 6(3), pp.159-171.

KLEPPER, S., 2010. *The origin and growth of industry clusters: The making of Silicon Valley and Detroit*. Journal of Urban Economics, 67(1), pp. 15-32.

KLEPPER, S., 2011. *Nano-economics, spinoffs, and the wealth of regions*. Small Business economics, 37(2), pp. 141-154.

KRUSKAL, J. B., 1964. *Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis*. Psychometrika, 29(1), pp. 1-27.

LEE, M. R. and CHEN, T. T., 2012. *Revealing research themes and trends in knowledge management: From 1995 to 2010*. Knowledge-Based Systems, 28, pp. 47-58.

LEJPRAS, A. and STEPHAN, A., 2011. *Locational conditions, cooperation, and innovativeness: evidence from research and company spin-offs*. The Annals of Regional Science, 46(3), pp. 543-575.

LEJPRAS, A., 2014. *How innovative are spin-offs at later stages of development? Comparing innovativeness of established research spin-offs and otherwise created firms*. Small Business Economics, 43(2), pp. 327-351.

LEYDESDORFF, L. and VAUGHAN, L., 2006. *Co-occurrence matrices and their applications in information science: Extending ACA to the Web environment*. Journal of the American Society for Information Science and Technology, 57(12), pp. 1616-1628.

LÖÖF, H. and NABAVI, P., 2014. *Survival, productivity and growth of new ventures across locations*. Small Business Economics, 43(2), pp. 477-491.

MAYER, H., 2013. *Entrepreneurship in a Hub-and-Spoke Industrial District: Firm Survey Evidence from Seattle's Technology Industry*. Regional Studies, 47(10), pp. 1715-1733.

MCCAIN, K.W., 1990. *Mapping authors in intellectual space: A technical overview*. Journal of the American Society for Information Science, 41(6), pp. 433-443.

MCKENDRICK, D. G., WADE, J. B. and JAFFEE, J., 2009. *A good riddance? Spin-offs and the technological performance of parent firms.* Organization Science, 20(6), pp. 979-992.

MOED, H. F., 2005. *Citation analysis in research evaluation*. Dordrecht, UK: Springer.

MONTANA, S., 2017. *Exploring the Student Entrepreneurship Phenomenon: The Case of The University of Padova*. Thesis discussed at the University of Padova. A.Y. 2016-2017. Available at: http://tesi.cab.unipd.it/56590/1/Montana_Salvatore.pdf [Date of access: 07/10/2017].

MOORE, G. and DAVIS, K., 2004. *Learning the Silicon Valley way*. In Bresnahan, T. and Gambardella, A. (Eds.), *Building high-tech clusters: Silicon Valley and beyond*, Cambridge, UK: Cambridge University Press, pp. 7-38.

MUENDLER, M., RAUCH, J. E. and TOCOICAN, O., 2012. *Employee spinoffs and other entrants: Stylized facts from Brazil*. International Journal of Industrial Organization, 30(5), pp. 447-458.

NELSON, R. R. and WINTER, S. G., 1982. *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press.

NERUR, S. P., RASHEED, A. A. and NATARAJAN, V., 2008. *The Intellectual Structure of the Strategic Management Field: An Author Co-Citation Analysis*. Strategic Management Journal, 29(3), pp. 319-336.

NEWMAN, M. E. J., 2010. Networks: an introduction. Oxford, UK: Oxford University Press.

NORMAN, G. R. and STREINER, D. L., 2008. Biostatistics: The Bare Essentials. 3rd edition. Hamilton, ON: B.C. Decker.

ÖZÇINAR, H., 2015. *Mapping teacher education domain: A document co-citation analysis from 1992 to 2012*. Teaching and Teacher Education, 47, pp. 42-61.

PHAM, G.-K., 2011. *Radical Innovation and Open Innovation: Creating New Growth Opportunities for Business*. Hamburg: Diplomica Verlag GmbH.

PHILLIPS, D. J., 2002. *A genealogical approach to organizational life chances: the parent progeny transfer among Silicon Valley law firms, 1946-1996*. Administrative Science Quarterly, 47(3), pp. 474-506.

POTTER, A. and WATTS, H. D., 2011. *Evolutionary agglomeration theory: increasing returns, diminishing returns, and the industry life cycle*. Journal of Economic Geography, 11(3), pp. 417-455.

RAMOS-RODRÍGUEZ, A. and RUÍZ-NAVARRO, J., 2004. *Changes in the intellectual structure of strategic management research: A bibliometric study of the strategic management journal, 1980-2000*. Strategic Management Journal, 25(10), pp. 981-1004.

RAUCH, J. E., 2016. *Dynastic entrepreneurship, entry, and non-compete enforcement*. European Economic Review, 86, pp. 188-201.

ROCHA, V., CARNEIRO, A. and VARUM, C., 2015. *What Explains the Survival Gap of Pushed and Pulled Corporate Spin-Offs?* Economics Letters, 126, pp. 127-130.

SCARINGELLA, L., MILES, R. E. and TRUONG, Y., 2017. *Customers involvement and firm absorptive capacity in radical innovation: The case of technological spin-offs*. Technological Forecasting and Social Change, 120, pp. 144-162.

SCHUMPETER J.A. (1934). *The Theory of Economic Development*. Cambridge, MA: Harvard University Press, pp. 61-116.

SHANE, S. and VENKATARAMAN, S., 2000. *The Promise of Entrepreneurship as a Field of Research*. Academy of Management Review, 25(1), pp. 217-226.

SMALL, H., 1973. *Co-citation in the scientific literature: A new measure of the relationship between two documents.* Journal of the American Society for Information Science, 24(4), pp. 265-269.

SMALL, H., and GRIFFITH, B. C., 1974. *The structure of scientific literatures I: Identifying and graphing specialties*. Science studies, 4(1), pp. 17-40.

SORENSON, O. and AUDIA, P. G., 2000. *The Social Structure of Entrepreneurial Activity: Geographic Concentration of Footwear Production in the United States, 1940–1989*. American Journal of Sociology, 106(2), pp. 424-462.

SORKUN, M. F. and FURLAN, A., 2017. *Product and Organizational Modularity: A Contingent View of the Mirroring Hypothesis*. European Management Review, 14(2), pp. 205-224.

STEENHUIS, H. and KIEFER, D., 2016. *Early stage cluster development: a manufacturers-led approach in the aircraft industry*. Competitiveness Review, 26(1), pp. 41-65.

STINCHCOMBE, A. L., 1965. *Social structure and organizations*. In March, J. G. (Ed.), *Handbook of Organizations*. Chicago: Rand McNally, pp. 142-193.

THELWALL, M., 2008. *Bibliometrics to webometrics*. Journal of Information Science, 34(4), pp. 605-621.

THOMPSON, P. and CHEN, J., 2011. *Disagreements, employee spinoffs and the choice of technology*. Review of Economic Dynamics, 14(3), pp. 455-474.

VAN ECK, N. J. and WALTMAN, L., 2014. *Visualizing bibliometric networks*. In Ding, Y., Rousseau, R. and Wolfram, D. (Eds.), *Measuring scholarly impact: Methods and practice*. Cham: Springer, pp. 285–320

WALLIN, M.W., 2012. *The bibliometric structure of spin-off literature*. Innovation, 14(2), pp. 162-177.

WALTER, S. G., HEINRICHS, S. and WALTER, A., 2014. *Parent hostility and spin-out performance*. Strategic Management Journal, 35(13), pp. 2031-2042.

WENNBERG, K., WIKLUND, J. and WRIGHT, M., 2011. *The effectiveness of university knowledge spillovers: Performance differences between university spinoffs and corporate spinoffs*. Research Policy, 40(8), pp. 1128-1143.

WENTING, R., 2008. *Spinoff dynamics and the spatial formation of the fashion design industry, 1858-2005*. Journal of Economic Geography, 8(5), pp. 593-614.

WHITE, H. D. and GRIFFITH, C., 1981. *Author cocitation: A literature measure of intellectual structure*. Journal of the American Society for Information Science, 32(3), pp. 163–171.

WIGGINS, S. N., 1995. *Entrepreneurial Enterprises, Endogenous Ownership, and the Limits to Firm Size*. Economic Inquiry, 33(1), pp. 54-69.

YEGANEGI, S., LAPLUME, A. O., DASS, P. and HUYNH, C., 2016. *Where do spinouts come from? The role of technology relatedness and institutional context*. Research Policy, 45(5), pp. 1103-1112.

WEB RESOURCES

http://www.analytictech.com/borgatti/mds.htm (Ucinet guide on Multidimensional Scaling) [Date of access: 14/07/2017]

http://www.analytictech.com/ucinet/documentation/quickstart.pdf (Ucinet quick start guide). [Date of access: 16/06/2017]

http://www.analytictech.com/ucinet/help/c_kchz.htm (Ucinet guide on Principal Components Analysis) [Date of access: 14/07/2017]

http://www.analytictech.com/ucinet/help/3j.x0e.htm (Ucinet guide on Cluster analysis) [Date of access: 14/07/2017]

https://ec.europa.eu (European Commission website, from which information on female entrepreneurship, number of workers in European firms, access to external capital, initiatives to support innovation in Eurozone and innovation statistics was obtained) [Date of access: 26/10/2017]

http://www.istat.it/it/ (Istat website from which information on employment per sector was obtained) [Date of access: 26/10/2017]

http://www.restore.ac.uk/srme (National Centre for Research Method website, from which information on how to run logistic regression by using SPSS was obtained) [Date of access: 06/10/2017]

https://www.scopus.com/search/form.uri?display=basic (Scopus website for articles retrieval) [Date of access: 26/07/2017]

https://www.spss-tutorials.com/basics/ (Tutorial on SPSS to get acquainted with this software) [Date of access: 20/08/2017]

http://www.statisticssolutions.com/assumptions-of-logistic-regression/ (Brief explanation on Logistic regression's assumptions) [Date of access: 06/10/2017]

http://www.vosviewer.com/ (VosViewer website, from which information on this software was obtained) [Date of access: 09/09/2017]

APPENDIX

APPENDIX A

PAPERS	PAPERS															
	#2	#3	#4	#5	#7	#8	#10	#11	#12	#13	#15	#16	#18	#19	#20	#21
#2	12.5	0	6	14	0	1	3	5	1	1	0	0	5	0	0	2
#3	0	3	0	0	3	1	0	0	0	0	2	0	0	0	0	0
#4	6	0	12	13	1	1	0	4	1	5	2	1	2	0	3	3
#5	14	0	13	19.5	1	6	12	7	1	3	1	0	8	0	2	3
#7	0	3	1	1	12	1	0	0	2	1	19	1	2	1	1	2
#8	1	1	1	6	1	6.5	1	2	1	0	0	0	0	0	0	0
#10	3	0	0	12	0	1	9.5	2	0	0	0	0	4	1	0	1
#11	5	0	4	7	0	2	2	8	1	0	0	0	2	0	0	2
#12	1	0	1	1	2	1	0	1	3.5	0	3	0	0	1	2	1
#13	1	0	5	3	1	0	0	0	0	4.5	1	1	1	0	0	1
#15	0	2	2	1	19	0	0	0	3	1	12.5	2	3	1	1	2
#16	0	0	1	0	1	0	0	0	0	1	2	2.5	1	0	0	2
#18	5	0	2	8	2	0	4	2	0	1	3	1	8.5	0	0	1
#19	0	0	0	0	1	0	1	0	1	0	1	0	0	1.5	1	0
#20	0	0	3	2	1	0	0	0	2	0	1	0	0	1	4	2
#21	2	0	3	3	2	0	1	2	1	1	2	2	1	0	2	4.5
#22	0	0	2	2	1	0	0	0	1	0	1	0	0	1	3	1
#24	1	0	0	2	0	1	3	1	0	0	1	0	1	0	1	1
#26	0	0	2	2	0	0	0	0	0	0	0	0	0	0	2	1
#27	2	0	0	2	0	1	1	2	1	0	0	0	1	0	0	0
#28	2	0	2	4	0	1	1	1	0	1	0	0	4	0	1	3
#29	0	0	1	1	1	0	0	0	1	0	1	0	0	1	2	1
#30	0	0	2	2	2	0	0	0	0	0	1	0	1	0	2	1
#31	1	0	0	3	0	5	0	1	1	0	0	0	0	0	0	1
#33	0	0	1	1	0	0	0	0	0	0	1	0	0	0	1	0
#34	1	0	0	1	0	0	1	1	0	0	0	0	1	0	0	1
#36	1	0	0	2	0	1	1	1	0	0	0	0	1	0	0	1
#37	2	0	1	4	0	2	2	3	0	1	0	0	2	0	0	1
#38	4	0	2	5	0	1	2	2	1	1	0	0	2	0	0	0
#39	2	0	1	3	0	1	1	1	0	1	0	0	2	0	0	1
#40	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0
#42	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0

		PAPERS															
PAPERS		#22	#24	#26	#27	#28	#29	#30	#31	#33	#34	#36	#37	#38	#39	#40	#42
	#2	0	1	0	2	2	0	0	1	0	1	1	2	4	2	0	0
	#3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	#4	2	0	2	0	2	1	2	0	1	0	0	1	2	1	1	0
	#5	2	2	2	2	4	1	2	3	1	1	2	4	5	3	1	0
	#7	1	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0
	#8	0	1	0	1	1	0	0	5	0	0	1	2	1	1	0	0
	#10	0	3	0	1	1	0	0	0	0	1	1	2	2	1	0	0
	#11	0	1	0	2	1	0	0	1	0	1	1	3	2	1	0	0
	#12	1	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0
	#13	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	0
	#15	1	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0
	#16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	#18	0	1	0	1	4	0	1	0	0	1	1	2	2	2	0	1
	#19	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
	#20	3	1	2	0	1	2	2	0	1	0	0	0	0	0	1	0
	#21	1	1	1	0	3	1	1	1	0	1	1	1	0	1	0	0
	#22	4	0	2	0	2	2	3	0	1	0	0	0	0	0	1	0
	#24	0	3	0	1	0	0	0	0	0	0	0	0	1	0	0	0
	#26	2	0	3	0	1	1	2	0	1	0	0	0	0	0	1	0
	#27	0	1	0	3	0	0	0	1	0	0	0	0	2	0	0	0
	#28	2	0	1	0	5.5	1	1	1	0	1	2	2	1	2	0	1
	#29	2	0	1	0	1	2.5	1	0	0	0	0	0	0	0	0	0
	#30	3	0	2	0	1	1	3.5	0	1	0	0	0	0	0	1	0
	#31	0	0	0	1	1	0	0	4.5	0	0	1	1	1	1	0	0
	#33	1	0	1	0	0	0	1	0	1.5	0	0	0	0	0	1	0
	#34	0	0	0	0	1	0	0	0	0	1.5	1	1	0	1	0	0
	#36	0	0	0	0	2	0	0	1	0	1	3	2	0	2	0	0
	#37	0	0	0	0	2	0	0	1	0	1	2	5	1	3	0	0
	#38	0	1	0	2	1	0	0	1	0	0	0	1	5.5	1	0	1
	#39	0	0	0	0	2	0	0	1	0	1	2	3	1	4	0	0
	#40	1	0	1	0	0	0	1	0	1	0	0	0	0	0	1.5	0
	#42	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1.5

APPENDIX B

	PAPERS																
		#2	**#3**	**#4**	**#5**	**#7**	**#8**	**#10**	**#11**	**#12**	**#13**	**#15**	**#16**	**#18**	**#19**	**#20**	**#21**
PAPERS	**#2**	1.00	-0.21	0.76	0.92	-0.17	0.42	0.70	0.84	0.04	0.45	-0.16	-0.08	0.76	-0.28	-0.05	0.41
	#3	-0.21	1.00	-0.15	-0.21	0.72	0.07	-0.17	-0.21	0.29	-0.04	0.74	0.23	-0.06	0.22	-0.11	-0.04
	#4	0.76	-0.15	1.00	0.78	-0.02	0.29	0.41	0.64	0.15	0.78	-0.01	0.15	0.49	-0.18	0.40	0.58
	#5	0.92	-0.21	0.78	1.00	-0.17	0.49	0.78	0.80	0.00	0.51	-0.16	-0.08	0.74	-0.19	0.04	0.41
	#7	-0.17	0.72	-0.02	-0.17	1.00	-0.11	-0.15	-0.21	0.58	0.08	0.89	0.53	0.11	0.42	0.09	0.21
	#8	0.42	0.07	0.29	0.49	-0.11	1.00	0.43	0.51	0.11	0.10	-0.12	-0.25	0.21	-0.24	-0.18	0.05
	#10	0.70	-0.17	0.41	0.78	-0.15	0.43	1.00	0.59	-0.12	0.17	-0.15	-0.16	0.73	-0.03	-0.11	0.21
	#11	0.84	-0.21	0.64	0.80	-0.21	0.51	0.59	1.00	0.07	0.31	-0.20	-0.12	0.56	-0.31	-0.11	0.40
	#12	0.04	0.29	0.15	0.00	0.58	0.11	-0.12	0.07	1.00	0.01	0.53	0.19	-0.06	0.60	0.40	0.24
	#13	0.45	-0.04	0.78	0.51	0.08	0.10	0.17	0.31	0.01	1.00	0.12	0.40	0.37	-0.22	0.13	0.46
	#15	-0.16	0.74	-0.01	-0.16	0.89	-0.12	-0.15	-0.20	0.53	0.12	1.00	0.49	0.11	0.43	0.10	0.24
	#16	-0.08	0.23	0.15	-0.08	0.53	-0.25	-0.16	-0.12	0.19	0.40	0.49	1.00	0.14	0.02	0.02	0.52
	#18	0.76	-0.06	0.49	0.74	0.11	0.21	0.73	0.56	-0.06	0.37	0.11	0.14	1.00	-0.16	-0.17	0.38
	#19	-0.28	0.22	-0.18	-0.19	0.42	-0.24	-0.03	-0.31	0.60	-0.22	0.43	0.02	-0.16	1.00	0.41	-0.03
	#20	-0.05	-0.11	0.40	0.04	0.09	-0.18	-0.11	-0.11	0.40	0.13	0.10	0.02	-0.17	0.41	1.00	0.40
	#21	0.41	-0.04	0.58	0.41	0.21	0.05	0.21	0.40	0.24	0.46	0.24	0.52	0.38	-0.03	0.40	1.00
	#22	-0.08	-0.09	0.32	0.01	0.09	-0.18	-0.11	-0.17	0.24	0.05	0.08	-0.08	-0.11	0.39	0.90	0.30
	#24	0.41	-0.08	0.15	0.51	0.00	0.27	0.73	0.41	0.04	-0.03	-0.05	-0.03	0.44	0.04	-0.05	0.16
	#26	0.07	-0.23	0.47	0.18	-0.13	-0.08	0.01	-0.02	-0.03	0.16	-0.14	-0.10	-0.05	0.02	0.81	0.31
	#27	0.61	-0.18	0.21	0.51	-0.20	0.45	0.49	0.65	0.19	-0.03	-0.20	-0.25	0.39	-0.20	-0.29	-0.07
	#28	0.51	-0.31	0.44	0.53	-0.23	0.21	0.40	0.39	-0.22	0.34	-0.23	-0.01	0.66	-0.29	0.10	0.55
	#29	-0.12	0.02	0.22	-0.07	0.24	-0.20	-0.14	-0.19	0.47	0.01	0.23	0.04	-0.14	0.63	0.85	0.36
	#30	0.00	0.06	0.38	0.10	0.22	-0.15	-0.04	-0.11	0.12	0.14	0.27	0.06	0.07	0.19	0.77	0.33
	#31	0.29	-0.04	0.13	0.33	-0.16	0.92	0.22	0.37	0.15	-0.02	-0.20	-0.20	0.08	-0.27	-0.22	0.03
	#33	0.01	-0.05	0.38	0.10	0.17	-0.12	-0.04	-0.09	0.10	0.12	0.06	-0.01	-0.06	0.07	0.64	0.08
	#34	0.49	-0.25	0.15	0.44	-0.22	0.11	0.51	0.53	-0.28	0.05	-0.22	-0.07	0.57	-0.27	-0.31	0.40
	#36	0.39	-0.21	0.11	0.38	-0.23	0.43	0.42	0.43	-0.27	0.05	-0.25	-0.18	0.44	-0.33	-0.32	0.29
	#37	0.59	-0.22	0.35	0.62	-0.24	0.52	0.57	0.70	-0.22	0.28	-0.26	-0.17	0.57	-0.35	-0.32	0.26
	#38	0.83	-0.23	0.54	0.76	-0.22	0.42	0.63	0.70	0.06	0.36	-0.21	-0.19	0.63	-0.28	-0.21	0.10
	#39	0.57	-0.24	0.34	0.55	-0.24	0.39	0.47	0.52	-0.26	0.35	-0.25	-0.12	0.59	-0.39	-0.31	0.28
	#40	0.05	-0.20	0.40	0.14	-0.13	-0.08	0.00	-0.04	-0.07	0.11	-0.13	-0.18	-0.10	-0.04	0.66	0.03
	#42	0.04	-0.13	-0.10	-0.01	-0.11	-0.13	0.02	-0.06	-0.20	-0.01	-0.10	-0.06	0.33	-0.21	-0.23	-0.12

	PAPERS															
PAPERS	**#22**	**#24**	**#26**	**#27**	**#28**	**#29**	**#30**	**#31**	**#33**	**#34**	**#36**	**#37**	**#38**	**#39**	**#40**	**#42**
#2	-0.08	0.41	0.07	0.61	0.51	-0.12	0.00	0.29	0.01	0.49	0.39	0.59	0.83	0.57	0.05	0.04
#3	-0.09	-0.08	-0.23	-0.18	-0.31	0.02	0.06	-0.04	-0.05	-0.25	-0.21	-0.22	-0.23	-0.24	-0.20	-0.13
#4	0.32	0.15	0.47	0.21	0.44	0.22	0.38	0.13	0.38	0.15	0.11	0.35	0.54	0.34	0.40	-0.10
#5	0.01	0.51	0.18	0.51	0.53	-0.07	0.10	0.33	0.10	0.44	0.38	0.62	0.76	0.55	0.14	-0.01
#7	0.09	0.00	-0.13	-0.20	-0.23	0.24	0.22	-0.16	0.17	-0.22	-0.23	-0.24	-0.22	-0.24	-0.13	-0.11
#8	-0.18	0.27	-0.08	0.45	0.21	-0.20	-0.15	0.92	-0.12	0.11	0.43	0.52	0.42	0.39	-0.08	-0.13
#10	-0.11	0.73	0.01	0.49	0.40	-0.14	-0.04	0.22	-0.04	0.51	0.42	0.57	0.63	0.47	0.00	0.02
#11	-0.17	0.41	-0.02	0.65	0.39	-0.19	-0.11	0.37	-0.09	0.53	0.43	0.70	0.70	0.52	-0.04	-0.06
#12	0.24	0.04	-0.03	0.19	-0.22	0.47	0.12	0.15	0.10	-0.28	-0.27	-0.22	0.06	-0.26	-0.07	-0.20
#13	0.05	-0.03	0.16	-0.03	0.34	0.01	0.14	-0.02	0.12	0.05	0.05	0.28	0.36	0.35	0.11	-0.01
#15	0.08	-0.05	-0.14	-0.20	-0.23	0.23	0.27	-0.20	0.06	-0.22	-0.25	-0.26	-0.21	-0.25	-0.13	-0.10
#16	-0.08	-0.03	-0.10	-0.25	-0.01	0.04	0.06	-0.20	-0.01	-0.07	-0.18	-0.17	-0.19	-0.12	-0.18	-0.06
#18	-0.11	0.44	-0.05	0.39	0.66	-0.14	0.07	0.08	-0.06	0.57	0.44	0.57	0.63	0.59	-0.10	0.33
#19	0.39	0.04	0.02	-0.20	-0.29	0.63	0.19	-0.27	0.07	-0.27	-0.33	-0.35	-0.28	-0.39	-0.04	-0.21
#20	0.90	-0.05	0.81	-0.29	0.10	0.85	0.77	-0.22	0.64	-0.31	-0.32	-0.32	-0.21	-0.31	0.66	-0.23
#21	0.30	0.16	0.31	-0.07	0.55	0.36	0.33	0.03	0.08	0.40	0.29	0.26	0.10	0.28	0.03	-0.12
#22	1.00	-0.18	0.85	-0.32	0.21	0.86	0.89	-0.21	0.68	-0.27	-0.24	-0.28	-0.23	-0.26	0.69	-0.14
#24	-0.18	1.00	-0.11	0.54	0.07	-0.15	-0.13	0.12	-0.08	0.22	0.09	0.21	0.47	0.09	-0.12	-0.04
#26	0.85	-0.11	1.00	-0.24	0.27	0.63	0.84	-0.14	0.76	-0.18	-0.16	-0.15	-0.10	-0.14	0.83	-0.15
#27	-0.32	0.54	-0.24	1.00	0.04	-0.31	-0.29	0.42	-0.24	0.14	0.09	0.27	0.77	0.17	-0.20	0.06
#28	0.21	0.07	0.27	0.04	1.00	0.14	0.22	0.22	-0.05	0.63	0.66	0.60	0.34	0.68	0.01	0.39
#29	0.86	-0.15	0.63	-0.31	0.14	1.00	0.67	-0.20	0.38	-0.26	-0.26	-0.31	-0.26	-0.29	0.34	-0.15
#30	0.89	-0.13	0.84	-0.29	0.22	0.67	1.00	-0.22	0.73	-0.22	-0.23	-0.22	-0.17	-0.20	0.73	-0.11
#31	-0.21	0.12	-0.14	0.42	0.22	-0.20	-0.22	1.00	-0.19	0.10	0.44	0.43	0.33	0.37	-0.16	-0.08
#33	0.68	-0.08	0.76	-0.24	-0.05	0.38	0.73	-0.19	1.00	-0.31	-0.29	-0.24	-0.13	-0.24	0.90	-0.23
#34	-0.27	0.22	-0.18	0.14	0.63	-0.26	-0.22	0.10	-0.31	1.00	0.80	0.74	0.26	0.75	-0.27	0.06
#36	-0.24	0.09	-0.16	0.09	0.66	-0.26	-0.23	0.44	-0.29	0.80	1.00	0.82	0.22	0.85	-0.25	0.03
#37	-0.28	0.21	-0.15	0.27	0.60	-0.31	-0.22	0.43	-0.24	0.74	0.82	1.00	0.48	0.90	-0.19	0.02
#38	-0.23	0.47	-0.10	0.77	0.34	-0.26	-0.17	0.33	-0.13	0.26	0.22	0.48	1.00	0.46	-0.09	0.30
#39	-0.26	0.09	-0.14	0.17	0.68	-0.29	-0.20	0.37	-0.24	0.75	0.85	0.90	0.46	1.00	-0.20	0.09
#40	0.69	-0.12	0.83	-0.20	0.01	0.34	0.73	-0.16	0.90	-0.27	-0.25	-0.19	-0.09	-0.20	1.00	-0.21
#42	-0.14	-0.04	-0.15	0.06	0.39	-0.15	-0.11	-0.08	-0.23	0.06	0.03	0.02	0.30	0.09	-0.21	1.00

APPENDIX C

Sensitivity Analysis: Missing values

In this Appendix results from raw co-citation matrix with missing diagonal values are reported.

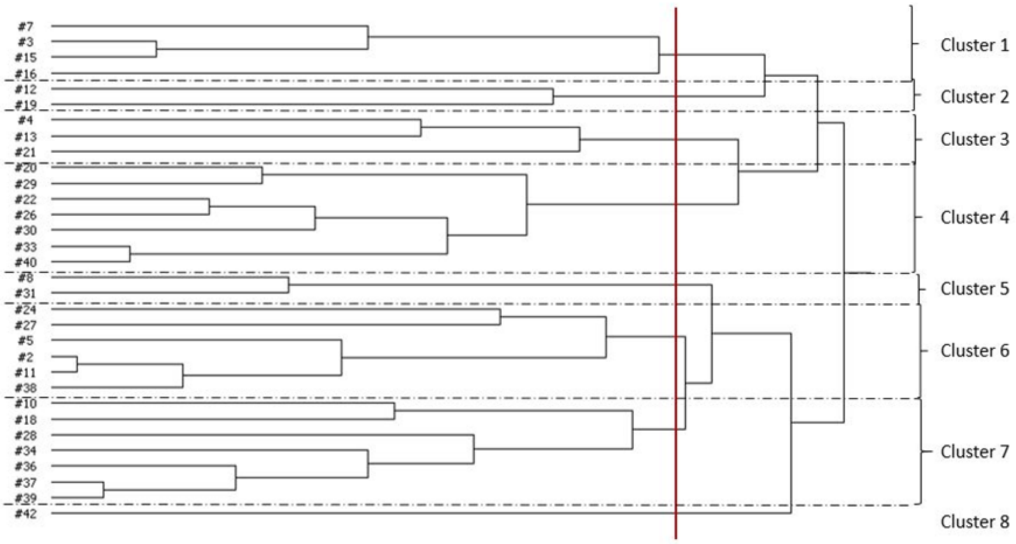
Principal Components Analysis

1	2	3	4	5	6
Paper #4	*Paper #2*	**Paper #21**	*Paper #3*	*Paper #8*	**Paper #19**
Paper #20	***Paper #4***	*Paper #28*	*Paper #7*	*Paper #31*	
Paper #21	*Paper #5*	*Paper #34*	Paper #12		
Paper #22	*Paper #10*	*Paper #36*	*Paper #15*		
Paper #26	*Paper #11*	*Paper #37*	*Paper #16*		
Paper #29	Paper #13	*Paper #39*	**Paper #19**		
Paper #30	*Paper #18*	Paper #42			
Paper #33	*Paper #24*				
Paper #40	*Paper #27*				
	Paper #38				

Papers with loadings higher than 0.7 are reported in italics. Paper #21 contributes equally both to Principal components #2 and #5, Paper #19 contributes almost equally to principal components #4 and #6, Paper #4 contributes to principal components #1 and more significantly to component #2: these three papers are reported in bold.

Cluster analysis

Dendrogram and icicle plot are reported as results of cluster analysis. The thicker red line in the dendrogram corresponds to the point in which the number of groups can be eyeballed.

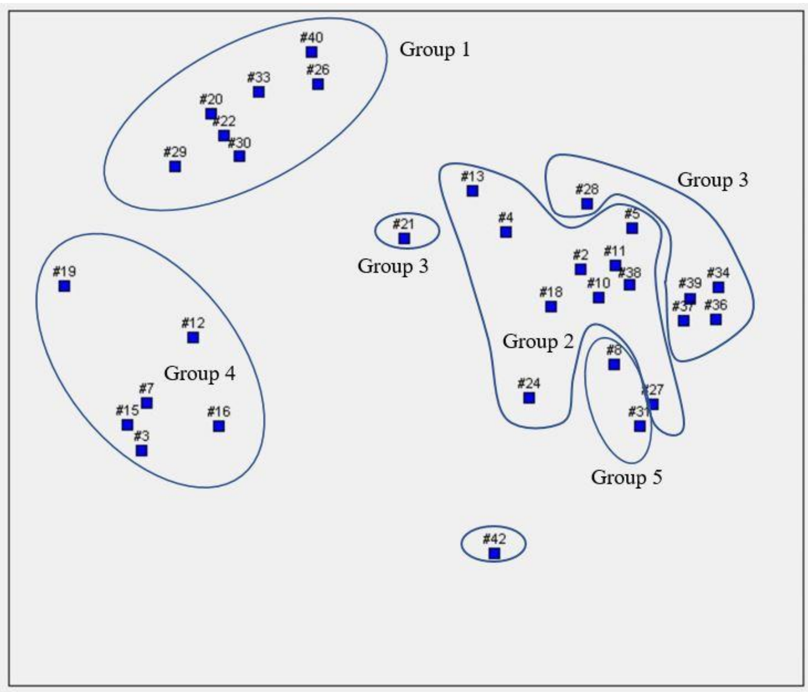


```
            # # # #   # # # # # # # # #   # # #     # # # # # # # # # #
          # # 1 1 1 1 # 1 2 2 2 2 2 3 3 4 # 3 2 2 # # 1 3 1 1 2 3 3 3 3 4
          7 3 5 6 2 9 4 3 1 0 9 2 6 0 3 0 8 1 4 7 5 2 1 8 0 8 8 4 6 7 9 2
  Level
-------   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 0.9340   . . . . . . . . . . . . . . . . . . . . . XXX . . . . . . . . .
 0.9160   . . . . . . . . . . . . . . . . . . . . . XXX . . . . . . XXX .
 0.9060   . . . . . . . . . . . . . . XXX . . . . . XXX . . . . . . XXX .
 0.9020   . XXX . . . . . . . . . . . XXX . . . . . XXX . . . . . . XXX .
 0.8970   . XXX . . . . . . . . . . . XXX . . . . . XXXXX . . . . . XXX .
 0.8900   . XXX . . . . . . . . XXX . XXX . . . . . XXXXX . . . . . XXX .
 0.8840   . XXX . . . . . . . . XXX . XXX . . . . . XXXXX . . . . XXXXX .
 0.8700   . XXX . . . . . . XXX XXX . XXX . . . . . XXXXX . . . . XXXXX .
 0.8430   . XXX . . . . . . XXX XXX . XXX XXX . . . XXXXX . . . . XXXXX .
 0.8400   . XXX . . . . . . XXX XXXXX XXX XXX . . . XXXXX . . . . XXXXX .
 0.8250   . XXX . . . . . . XXX XXXXX XXX XXX . . XXXXXXX . . . . XXXXX .
 0.8150   XXXXX . . . . . . XXX XXXXX XXX XXX . . XXXXXXX . . . XXXXXXX .
 0.7960   XXXXX . . . . . . XXX XXXXX XXX XXX . . XXXXXXX XXX . XXXXXXX .
 0.7520   XXXXX . . . XXX . XXX XXXXX XXX XXX . . XXXXXXX XXX . XXXXXXX .
 0.7250   XXXXX . . . XXX . XXX XXXXXXXXX XXX . . XXXXXXX XXX . XXXXXXX .
 0.6570   XXXXX . . . XXX . XXX XXXXXXXXX XXX . . XXXXXXX XXX XXXXXXXXX .
 0.6100   XXXXX . . . XXX . XXX XXXXXXXXX XXX XXX XXXXXXX XXX XXXXXXXXX .
 0.6030   XXXXX . . . XXX . XXXXXXXXXXXXX XXX XXX XXXXXXX XXX XXXXXXXXX .
 0.5940   XXXXX . XXX XXX . XXXXXXXXXXXXX XXX XXX XXXXXXX XXX XXXXXXXXX .
 0.5880   XXXXX . XXX XXXXX XXXXXXXXXXXXX XXX XXX XXXXXXX XXX XXXXXXXXX .
 0.5300   XXXXX . XXX XXXXX XXXXXXXXXXXXX XXX XXXXXXXXXXX XXX XXXXXXXXX .
 0.5260   XXXXX . XXX XXXXX XXXXXXXXXXXXX XXX XXXXXXXXXXX XXXXXXXXXXXXX .
 0.4330   XXXXXXX XXX XXXXX XXXXXXXXXXXXX XXX XXXXXXXXXXX XXXXXXXXXXXXX .
 0.2820   XXXXXXX XXX XXXXX XXXXXXXXXXXXX XXX XXXXXXXXXXXXXXXXXXXXXXXXX .
 0.2180   XXXXXXX XXX XXXXX XXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXX .
 0.1850   XXXXXXX XXX XXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXX .
 0.1560   XXXXXXXXXXX XXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXX .
-0.0570   XXXXXXXXXXX XXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
-0.2000   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
-0.3740   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
```

The maximum decrease in the level of similarity is recorded between the fourth to last and the third to last level. The second largest decrease is recorded between the eighth to last and the seventh to last level. Therefore, these two solutions may provide an indication for detecting clusters. The latter option provides the same result as the one identified by the dendrogram.

MDS

(stress value: 0.135, acceptable)



The identification of groups in MDS is made in compliance with the results provided by PCA. Papers are plotted in a way that resembles MDS with artificial value.

APPENDIX D

UNIVERSITÀ DEGLI STUDI DI PADOVA Students Entrepreneurship Analysis JESSICA RAGAZZON

Salve,

sono Jessica Ragazzon e sono una studentessa della magistrale in business administration dell'Università di Padova. La contatto perchè sto portando avanti un progetto di tesi in cui affronto il tema della students entrepreneurship, quindi (ex)studenti dell'Università di Padova che sono diventati imprenditori. Nella fattispecie andrò ad analizzare se ci sono degli elementi sociali, personali e/o legati all'Università che stimolano ed inducono il soggetto alla creazione di un'attività imprenditoriale.

Il questionario è diviso in più parti distinte (poche domande per sezione): in una andiamo ad indagare le motivazioni personali, in un'altra i fattori ambientali del perchè localizzarsi in Veneto e non in altre regioni d'Italia, in un'altra sezione si domanda se ci sono state delle influenze dell'Università e in un'altra ancora si valutano le performances aziendali per studiare e valutare le evoluzioni negli anni degli ex studenti imprenditori (anche in relazione al loro background universitario).

Ci vorranno solo pochissimi minuti ed il suo contributo è veramente prezioso. Per onor di cronaca: non è stato scelto in maniera casuale tra millemila persone, ma selezionato appositamente (tramite un database fornito da infocamere ed uno da almalaurea) perchè rientra alla perfezione (e siete in pochi) nei criteri della mia ricerca.

Grazie in anticipo per la sua disponibilità, non le rubo altro tempo!

UNIVERSITÀ DEGLI STUDI DI PADOVA Students Entrepreneurship Analysis JESSICA RAGAZZON

* 1. Nome

* 2. Cognome

1

3. Codice fiscale

4. Qual è il nome dell'azienda?

* 5. Fa parte del team che ha fondato la società?

- Si
- No

UNIVERSITÀ DEGLI STUDI DI PADOVA Students Entrepreneurship Analysis JESSICA RAGAZZON

* 6. Può brevemente descrivere qual è l'attività prevalente dell'azienda che ha fondato?

* 7. Ha accumulato delle esperienze professionali prima della creazione dell'impresa?

- Si
- No

UNIVERSITÀ DEGLI STUDI DI PADOVA Students Entrepreneurship Analysis JESSICA RAGAZZON

8. Che tipo di esperienza professionale ha accumulato? Sono possibili più risposte.

- ☐ Lavoro nell'impresa di famiglia operante nello stesso settore della sua azienda
- ☐ Lavoro nell'impresa di famiglia operante in un settore diverso rispetto alla sua azienda
- ☐ Lavoro in un'altra impresa operante nello stesso settore della sua azienda
- ☐ Lavoro in un'altra impresa operante in un settore diverso rispetto alla sua azienda
- ☐ Attività di consulenza svolta in proprio
- ☐ Attività di consulenza svolta per un'altra società
- ☐ Fondazione di un'altra società operante nello stesso settore della sua azienda
- ☐ Fondazione di un'altra società operante in un settore diverso rispetto alla sua azienda

9. Indichi quanto sono stati importanti i seguenti elementi maturati nel corso delle pregresse esperienze

	Per nulla importante	Poco importante	Abbastanza importante	Molto importante	Estremamente importante	Non applicabile
Conoscenza del mercato di riferimento (players, dimensione, barriere...)	○	○	○	○	○	○
Acquisizione di competenze tecniche	○	○	○	○	○	○
Relazioni personali con i colleghi	○	○	○	○	○	○
Relazioni con clienti e fornitori	○	○	○	○	○	○
Altro (specificare sotto e riportare la sua importanza)	○	○	○	○	○	○

.

* 10. La sua azienda, al momento della costituzione, è stata:

- ◯ Il risultato di una operazione straordinaria (scissione, LBO, conferimento) di una impresa già esistente, DE ALIO
- ◯ Creata ex-novo, DE NOVO (start-up, spin-off)
- ◯ Il risultato di un investimento di un'impresa e di una persona fisica che svolge un ruolo gestionale nell'azienda (franchising)

11. Alla fondazione, la sua impresa: (può indicare più di una risposta)

- ☐ Si è inserita in un mercato già esistente producendo beni/servizi leggermente migliorati
- ☐ Si è inserita in un mercato già esistente producendo beni/servizi radicalmente modificati
- ☐ Si è inserita in un mercato già esistente producendo beni/servizi con costi inferiori
- ☐ Ha lanciato uno o più nuovi beni/servizi che creano un nuovo mercato
- ☐ Ha lanciato uno o più nuovi beni/servizi in un mercato emergente, ma non ancora consolidato
- ☐ Altro (specificare)

12. Quante persone lavoravano (soci + dipendenti) nella sua impresa

Alla fine del primo anno di attività	
Alla fine del secondo anno	
Alla fine del terzo anno	
Al 2015	

13. Indichi approssimativamente il valore dei ricavi della sua impresa

Alla fine del primo anno di attività	
Alla fine del secondo anno	
Ala fine del terzo anno	
Al 2015	

4

14. Indichi la percentuale del fatturato derivante da prodotti/servizi nuovi sul mercato

Alla fine del primo anno di attività	
Alla fine del secondo anno	
Alla fine del terzo anno	
Al 2015	

15. La sua impresa è localizzata in Veneto?

- Si
- No

16.
Esprima il grado di importanza della presenza dei seguenti fattori nella decisione di creare la sua impresa in Veneto e non in altre regioni

	Per nulla importante	Poco importante	Abbastanza importante	Molto importante	Estremamente importante
Infrastrutture economiche (Reti stradali, Aeroporti, Strutture per il trasferimento delle merci..)	○	○	○	○	○
Manodopera specializzata	○	○	○	○	○
Presenza di distretti industriali (fiducia tra operatori, riduzione costi di transazione)	○	○	○	○	○
Presenza di fornitori specializzati (di beni e/o servizi)	○	○	○	○	○
Presenza di un potenziale mercato per i propri beni/servizi	○	○	○	○	○
Facilità di accesso al credito	○	○	○	○	○
Agevolazioni fiscali	○	○	○	○	○

	Per nulla importante	Poco importante	Abbastanza importante	Molto importante	Estremamente importante
Vantaggi di immagine/reputazione legata al territorio (es. certificazioni IGP, DOCG...)					
Presenza di società/consulenti che offrono servizi alle imprese in tema di qualità/formazione/accesso al credito/bandi regionali e/o comunitari (KIBS)					
Presenza di incubatori di impresa/parchi scientifici tecnologici					
Presenza dell'università dove si è laureato (l'università ha influito sulla scelta di creare impresa in Veneto?)					
Motivi personali (vicinanza casa/famiglia/amici)					
Vi erano localizzate le imprese dove ha lavorato in precedenza o i loro clienti/fornitori					
Altro (specificare sotto e riportare la sua importanza)					

.

6

17. **SPECIFICARE CHE LA DOMANDA E' INDIPENDENTE DAL VENETO**

Indichi il grado di importanza dei seguenti attori chiave nella sua decisione di creare impresa.
Indichi "Non ne ho usufruito" se non si è servito di quella realtà

	Per niente importante	Poco importante	Abbastanza importante	Molto importante	Estremamente importante	Non ne ho usufruito
Incubatori di impresa	◯	◯	◯	◯	◯	◯
Camera di Commercio ed enti pubblici/privati a supporto dell'imprenditorialità	◯	◯	◯	◯	◯	◯
Associazioni di categoria (Confindustria, Associazioni Artigiani, ecc.)	◯	◯	◯	◯	◯	◯
Parchi scientifici tecnologici	◯	◯	◯	◯	◯	◯
Network relazionale personale	◯	◯	◯	◯	◯	◯
Altro (specificare sotto e riportare la sua importanza)	◯	◯	◯	◯	◯	◯

.

UNIVERSITÀ DEGLI STUDI DI PADOVA Students Entrepreneurship Analysis JESSICA RAGAZZON

* 18. Ha costituito la sua impresa da solo o con dei soci?

◯ Solo

◯ Con dei soci

UNIVERSITÀ DEGLI STUDI DI PADOVA Students Entrepreneurship Analysis JESSICA RAGAZZON

19. Con quanti soci?

0 10

20. A parte lei, quanti componenti del team che hanno fondato l'impresa hanno avuto esperienze pregresse di **CREAZIONE** di imprese?

0 10

21. Che relazione esiste tra i fondatori? (Può fornire più di una risposta)

- [ ] Parentela
- [ ] Amicizia
- [ ] Precedenti esperienze lavorative
- [ ] Si sono conosciuti ad eventi per stimolare/promuovere l'imprenditorialità organizzati dall'Università di Padova
- [ ] Hanno condiviso lo stesso percorso di studio
- [ ] **SE HANNO CONDIVISO LO STESSO PERCORSO DI STUDIO RISPONDERE ALLE DUE SUCCESSIVE ALTERNATIVE**
- [ ] Si sono conosciuti tra i banchi (coevi. Compagni di corso)
- [ ] Hanno frequentato lo stesso corso di studio (anni diversi ma stesso percorso)
- [ ] Altro (specificare)

8

22. Esprima una valutazione dell'importanza dei seguenti fattori nella scelta di coinvolgere dei soci

	Per niente rilevante	Poco rilevante	Abbastanza rilevante	Molto rilevante	Estremamente rilevante	Non applicabile
Colmare la mancanza di specifiche conoscenze tecniche						
Colmare la mancanza di conoscenze economico gestionali						
Per avere il capitale minimo per l'avvio						
Per il network del/i socio/i						
(DA CHIEDERE SOLO SE NELLA PRECEDENTE DICE "PARENTELA") Sono miei familiari e la mia è una azienda di famiglia						
Altro (specificare sotto e riportare la sua importanza)						

* 23. L'esperienza come studente universitario ha avuto un ruolo nella formazione del team (per trovare soci, dipendenti, stagisti, collaboratori etc.)?

- Si
- No

UNIVERSITÀ DEGLI STUDI DI PADOVA Students Entrepreneurship Analysis JESSICA RAGAZZON

24. Che ruolo ha avuto l'esperienza universitaria nella sua impresa?

Per niente rilevante	Poco rilevante	Abbastanza rilevante	Molto rilevante	Estremamente rilevante

25. In che termini?

26. Qual è stato l'impatto delle seguenti situazioni nella formazione del team?

	Per niente importante	Poco importante	Abbastanza importante	Molto importante	Estremamente importante	Non applicabile
Partecipazione a lavori di gruppo in aula	○	○	○	○	○	○
Partecipazione alle stesse lezioni	○	○	○	○	○	○
Partecipazione ad incontri/aperitivi informali (Academy night, ecc.) promossi dall'Università	○	○	○	○	○	○
Partecipazione a visite aziendali	○	○	○	○	○	○
Partecipazione a seminari o eventi sull'imprenditorialità o di orientamento al mondo del lavoro	○	○	○	○	○	○
Frequentazione delle stesse aule studio universitarie	○	○	○	○	○	○
Altro (specificare sotto e riportare la sua importanza)	○	○	○	○	○	○

.

27. Quanti dei suoi soci sono laureati?

0 10

DA CHIEDERE SE LA PRECEDENTE E' DIVERSA DA 0

28. Quanti di questi sono laureati a Padova?

0 10

29. Cosa hanno studiato i soci? (Facoltà)

30. Le conoscenze apprese durante l'università quanto sono state importanti per la sua impresa?

Per niente rilevante	Poco rilevante	Abbastanza rilevante	Molto rilevante	Estremamente rilevante
○	○	○	○	○

31. Che tipo di finanziamenti ha utilizzato per avviare l'azienda? Può indicare più di una risposta

- ☐ Prestito bancario
- ☐ Venture/ seed capital/ business angels privati
- ☐ Venture/seed capital/ prestiti partecipativi o altri fondi pubblici
- ☐ Capitali propri, familiari o di amici
- ☐ Altre imprese
- ☐ Altro (specificare)

32. Esprima una valutazione dell'importanza delle seguenti motivazioni per la fondazione della sua impresa:

	Per niente importante	Poco importante	Abbastanza importante	Molto importante	Estremamente importante
Volevo sfruttare opportunità di business intraviste nelle esperienze lavorative passate	◯	◯	◯	◯	◯
Volevo sfruttare opportunità nate durante l'Università	◯	◯	◯	◯	◯
Per sfruttare opportunità che prescindono da esperienze passate	◯	◯	◯	◯	◯
Per sfruttare opportunità scaturite da bisogni personali non soddisfatti dalle soluzioni presenti sul mercato	◯	◯	◯	◯	◯
Altro (specificare)	◯	◯	◯	◯	◯

.

33. Esprima un parere sul grado di importanza dei seguenti elementi legati all'università, che possono avere influenzato la scelta di avviare la sua impresa

	Per niente importante	Poco importante	Abbastanza importante	Molto importante	Estremamente importante	Non applicabile
Professori stimolanti	○	○	○	○	○	○
Materie del suo percorso di studio	○	○	○	○	○	○
Incontri con manager in aula	○	○	○	○	○	○
Case studies di aziende, simulation games	○	○	○	○	○	○
Visite aziendali	○	○	○	○	○	○
Seminari/cicli di incontri sull'imprenditorialità	○	○	○	○	○	○
Associazioni studentesche	○	○	○	○	○	○
Eventi organizzati dall'università (Galileo festival, aperitivi con managers/imprese...)	○	○	○	○	○	○
Altro (specificare sotto e riportare la sua importanza)	○	○	○	○	○	○

34. Indichi quanto sono stati importanti per il buon funzionamento dell'impresa le sue relazioni interpersonali indicate di seguito:

	Per nulla importante	Poco importante	Abbastanza importante	Molto importante	Estremamente importante	Non applicabile
Maturate durante l'università	○	○	○	○	○	○
Maturate in esperienze professionali precedenti alla creazione dell'impresa	○	○	○	○	○	○
Maturate in esperienze professionali durante l'attività nell'impresa che ha fondato	○	○	○	○	○	○
Maturate durante la partecipazione ad un programma di incubazione	○	○	○	○	○	○
Maturate sul web	○	○	○	○	○	○
Family and friends	○	○	○	○	○	○

35. Guardando al periodo dopo la laurea, esprima una valutazione sulla rilevanza delle relazioni con l'Università di Padova

	Per niente imporante	Poco imporante	Abbastanza imporante	Molto imporante	Estremamente imporante
Relazioni con professori	○	○	○	○	○
Rapporti con i relatori di tesi	○	○	○	○	○
Ricerca di stagisti	○	○	○	○	○
Progetto di ricerca	○	○	○	○	○
Partecipazione a convegni/seminari in aula	○	○	○	○	○
Altro (specificare sotto e riportare la sua importanza)	○	○	○	○	○

.

14

Il questionario è finito!

Grazie infinite per il suo prezioso aiuto.

36. Se vuole ricevere i risultati della nostra ricerca lasci una sua mail!

15

APPENDIX E

First χ^2 test

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6.149[a]	2	0.046
N of Valid Cases	195		

a. 1 cells (16.7%) have expected count less than 5. The minimum expected count is 4.00.

Second χ^2 test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.789[a]	1	0.052
N of Valid Cases	195		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 17.33.

Third χ^2 test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5.705[a]	1	0.017
N of Valid Cases	180		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.61.

APPENDIX F

Logit Model 1

Correlation matrix

Pearson Correlations

		spinoff	size	location	incu.scipark	org.assoc	network	partner	funds	primary	secondary
spinoff	corr[20]	1	-.019	.066	.052	-.066	.016	.008	-.037	.070	-.048
	sig.		.791	.362	.470	.358	.827	.914	.603	.330	.507
size	corr	-.019	1	.079	.153*	.012	.048	.407**	.067	-.090	-.014
	sig.	.791		.274	.033	.864	.509	.000	.355	.209	.843
location	corr	.066	.079	1	.054	.021	.104	.132	-.085	-.165*	.013
	sig.	.362	.274		.454	.769	.148	.066	.239	.021	.860
incu.scipark	corr	.052	.153*	.054	1	.203**	.150*	.005	.034	.028	-.132
	sig.	.470	.033	.454		.004	.037	.943	.640	.701	.067
org.assoc	corr	-.066	.012	.021	.203**	1	.314**	-.151*	.000	.080	-.182*
	sig.	.358	.864	.769	.004		.000	.035	.995	.267	.011
network	corr	.016	.048	.104	.150*	.314**	1	.021	-.086	.013	-.064
	sig.	.827	.509	.148	.037	.000		.769	.230	.854	.375
partner	corr	.008	.407**	.132	.005	-.151*	.021	1	.027	-.127	-.009
	sig.	.914	.000	.066	.943	.035	.769		.712	.076	.897
funds	corr	-.037	.067	-.085	.034	.000	-.086	.027	1	.107	.053
	sig.	.603	.355	.239	.640	.995	.230	.712		.135	.463
primary	corr	.070	-.090	-.165*	.028	.080	.013	-.127	.107	1	-.117
	sig.	.330	.209	.021	.701	.267	.854	.076	.135		.102
secondary	corr	-.048	-.014	.013	-.132	-.182*	-.064	-.009	.053	-.117	1
	sig.	.507	.843	.860	.067	.011	.375	.897	.463	.102	

*. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed).

Multicollinearity statistics

	Tolerance	VIF
spinoff	0.975	1.026
size	0.803	1.245
location	0.938	1.066
incu.scipark	0.912	1.096
org.assoc	0.819	1.222
network	0.874	1.144
partner	0.789	1.267
funds	0.962	1.040
primary	0.926	1.080
secondary	0.938	1.066

[20] “corr” means correlation, whereas “sig” significance

Logit Model 2A

Correlation matrix

Pearson Correlations

		no.exper	no.ind.spec	size	location	incu.scipark	org.assoc	network	partner	funds	primary	secondary
no.exper	cor[21]r	1.0	-.170*	0.0	0.1	-0.1	0.1	0.0	0.0	0.1	-0.1	0.0
	sig.		0.0	0.7	0.1	0.3	0.1	0.9	0.7	0.4	0.2	0.7
no.ind.spec	corr	-.170*	1.0	0.0	-.150*	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	sig.	0.0		0.9	0.0	0.9	0.9	0.9	0.9	0.9	0.8	0.6
size	corr	0.0	0.0	1.0	0.1	.153*	0.0	0.0	.407**	0.1	-0.1	0.0
	sig.	0.7	0.9		0.3	0.0	0.9	0.5	0.0	0.4	0.2	0.8
location	corr	0.1	-.150*	0.1	1.0	0.1	0.0	0.1	0.1	-0.1	-.165*	0.0
	sig.	0.1	0.0	0.3		0.5	0.8	0.1	0.1	0.2	0.0	0.9
incu.scipark	corr	-0.1	0.0	.153*	0.1	1.0	.203**	.150*	0.0	0.0	0.0	-0.1
	sig.	0.3	0.9	0.0	0.5		0.0	0.0	0.9	0.6	0.7	0.1
org.assoc	corr	0.1	0.0	0.0	0.0	.203**	1.0	.314**	-.151*	0.0	0.1	-.182*
	sig.	0.1	0.9	0.9	0.8	0.0		0.0	0.0	1.0	0.3	0.0
network	corr	0.0	0.0	0.0	0.1	.150*	.314**	1.0	0.0	-0.1	0.0	-0.1
	sig.	0.9	0.9	0.5	0.1	0.0	0.0		0.8	0.2	0.9	0.4
partner	corr	0.0	0.0	.407**	0.1	0.0	-.151*	0.0	1.0	0.0	-0.1	0.0
	sig.	0.7	0.9	0.0	0.1	0.9	0.0	0.8		0.7	0.1	0.9
funds	corr	0.1	0.0	0.1	-0.1	0.0	0.0	-0.1	0.0	1.0	0.1	0.1
	sig.	0.4	0.9	0.4	0.2	0.6	1.0	0.2	0.7		0.1	0.5
primary	corr	-0.1	0.0	-0.1	-.165*	0.0	0.1	0.0	-0.1	0.1	1.0	-0.1
	sig.	0.2	0.8	0.2	0.0	0.7	0.3	0.9	0.1	0.1		0.1
secondary	corr	0.0	0.0	0.0	0.0	-0.1	-.182*	-0.1	0.0	0.1	-0.1	1.0
	sig.	0.7	0.6	0.8	0.9	0.1	0.0	0.4	0.9	0.5	0.1	

*. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed).

Multicollinearity statistics

	Tolerance	VIF
no.exper	0.918	1.090
no.ind.spec	0.948	1.055
size	0.803	1.246
location	0.909	1.100
incu.scipark	0.906	1.104
org.assoc	0.810	1.235
network	0.873	1.146
partner	0.787	1.270
funds	0.958	1.044
primary	0.922	1.085
secondary	0.938	1.066

[21] "corr" means correlation, whereas "sig" significance

Logit Model 2B

Correlation matrix

Pearson Correlations

		spinoff	no.ind.spec	size	location	incu.scipark	org.assoc	network	partner	funds	primary	secondary
spinoff	corr[22]	1	-.830**	-.019	.066	.052	-.066	.016	.008	-.037	.070	-.048
	sig.		.000	.791	.362	.470	.358	.827	.914	.603	.330	.507
no.ind.spec	corr	-.830**	1	.006	-.150*	-.012	.005	-.009	.009	.007	-.015	.034
	sig.	.000		.933	.036	.871	.939	.899	.900	.924	.840	.640
size	corr	-.019	.006	1	.079	.153*	.012	.048	.407**	.067	-.090	-.014
	sig.	.791	.933		.274	.033	.864	.509	.000	.355	.209	.843
location	corr	.066	-.150*	.079	1	.054	.021	.104	.132	-.085	-.165*	.013
	sig.	.362	.036	.274		.454	.769	.148	.066	.239	.021	.860
incu.scipark	corr	.052	-.012	.153*	.054	1	.203**	.150*	.005	.034	.028	-.132
	sig.	.470	.871	.033	.454		.004	.037	.943	.640	.701	.067
org.assoc	corr	-.066	.005	.012	.021	.203**	1	.314**	-.151*	.000	.080	-.182*
	sig.	.358	.939	.864	.769	.004		.000	.035	.995	.267	.011
network	corr	.016	-.009	.048	.104	.150*	.314**	1	.021	-.086	.013	-.064
	sig.	.827	.899	.509	.148	.037	.000		.769	.230	.854	.375
partner	corr	.008	.009	.407**	.132	.005	-.151*	.021	1	.027	-.127	-.009
	sig.	.914	.900	.000	.066	.943	.035	.769		.712	.076	.897
funds	corr	-.037	.007	.067	-.085	.034	.000	-.086	.027	1	.107	.053
	sig.	.603	.924	.355	.239	.640	.995	.230	.712		.135	.463
primary	corr	.070	-.015	-.090	-.165*	.028	.080	.013	-.127	.107	1	-.117
	sig.	.330	.840	.209	.021	.701	.267	.854	.076	.135		.102
secondary	corr	-.048	.034	-.014	.013	-.132	-.182*	-.064	-.009	.053	-.117	1
	sig.	.507	.640	.843	.860	.067	.011	.375	.897	.463	.102	

**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

Multicollinearity statistics

	Tolerance	VIF
spinoff	0.293	3.410
no.ind.spec	0.293	3.414
size	0.803	1.246
location	0.909	1.100
incu.scipark	0.906	1.104
org.assoc	0.810	1.235
network	0.873	1.146
partner	0.787	1.270
funds	0.958	1.044
primary	0.922	1.085
secondary	0.938	1.066

[22] "corr" means correlation, whereas "sig" significance

Logit Model 3

Correlation matrix

Pearson Correlations

		spinoff	size	location	incu.scipark	org.assoc	network	partner	funds	primary	secondary
spinoff	corr[23]	1	-.011	.133	.025	-.024	.012	-.004	-.017	.032	-.041
	sig.		.888	.075	.736	.746	.873	.953	.819	.668	.582
size	corr	-.011	1	.073	.131	.010	.038	.417**	.089	-.091	-.057
	sig.	.888		.330	.079	.892	.608	.000	.233	.223	.448
location	corr	.133	.073	1	.060	.007	.063	.136	-.084	-.156*	.004
	sig.	.075	.330		.421	.923	.402	.068	.260	.037	.956
incu.scipark	corr	.025	.131	.060	1	.220**	.149*	.005	.038	.021	-.159*
	sig.	.736	.079	.421		.003	.047	.948	.616	.778	.033
org.assoc	corr	-.024	.010	.007	.220**	1	.333**	-.155*	-.007	.092	-.196**
	sig.	.746	.892	.923	.003		.000	.038	.928	.221	.008
network	corr	.012	.038	.063	.149*	.333**	1	.009	-.067	.013	-.081
	sig.	.873	.608	.402	.047	.000		.908	.371	.867	.281
partner	corr	-.004	.417**	.136	.005	-.155*	.009	1	.027	-.136	.012
	sig.	.953	.000	.068	.948	.038	.908		.716	.069	.874
funds	corr	-.017	.089	-.084	.038	-.007	-.067	.027	1	.119	.092
	sig.	.819	.233	.260	.616	.928	.371	.716		.113	.218
primary	corr	.032	-.091	-.156*	.021	.092	.013	-.136	.119	1	-.121
	sig.	.668	.223	.037	.778	.221	.867	.069	.113		.105
secondary	corr	-.041	-.057	.004	-.159*	-.196**	-.081	.012	.092	-.121	1
	sig.	.582	.448	.956	.033	.008	.281	.874	.218	.105	

**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

Multicollinearity statistics

	Tolerance	VIF
spinoff	0.976	1.025
size	0.796	1.256
location	0.932	1.073
incu.scipark	0.912	1.097
org.assoc	0.808	1.238
network	0.873	1.145
partner	0.782	1.279
funds	0.951	1.052
primary	0.925	1.081
secondary	0.920	1.087

[23] "corr" means correlation, whereas "sig" significance