



Networks, geography and the survival of the firm

Mark J. O. Bagley^{1,2,3} 

Published online: 17 May 2019

© The Author(s) 2019

Abstract

Prior studies show that the success of firms in industrial clusters is the result of two main reasons; the transfer of knowledge and routines from parent firms to spinoffs that locate in the same locality, and the returns from co-location of firms. While previous research has largely inferred the presence of parent-spinoff networks, few studies have measured them. Furthermore, the lack of geographic precision has led to conflicting results for evidence of returns from location, as the gains from geographic proximity may not always be linear. This paper introduces network measurement and a refined geographic measure to separate these two respective channels of knowledge transfer, and analyzes their impact on firm survival (as a proxy for firm success). It is found that the gains with respect to location are nonlinear. Furthermore, a firm's historical links formed through parent-spinoff linkages have a significant impact on survival, which differ depending on the motivations of the entrepreneur. Moreover, these channels of knowledge are complementary in nature.

Keywords Industrial clusters · Spinoffs · Networks · Knowledge flows · Firm survival · Schumpeter

JEL Classification B52 · D85 · L26 · L86 · O31 · R11

1 Introduction

Prior studies have shown that industrial clusters may arise from an entrepreneurial spawning process where employees of incumbent firms depart and establish spinoffs,

✉ Mark J. O. Bagley
mark.bagley@gu.se

¹ Institute of Innovation and Entrepreneurship, University of Gothenburg, Gothenburg, Sweden

² Centre for Entrepreneurship and Spatial Economics (CEnSE), Jönköping University, Jönköping, Sweden

³ Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE), Lund University, Lund, Sweden

which occupy the same or a related sector (Klepper 2007a; Boschma and Wenting 2007). Additionally, firms in industrial clusters are said to take advantage of localization economies, which includes the sharing of information regarding local labor markets and other types of industry-specific knowledge¹ (Marshall 1890). Evidence suggests that there are two main sources of knowledge that give spinoffs in clusters a performance premium compared to other types of new firm. The first channel, which draws from the literature on spinoffs and industrial dynamics, enable the passive transfer of knowledge and routines from parent firms to spinoffs (Renski 2011; Rosenthal and Strange 2005). Spinoffs locate close to their parents and inherit tried and tested routines that convey superior performance. The second channel, gleaned from the literature on industrial clusters, concerns the returns from locating alongside other establishments in their sector, taking advantage of knowledge spillovers and face-to-face contact with other firms and clients (Klepper 2001; Agarwal et al. 2004). Firms that happen to locate in clusters thus benefit from external scale economies. While the latter knowledge channel is a true ‘environmental effect’, the former is not. In this paper, we measure these knowledge channels separately and test them in a common analysis to gauge their effect on firm survival² in industrial clusters. We test the hypothesis that spinoff firms gain a performance premium via the occupational background of the entrepreneur. Those firms that survive then provide a source of potential knowledge for other firms in the same geographic area. The aim of this paper is thus to explore how the varying structure of inherited knowledge channels affect the prospects of survival among spinoff firms, and to disentangle these effects from those that arise from co-location.

Studies disentangling this mechanism of localization have been sparse. Frenken et al. (2015), in their survey of recent literature in industrial dynamics and clusters, found that there is an empirical gap in studies that analyze the effect of localization economies on firm survival, and the conditions, mechanisms and spatial scale that affect different types of firms are not well understood. Due to the lack of longitudinal data, knowledge channels between parent firms to spinoffs have mostly been inferred rather than identified and measured. But this does not fully capture the underlying mechanisms that may give certain firms an advantage in performance, namely the knowledge and routines inherited through the founder’s history. Firm success may not derive from geographic clustering in its own right but from the experience of the entrepreneur. Routines and knowledge gained from experience are genealogical in nature, as they reflect multiple generations of parent-progeny relationships. Furthermore, the lack of studies using continuous spatial scale have contributed to conflicting results concerning the survivability of spinoff firms in and outside industrial clusters. Some authors, such as Arzaghi and Henderson (2008), found that the benefits of co-location attenuate sharply with distance, with any gains from such externalities restricted to the neighborhood level rather than the city level. We argue

¹In addition to input-sharing.

²Of course, there are other indicators of firm success (growth, sales, etc.). However, it should be noted that many technology firms in their first few years tend not to exhibit growth or profit, but remain in the seed funding stage or series A round of venture capital financing, which does not tend to entail sales or revenue.

that geographic proximity has a role to play in localization after controlling for other types of knowledge channels. This role may not necessarily be linear. Due to supply- and demand-side competition effects, geographic proximity may act as a cost (Sorenson and Audia 2000; Stuart and Sorenson 2003) as well as a benefit.

To separate these two knowledge channels, this paper examines the Information and Communications Technology (ICT) cluster in Stockholm county, a research-intensive sector that exhibits a high degree of clustering in that region. Using Swedish employer-employee matched data for 1990 to 2010, we construct a network based on parent-spinoff relationships. We may then measure the level of connectivity of a firm relative to the rest of the network, and this serves to measure the inherited and passive knowledge channels of the firm, which may act as a conduit of established knowledge and routines. To account for knowledge channels that derive from geographic proximity, coordinate data allows for a continuous measure of distance from a firm's location to the cluster's core, which may shed some light on any potential nonlinearity of the returns from co-location. We use a Cox proportional hazard model to understand the effects these measures have on the survival rates of new firms. We expect that a firm's genealogical background is a major explanator of firm survival rates, and that the effect of geographic distance from a cluster's core on firm survivability is nonlinear in nature. These effects are expected to differ depending on the underlying motivations of the entrepreneur, and the reasons for establishing a new firm.

The rest of this paper is structured as follows. We begin by introducing the evolutionary framework of industrial clusters, which describes the inherited knowledge and routines of new firms, as well as an overview of past survival studies regarding spinoffs in industrial clusters. This leads us to provide an overview of the current research gaps in this field, and to introduce two types of information flow for the firm: realized and unrealized knowledge channels. The next section introduces the data, which includes spinoff identification, the rules for assembling the network, and an overview of the variables used in regression analysis, including the measurements used to proxy different knowledge channels. The fourth section gives the results of the Cox survival analysis, which we expect will unmask some of the underlying mechanics of localization economies. The final section concludes.

2 Background and motivation

2.1 An evolutionary framework of industrial clusters

Traditionally, the study of industrial clusters has been the preserve of 'agglomeration economies', which came to prominence via the works of Marshall (1890) and Jacobs (1969) and Porter (1990). Firms of a given industry were said to choose a certain geographic location to take advantage of local and external returns to scale that suit that firm's industry. Although agglomeration economies resulting from Marshallian externalities cannot be completely ruled out (Boschma 2015), a more recent stream of literature can be found within the realms of evolutionary economic geography, which stresses the role of entrepreneurship by introducing notions put forward by

Schumpeter (1934). This transfer of knowledge involves the creation of a new firm by an employee of an incumbent³ (Agarwal et al. 2004; Klepper 2001; Klepper and Sleeper 2005). The entrepreneur, who establishes a new firm (from here on referred to as a spinoff), transfers organizational routines from the incumbent, thereby giving the spinoff the acquired experience over other types of new firms (Klepper 2001). There are of course other types of entrants in industrial clusters, such as those that, after entering the market, choose to actively relocate their activities to regions with a high degree of sectoral specialization. It is however spinoffs that are the focus of this paper as it is such firms that tend to make up a disproportionately large percentage of firms in industrial clusters, and are often seen as the explanation of cluster emergence, growth and evolution (Dahl et al. 2010). Spinoffs therefore act as vehicles for the transfer of knowledge.⁴ This transfer of knowledge becomes more useful when the industry of the spinoff and parent are increasingly similar. The crux of this stream of thought is the heterogeneity of organizational routines and their role in the performance of firms (Nelson and Winter 1982). Organizational routines, which consist of tacit knowledge acquired by learning-by-doing, are difficult to imitate (Teece et al. 1997), but may be passed down from firm to firm via an entrepreneurial spawning process.

Most prior empirical studies on firm survival in industrial clusters have emphasized the *identification* of localization economies and the relationship with industrial dynamics (e.g. firm entry, growth and exit), but less so on the underlying *mechanisms* that bring about those Marshallian externalities (Rigby and Brown 2015; Van Oort et al. 2012). Frenken et al. (2015) provide a broad survey of empirical studies on industrial clusters. Earlier studies focused on firm characteristics (Geroski 1995; Audretsch et al. 2000), barriers to entry (Geroski 1995), technological conditions within the industry (Agarwal 1998), as well as innovation intensity (Hall 1987) and innovation performance (Cefis and Marsili 2015). Evidence of the returns from location on firm survival is still somewhat weak, if not unclear (Frenken et al. 2015). Nyström (2007) and Renski (2011) find evidence of returns from co-location in only some industrial sectors. Moreover, the lack of geographic precision may contribute to such conflicting results as returns may attenuate sharply with distance (Wennberg and Lindqvist 2010; Rosenthal and Strange 2001).

Other studies of firm survival have found that the absence of cluster effects becomes clear after controlling for the pre-entry experience of spinoff firms. This is an important result as it shows that success may not derive from geographic clustering in its own right but from the experience of the entrepreneur in prior employment before the establishment of the spinoff. Klepper (2007a), in a study of the U.S. car

³Coined as *heritage theory*.

⁴In this paper, we use the definition that a spinoff is a new, distinct and independent firm. This differs from some other definitions in the literature, namely that of Anton and Yao (1995) and Wright et al. (2004), where spinoffs are entities that maintain an organizational link with the parent firm. This can include retained ownership and control of the spinoff's innovations and ideas, equity splits, and in terms of academic spinoffs, use of university resources and academic secondment. This paper instead uses a definition of spinoff closer to that of Chesbrough (2003), where spinoffs are start-ups established by a former employee (now entrepreneur) who, for some reason or another, did not pursue her newly found ideas with the parent firm.

industry, found that spinoffs in particular appear to possess valuable local knowledge which makes it profitable to locate in the same region as the founders' origin, which is the same region as their parent firms. The geographic concentration of automotive firms in the Detroit area, which was the result of four early successful entrants, was the result of a spinoff process alone that caused the industry to agglomerate. Likewise, Moore and Davis (2001) argue that spinoffs in Silicon Valley, which were key to that region's growth, can trace their heritage back to Fairchild Semiconductor. Buenstorf and Klepper (2009) conduct a similar study on the U.S. tire industry and found that the Akron tire cluster grew due to a process associated with the inheritance of knowledge. Heebels and Boschma (2011) reached similar conclusions with the Dutch publishing industry, as new firms with prior experience in publishing had a positive influence on firm survival. Similar conclusions were found with the global fashion industry (Wenting 2008), the German machine tool industry (Buenstorf and Guenther 2011), the U.S. semiconductor industry (Klepper 2010), and the television receiver industry (Klepper 2007a). These studies thus account for the variation of acquired competences from firms in related industries. This advantage even shapes spinoff success after the founder leaves the firm shortly after establishment (Klepper 2002).

Thus, industrial clusters arise from a process where early and successful entrants (with superior organizational routines) give way to a series of successful spinoff firms who in turn inherit those productive organizational routines. Incumbent firms thus serve as 'training grounds' for future firms (Buenstorf and Klepper 2009). Crucially, spinoffs tend to locate in the same geographical area as their parent firms which effectively creates a positive feedback loop where successive generations of firms take advantage of the superior attributes of a region bestowed by previous generations, while simultaneously reinforcing those geographic differences for future generations.⁵ A region hosting an industrial cluster does not even have to enjoy any specific or unique geographic attributes in a physical sense.⁶ The localization economies that result, which include labor market sharing and knowledge spillovers Marshall (1890) as well as the reduced costs of experimentation (Duranton and Puga 2001), are predominantly a consequence of an entrepreneurial spawning process. Carias and Klepper (2010), using matched employer-employee data for Portuguese firms, developed a model of location choice in which entrepreneurs locate their firms in home regions to take advantage of their knowledge of local employees. It was found that for spinoffs in the same industry as their founders, firms are more likely to hire workers from the entrepreneur's prior employer as well as from other firms in the same industry and region, and were also more likely to enjoy superior performance compared to other firms. Specifically, there is a preference for hiring old colleagues.

⁵Klepper (2007b) captures this process in a general model, which shows how organizational heterogeneity and technological change provide the conditions for the entry, exit and market structure of firms in a geographic locale.

⁶Buenstorf and Klepper (2009), in their study of the Akron tire industry, show that apart from benefiting from localization economies arising from the endogenous spawning process of the industrial cluster; the city of Akron, Ohio has few discernible geographic characteristics that would give it any meaningful advantages in the manufacture of tires.

Moreover, firms that locate in their home county as the parent firm had lower annual hazards of exit. The resulting cluster of firms may then even act as a signal for the location decision of non-spinoffs (Suire and Vicente 2009), a Marshallian effect in its own right.

The occupational background of the entrepreneur thus plays a key role in firm survival as well as an explanation of regional industrial clusters. Further studies have shed light on not only the background but also the *incentives* of the entrepreneur. Andersson and Klepper (2013), using matched employer-employee data for establishments in Sweden, focused on the characteristics of employees and eventual founders by identifying *pulled* spinoffs, i.e. firms that continue to have active parents after the time of creation. This contrasts with *pushed* spinoffs, which are firms that do not have a surviving parent when founded. This distinction is important as it may highlight the underlying motives for starting a new firm. Pulled spinoffs are typically the result of individuals with high entrepreneurial talent as well as Schumpeterian motivations.⁷ The entrepreneurs that establish pushed spinoffs, on the other hand, tend to do so to escape unemployment resulting from the closing of their previous firm (Cabral and Wang 2009; Bruneel et al. 2013). It was found that pulled spinoffs have lower hazard rates (i.e. have a higher probability of survival) than other types of new firms in the sample. This was after controlling for more educated workers as well as the number of employees. These results were similar to that of Eriksson and Kuhn (2006) in their study of Danish firms, who found that pulled spinoffs have substantially lower risks of exit. Similar findings can be found in more recent studies for firms in Italy (Furlan 2016) and Germany (Fackler and Schnabel 2016). Furthermore, both Andersson and Klepper (2013) and Eriksson and Kuhn (2006) find that the advantages associated with being a pulled spinoff persist for many years after birth. Andersson and Klepper (2013) find that the benefit of entering the same sector as their parent diminishes after 3 years, concluding that industry-specific knowledge depreciates at a high rate. Another reason for the higher hazard rates in pushed spinoffs may be due to the nature of the inherited knowledge itself, in addition to the incentives of the entrepreneur. Much of the knowledge spinoffs inherit from their parents is tacit, and not easily imitated. An exiting parent firm may suggest tacit knowledge of inferior routines, and the diminished success of pushed spinoffs compared to pulled spinoffs may be the result of inheriting such inferior routines.

Spinoffs in the same or related industry as their parent also enjoy advantages that are not restricted to organizational routines. In comparison to other types of startups, spinoffs enjoy notable advantages that could extend to recruiting the right type of employee out of the result of a better local knowledge of that industry (Agarwal et al. 2004). Furthermore, new firms that maintain their parent-spinoff links can exploit new information and technologies as well as knowledge of new markets that pertain to their industry that they may later exploit (Phillips 2002; Eriksson and Kuhn

⁷The 'Schumpeterian entrepreneur' is a rare and radical generator of ideas who opens up new market landscapes. "It is the entrepreneur who carries out new combinations, who 'leads' the means of production into new channels" . . . "He also leads in a sense that he draws other producers in his branch after him. But as they are his competitors, who first reduce and then annihilate his profit, this is, as it were, leadership against one's own will" (Schumpeter, 1934, 89)

2006). Such relationships may also serve to further the reputation and credibility of the spinoff (Hitt et al. 2001). This may even provide a spinoff several advantages over its parent as they may enjoy the benefits associated with being small while also enjoying the qualities of a larger firm (Parhankangas and Arenius 2003). Moreover, these parent-spinoff linkages tend to serve as an explanation to the empirical regularity that more successful firms tend to spawn more successful spinoffs (Cabral and Wang 2009) and vice versa (Eriksson and Kuhn 2006). Thus there is a distinct merit in studying the parent-progeny relationship between firms.

Agarwal et al. (2004), in a study of the rigid disk drive industry, found that a spinoff's probability of survival increase with the incumbent's (or parent firm's) capabilities at the time of spawning. Specifically, the technological and market-pioneering knowledge of the parent in one year can act as a powerful predictor of the spinoff's knowledge in these areas in the following year. It is this inherited knowledge that gives spinoffs an advantage over other types of new firms. Furthermore, it was found that spinoffs with higher technological knowledge had increased chances of survival. Shane and Stuart (2002), in a separate study of MIT-based university start-ups, found that the social capital of founders is positively related to firm survival. Thus, considering these conclusions, what remains is whether the survival prospects *among* spinoffs can be linked to their comparative levels of inherent social linkages that are resultant of historical genealogical sequences. Controlling for this, one may then consider what role geographic proximity has to play.

2.2 Channels of knowledge and the mechanism of localization

Prior empirical studies on firm survival suggest that firm location is the result of both entrepreneur knowledge of prospective hires as well as industry-specific knowledge. However, the mechanism of knowledge transfer remains unclear. In this paper, we propose that the knowledge gained by the entrepreneur may derive from two different routes; *realized* and *unrealized* knowledge channels.

Figure 1 illustrates the separation of inherited and unrealized knowledge channels, and how these relate to channels the firm may inherit at the time of entry, as well as future post-entry knowledge that may be acquired via informal yet geographically proximate contacts. *Realized* knowledge channels, whether in terms of product technology, routines or information of prospective hires and general industry experience, relays through the genealogical relationships with other firms, both current as well as exited. These knowledge channels are *passively* acquired via pre-entry experience. Often, the knowledge and information relayed through these channels is tacit and is most efficiently transmitted within the boundaries of the firm (Nelson and Winter 1982). The 'liability of newness' of new firms (Stinchcombe 1965), may be mitigated by the pre-entry experience of the entrepreneur and employees via this route. The knowledge transmitted through realized knowledge channels may be embedded in the entrepreneur at the time of entry, *and/or* transmitted at a later date through these established channels. It is therefore important to understand that it is not only *knowledge* per se that gives spinoffs a relative advantage, but the inherited *channels* that carry such knowledge.

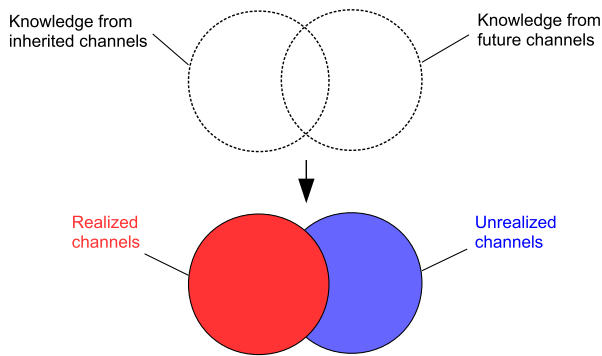


Fig. 1 Realized and unrealized knowledge channels

Wenting (2008) addressed this knowledge channel by mapping the experience of the entrepreneur with past employers in order to analyze localization economies, treating the parent-progeny relationship as the main pathway for routine inheritance. This paper takes this approach further by also addressing multiple generations of firm genealogy. Hence, a firm's characteristics may derive from not only the entrepreneur's previous employer but also the inherent history of those previous employers as well. The result is a genealogical network that arises out of historical relationships. It is through this network structure where the effects of routine replication may be mapped and, importantly, measured. Up until this point, the majority of work on localization economies has only inferred the existence of such a network, and this paper aims to close this research gap by not only measuring such a network, but also analyzing network effects on firm survival within an industrial cluster. Prior network analysis of industrial clusters has focused more on collaboration networks, whether in terms of inventor or intra-industry collaboration (Giuliani and Bell 2005; Breschi and Lissoni 2009; Ter Wal 2013; Boschma and Ter Wal 2007; Fleming and Frenken 2007) or university-industry collaboration (Ponds et al. 2010). In terms of collaboration networks, perhaps a closer comparison of the network and themes in this paper may be made to that of *regional innovation networks*, which are composed of individuals that cooperatively engage in the development of new ideas and subsequently economize on them (Canter and Graf 2007). Using collaborative patent data combined with a firm database, Canter and Wolf (2016) found that individuals' connectivity to an innovation network positively affects its survival, while Canter and Wolf (2018) found a significant (and non-linear) relationship between overall network structure and a firm's probability of survival, but only in the firm's very early stages of its organizational lifecycle. Furthermore, the firm's relationship to that network played no role in future prospects. Such networks, however are more the result of *active* participation rather than an emergent and *passive* one resulting from past historical relationships. A *passive* network of knowledge transmission, on the other hand, is one formed by existing relationships, and these relationships endure over time, space and organizational boundaries (Agrawal et al. 2006). Furthermore, by mapping a firm's connectivity and position in such a network, we get a sense of a

firm's inherited knowledge and routine paths that come about upon a firm's entry to the market. Measuring such a network may thus quantify the origins of that inherited knowledge. In addition, we may also get a sense of the firm's ability to receive knowledge and information via its acquired historical relationships with other firms.

Unrealized knowledge channels refer to sources of information that may be actively sought out, and, as illustrated in Fig. 1, exclude inherited channels. Such channels require some dimension of proximity that enable the potential transmission of knowledge post-entry. We propose that this dimension is a geographical one, which offers an account for any returns not realized through other forms of knowledge transmission. In other words, geographic proximity serves as an enabler for any information flows that don't arrive through acquired channels. There may also be further location advantages that do not necessarily pertain to industry knowledge but access to face-to-face contacts and personal access to prospective clients. This sort of non-market interaction resulting from close geographic proximity, according to Glaeser (2000), not only makes existing relationships easier to maintain but also enables future relationships, as the "transmission of ideas and values depend on sight or hearing. Even if the affected person has not seen or heard the influential person himself, it is often true that he knows someone who has had this personal contact. Obviously, the ability to see or hear depreciates sharply with space" (Glaeser, 2000, 103). Learning may facilitate at the neighborhood level rather than the city level (Durlauf 2004). Rosenthal and Strange (2001, 2008) found that knowledge spillovers tend to occur primarily at smaller distances in certain industries, while the benefits of spatial concentration attenuate sharply with distance. This was after using more detailed units such as zip codes or concentric circles in favor of aggregate geographic units. Maine et al. (2010), using geographic distance to (but not *within*) a cluster to analyze distance effects, found that for firms in certain high-technology sectors, there is a negative effect of distance on firm growth. Andersson et al. (2016), who find that knowledge spillovers fall away after less than 1 km for university workers, argue that knowledge and information spillovers happen at the neighborhood level, as learning effects are only realized through close proximity with other people (Larsson 2014; 2017; Andersson et al. 2014). Geographic proximity to other firms may also act as a hindrance to firm prospects via inadvertent knowledge spillovers and competition effects. Da Silva and McComb (2012) found that among high-technology firms in Texas, co-locating firms within 1 mile (~1.6 km) experienced localization diseconomies, while co-locating firms within 1 and 25 miles experienced positive effects. Firms with advanced capabilities that locate close to other firms may in fact have most to lose via knowledge spillovers unless such knowledge is difficult or costly to learn (Alc cer and Chung 2014). Arzaghi and Henderson (2008), in their study of New York advertising agencies found that even in the narrow confines of Manhattan, localization externalities evaporate after a mere 750 meters. Hence, given this rapid spatial decay, the use of aggregate geographic units comes into question. Moreover, geographic distance plays a key role in face-to-face meetings with clients, employees, and crucially, other advertising agencies in an industry where the sharing of information plays a critical role. Geographic distance therefore acts as a networking cost in this regard. Therefore, to gauge the effect of geography on the survival of the firm, a measure of some degree of precision is required. Detailed or continuous

geographic units may reveal how the various microfoundations of localization economies operate at different spatial scales (Andersson et al. 2016).

We take the view that surviving firms provide a surface depiction of an economic system (in this case an industrial cluster), as firms that offer the set of attributes that are most appropriate to that system are 'selected' to be a part of it (Alchian 1950). Acquired attributes, such as realized knowledge channels, give new entrants a performance premium in a system where they would otherwise be at a disadvantage (Stinchcombe 1965). Surviving firms in a given region then act as sources of realized and unrealized knowledge channels for successive generations of entrants. Thus, we propose that realized and unrealized knowledge channels together provide a mechanism of localization economies, which are in-turn defined by surviving firms. There may even exist a complimentary overlap of these channels. Stuart and Sorenson (2003) show that entrepreneurs find it increasingly difficult to leverage existing social ties when locating further away from their sector's industrial cluster.

3 Data

This paper draws upon the Swedish matched employer-employee data set for the period 1990-2010, which comprises of all individuals, establishments and firms in the country. Surveys are conducted every November. For individuals; records exist for education, gender, employment status, professional status, as well as place of employment. For establishments and firms; the number of employees, geographic coordinates,⁸ county,⁹ as well as the 5-digit Statistical Classification of Economic Activities (NACE) is reported. As this paper focuses on the ICT sector, we restrict analysis to firms with the NACE code of that sector.¹⁰ Table 8 in the appendix provides a summary of these classification codes. We do not consider firms outside these classification codes, and hence cross-sector and academic spin-offs are not considered as a focus of this study. We are hence only interested in the transfer of knowledge and organizational routines of incumbents in the same or related industry of the spin-off firm, i.e. the tacit knowledge that can only be acquired by learning-by-doing from similar or related firms¹¹ (Klepper 2001; Teece et al. 1997).

3.1 Identification of spinoffs

To identify the presence of new firms, we use firm and establishment id-codes (organization numbers) on a yearly basis from 1991 to 2010. Using data for individuals'

⁸Coordinates are based on the SWEREF99 coordinate system, which is a Swedish customization of the WGS84 coordinate system.

⁹This is a top-level local geographic subdivision. There are 21 in total, as of 2017.

¹⁰Due to a change in NACE classification after 2002, Table 8 also provides a translation from its previous 1992 scheme, which was used for all observations before 2002.

¹¹A network incorporating other spinoff types, for example academic spin-offs, would be more appropriately modeled using a second network layer, i.e. a bipartite network, that comes with an implied alternative set of assumptions regarding knowledge and routine transfers.

place of work (or establishment id-code), employees are matched with their employers. To identify employee flows at the establishment level each year, we may compare this information to individuals' place of work for the previous year.¹² Information for founders, however, is limited, and thus new spinoff firms are identified based on a number of conditions. To identify parent-spinoff relationships, we follow three criteria. *First*, if there is only one employee at the firm and establishment level, then we consider that single employee as the firm's founder. It then follows that the founder's previous establishment is considered the spinoff's parent. *Second*, if any individual at a new establishment has a professional status recorded as 'entrepreneur', then that individual is considered the founder of the new firm and the parent-spinoff progeny is identified using the entrepreneur's previous workplace. *Third*, and if a founder still cannot be identified using the first two methods, then we identify the parent-spinoff relationship by the largest contingent of employees in the new firm that have the same previous employer. Spinoffs with multiple parents are possible, using these three criteria.¹³ If we exhaust these three criteria, the new firm cannot be identified as a spinoff and is thus treated as a *de novo* firm instead.

Finally, using the same logic as Andersson and Klepper (2013) and Eriksson and Kuhn (2006), if a spinoff firm came into being in the same year as the parent firm exited, then that firm is considered a pushed spinoff. If not, then it is considered a pulled spinoff.

3.2 Network identification

Wenting (2008) constructed a genealogy of parent-spinoff relationships to study the effects of routine replication upon firm success by identifying an entrepreneur's previous employers. Here, we expand upon that principle by constructing a network based upon a firm's ancestral background by considering multiple generations of spinoffs, a model that is refinement of that introduced by Bagley (2018). Upon identification of a parent-spinoff relationship, a link¹⁴ is established between those two firms, as it is assumed that spinoffs, through their founders, have knowledge of their former workplaces. Furthermore, a new spinoff firm may also form several additional links per a series of genealogical rules. The first of these is a 'sibling' relationship with other spinoff firms founded in the same time period and from the same parent firm. This is based on the assumption that if a firm spawns two or more firms in the

¹²Spinoffs are formally identified as new firms with a founder employed at a prior firm in an *adjacent year*, a system used by Eriksson and Kuhn (2006). Extending the number of years between an employee leaving a firm and becoming an entrepreneur broadens our assumptions on what defines a spinoff (both in terms of entrepreneurial incentives as well as parent-spinoff linkage strength), and hence we keep the time dimension strict.

¹³However, after the third criterion, no such cases were identified

¹⁴Links are undirected, meaning that they have no orientation. Thus, information may flow in both directions of a link. Agrawal et al. (2006) showed that enduring social relationships contribute to information flows from individuals to their former workplaces, as a result of prior employers maintaining their relationships with former employees.

same period, there is a reasonable chance that the founders of each firm are familiar with each other as they had previously worked at the same firm and at the same time. Figure 2 illustrates this. Consider the diagram in the upper left corner, which maps spinoffs in four time periods. Note that the arrows represent parent-spinoff flows, not strictly the network linkages themselves. In time period $t + 1$, firm **A** spawns firms **B** and **C**. Now consider the network diagram for time period $t + 1$ in the upper right corner. Each firm is represented by a node and labeled as such. Firm **A** forms a link with **B** and **C** due to their parent-spinoff relationship. Moreover, firms **B** and **C** also establish links with each other. This rule would only hold if **B** and **C** spawned from the same firm at the same time period. There is therefore an assumption of diminished familiarity between firms that spawned from the same parent but in different time

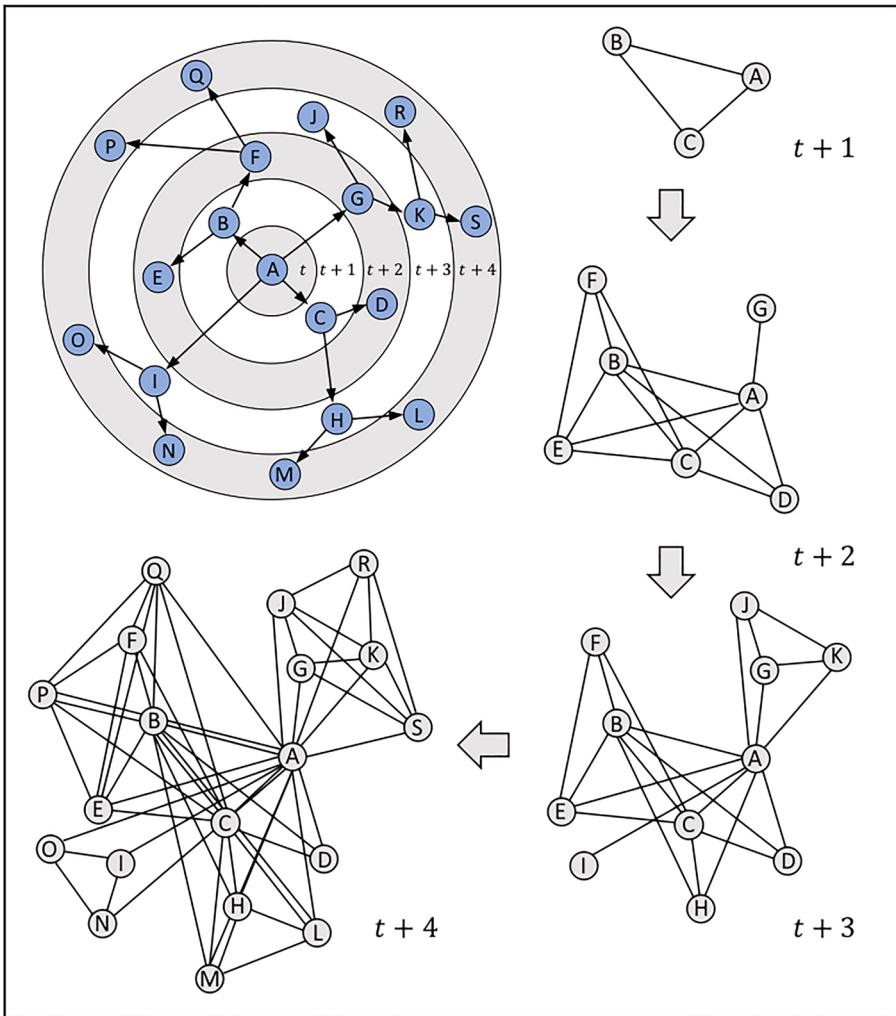


Fig. 2 Spinoff network rules

Table 1 Spinoff network rules

Type of link	Explanation
Parent-spinoff	A link formed between a parent and spinoff firm.
Sibling	If a firm spawns more than one firm in a single time period, then links form between those spinoffs.
Inheritance	A spinoff will inherit all its parent firm's links, but only at the time of spinoff.

periods. This is shown in period $t + 2$, when **A** spawns **G**, but **G** does not form a link with **B** or **C**.

In $t + 2$, **C** spawns **D**. Firm **D** does not form any sibling relationships as **C** does not spawn any other firm in that time period. However, firm **D** does inherit several links from **C**, namely to **A** and **B**. This is because it is assumed that the founder of firm **D** has knowledge of its parent's connections. If firm **C** derives its genealogy from **A** and **B**, it should not be overlooked that its offspring does so as well. These inherited links signify potential knowledge flows that *originate* from a given firm, but not necessarily directly from it in a tangible sense. Thus, the inheritance rule is one that recognizes the genealogical legacy of firms, and the knowledge and routines that this entails.¹⁵

The remaining network diagrams in Fig. 2 for time periods $t + 3$ and $t + 4$ further illustrate these rules, which may be summarized in Table 1. Links can only be inherited backwards in time, not forward. Hence, all firms in $t + 1$ onward, **B** through **S**, can trace a relation to firm **A**. Furthermore, as the rules stipulate that links between sibling firms may only occur between siblings born in the same period, this creates a 'branching' effect of localized network grouping. This becomes more apparent in the network diagram for $t + 4$. It should be noted that network links are not weighted in any way. This is to simplify the analysis while also avoiding any assumptions to what those weights may be.

The resulting network does not only consider spinoffs. It is not unreasonable to further address the linkages that arise from multiple establishments for one firm, as well as mergers and splits. A summary of these rules (which are not as complex as those that concern spinoffs) may be found in Table 9 in the appendix. It should be noted that in this framework, nodes representing exited firms are not removed from the network. Although a firm may cease to exist, knowledge of that firm lingers. Thus, the network preserves memory. We derive the network using all observations in Sweden without restricting it to any geographic sub region (later empirical analysis uses observations in Stockholm county only).

¹⁵To give an example, consider a chain of spinoffs. Assume no siblings. If we would consider a chain of spinoffs over 20 years, with each subsequent spinoff spawning a new firm in the next, there would be no way to recognize the importance of this legacy as it would be one simple chain with a long path from the newest to the oldest firm. But if we were to "reward" nodes with their genealogical legacy, the resulting network would have a greater meaning. This is to avoid the potentially ambiguous weighting of links and nodes.

3.3 Variables

3.3.1 Closeness centrality

The passive network derived in Section 3.2 provides a basis for identifying realized knowledge channels. In networks, centrality measures the ‘importance’ of nodes. ‘Importance’ can mean many different things, however, and how one wishes to define ‘importance’ depends on what one wishes to consider. If ‘importance’ characterizes a network’s walk structure, one of the simpler measures of centrality is closeness centrality, originally introduced by Bavelas (1950), who defined the measure as a means of assessing the effective flow of information that is also efficient for human effort. Formally, a normalized¹⁶ measure of closeness centrality of firm i may be defined as:

$$C_i = \frac{N}{\sum_j d_{ij}} \quad (1)$$

where d_{ij} is the geodesic distance, or shortest path, from firm i to firm j , and N is the number of firms in the network. If firm i and firm j share a link between them, then the geodesic distance d_{ij} would be 1. Similarly, if it took two links for i to reach j (thereby passing through an intermediate firm), then the geodesic distance would be 2 and so on. Thus, closeness centrality measures the average of shortest paths between firm i and all other firms in the graph. Therefore, the higher the closeness centrality of a firm, the closer it is to *all* other firms in a network sense. In other words, the more central a firm’s position in a network, fewer links are required to reach other firms, hence its higher closeness centrality score. Thus, the measure does not ignore *how* a firm is connected, and considers the firm’s *position* in the network. In economic intuition, closeness centrality may be interpreted as the efficiency of transmission of information. The higher a firm’s closeness centrality, information from other firms is more accessible, and this level of access is less dependent on other firms (Powell et al. 1996). Higher closeness centrality implies further *acquired* understanding of other firms, which in-turn enhances the exchange of information between them. In the context of this paper, this inter-firm understanding is the result of genealogical linkages.

Furthermore, we only measure firms in the *main component* (the largest connected component) of the network. This is to give further weight to the measurement by ‘rewarding’ firms that are in the main component. All other firms are automatically given a centrality score of zero.¹⁷ The main component of the network in 2010 may

¹⁶Often, one considers the *normalized* closeness centrality measure as it allows for comparison between networks of different sizes. As network sizes differ from year to year, the normalized form of the measure is used in this paper.

¹⁷It should be noted that if all firms were given a centrality score, then not only would firms outside the main component have a centrality very close to zero, but all firms in the network would also. This is because the resultant network compiled from the rules in Section 3.2 is quite disconnected. Other network measurements are of course possible to measure nodes in disconnected networks (for example, *harmonic centrality* as first proposed by Marchiori and Latora (2000)). Such measures are however not suitable for networks with so many disconnected components, and would also fail to give precedence to firms in the main component, which is one goal in seeking an appropriate measure. Measuring the network from the perspective of the main component is not a new concept. See Bagley (2018), Borgatti (2006), Gulati et al. (2012), and Davis et al. (2003) and Uzzi and Spiro (2005) for further reasoning behind this method.

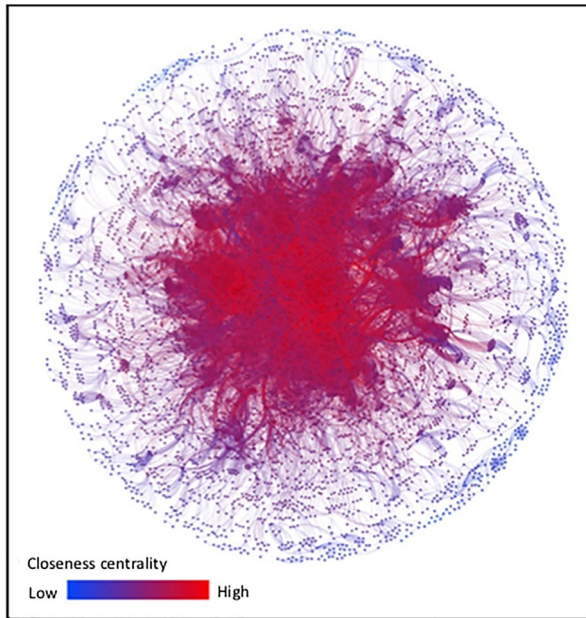


Fig. 3 The Main Component of the Network, 2010

be seen¹⁸ in Fig. 3. There is some suggestion of local clustering, and some of these local clusters display a higher level of closeness centrality than others. Closeness centrality is very high in areas (illustrated in the core of the network diagram) where there is a high level of interconnectivity between these local clusters. Furthermore, higher closeness centrality is the result of being located within the middle of the network, implying shorter average paths to other firms. The network, which is used for empirical analysis, considers all of Sweden as one cannot discount the inter-firm linkages between geographical locales outside of Stockholm. As closeness centrality proxies for the *realized* knowledge channels of firms, it has no geographic restrictions. Firms may access such knowledge from any geographic source.

3.3.2 Geographic distance

Where closeness centrality may proxy for realized knowledge channels and hence *acquired* knowledge and routines, we propose that geographic distance proxies for *unrealized* knowledge channels and the *potential* of future post-entry knowledge. For the most part, and mostly due to the lack of appropriate data, previous studies have relied on aggregate areal geographic units (such as counties, metropolitan areas,

¹⁸This network has been drawn using the Fruchterman and Reingold (1991) layout, which is a force-directed network drawing. Links tend to have an equal length, and do not cross. An efficient way of achieving this is by placing firms with a high centrality in the middle of the displayed network. Note that this representation has no geographic interpretation in any way.

etc.) to compare firms inside and outside an industrial cluster. Such an approach is problematic as it ignores any potential learning effects that may decay sharply with geographic distance, which in-turn may reflect different mechanisms that instigate localization economies in a cluster (Wennberg and Lindqvist 2010; Rosenthal and Strange 2001).

Here, geographic distance is measured as the distance from the cluster's geographic centroid (which in-turn is simply the mean of the x - and y -coordinates of each existing ICT firm in Stockholm county). Using the same methodology as De Vaan et al. (2013), it is measured 'as the crow flies', or the firm's Euclidean distance from the cluster's centroid.¹⁹ In general, one may envisage that geographic distance from the cluster core increases the hazard rate of firm exit. However it is not expected that the effect of geographic distance from the cluster's core on firm survival is linear. Geographic proximity can be a cost as well as a benefit (Da Silva and McComb 2012; Alc ecer and Chung 2014). One can imagine that firms located closer to the core may expect negative effects from competition and unintended knowledge spillovers, which outstrip any gains from localization. Hence, geographic distance is broken down into two concentric rings; one at 1 km, another at 5 km, and a final category for firms outside the 5 km ring and terminating at the institutional boundary of Stockholm county. There is intuition behind this choice of radii, as there is a clear and sharp drop in firm density after the 5 km threshold.²⁰ A firm density map of ICT firms in Stockholm county in 2010 may be seen in Fig. 4, which also depicts the two concentric rings. There is a noticeable concentration of firms in and around the boundaries of the first two distance categories, and this concentration gradually fades away with increased distance from the cluster's centroid.

Other geographic measurements are of course possible. For example, given a firm or establishment, one could count the number of other firms within a certain radius (or radii) as a means of measuring co-location and thus unrealized knowledge channels. However, such a measure would fail to account for other externalities associated with the urban landscape that may have an effect of firm survival, such as costs. Furthermore, given the geographic scale of this paper (the Stockholm ICT cluster), such a measure may not be entirely appropriate when understanding the dynamics of firms *within* clusters, rather than *between* clusters. In other words, firm counts would be better suited for global-, rather than local-level analysis. Even though the Stockholm cluster is not spatially distributed evenly (as seen in Fig. 4 and largely due to urban development patterns), firms do tend to be concentrated toward a hypothetical core, and within a compact geographic area. Distance from the cluster's geographic centroid, on the other hand, also incorporates aspects of urbanization and the externalities associated with co-location with unrelated firms (Jacobs 1969), as well as costs. Furthermore, location alongside firms that are related but not in the five-digit NACE sector may reveal externalities associated with related variety (Frenken et al.

¹⁹Formally, this geographic distance may be defined by $D_i = \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2}$, where (x_i, y_i) represent the geographic coordinates of firm i , and (\bar{x}, \bar{y}) is the cluster's centroid, calculated as the mean x - and y -coordinates of all ICT firms recorded by Statistics Sweden in a given year.

²⁰One may gauge this from the descriptive statistics in Table 3.

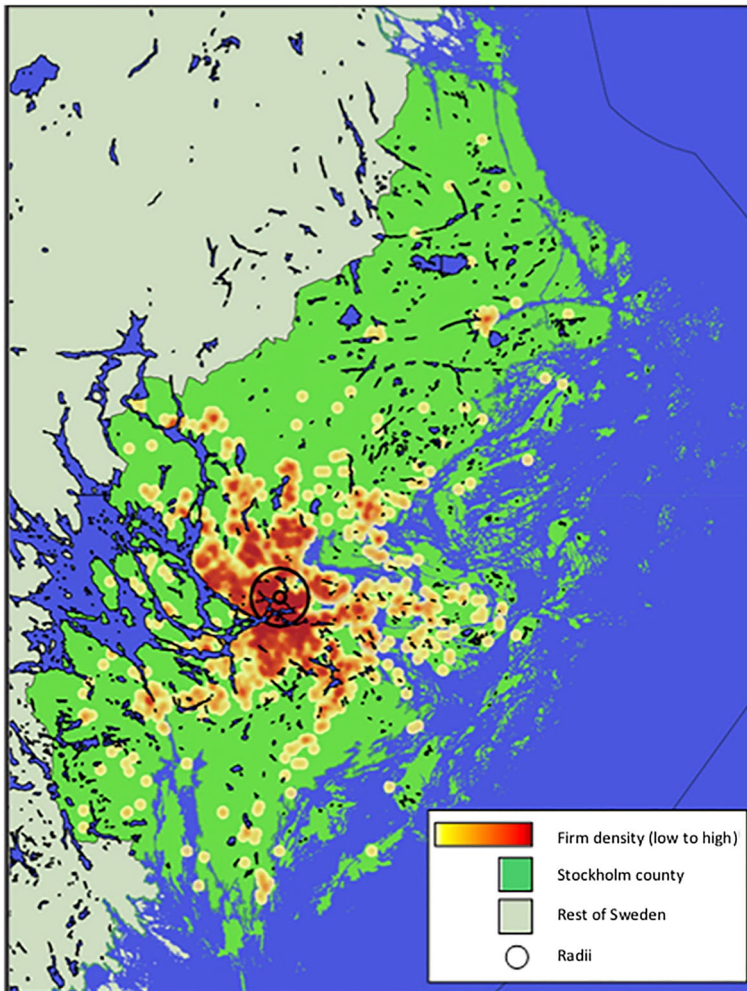


Fig. 4 ICT firm density in Stockholm county, 2010

2007)). This is in addition to proxying for access to specialized labour inputs as well as unrelated but associated sectors.

3.3.3 Control variables

In addition to the main variables of interest we also include several control variables that may explain firm survival rates. This includes the founding size of the firm, measured by number of employees (including the founder); the size of the spinoff's parent firm (at the time of spinoff), also measured in the number of employees, the fraction of the firm's initial workers that are highly educated (measured by the percentage of employees with ≥ 3 years of university education); and the fraction of the

firm's initial workers that are male. Similar to Andersson and Klepper (2013), we also include a number of dummy variables including those for time periods in five-year increments (adjusted to take into account a three-year increment for 1993-1995 when Sweden experienced a sharp recession); if the spinoff is in a different five-digit NACE industry as its parent firm and whether or not the spinoff is a pulled spinoff. A summary of the variables may be found in Table 2.

We hypothesize that the initial size of a spinoff has a negative influence on hazard of exit, as found in prior research (Audretsch and Agarwal 2001; Mata and Portugal 1994; Andersson and Klepper 2013). It can be said that increased firm size increases access to capital and knowledge (Falck 2007). The effect of *parent* size on spinoff performance, however, has been conflicting in previous literature. On the one hand, Andersson and Klepper (2013) found that spinoffs of larger firms generally perform better. Hvide (2009), when considering firms of two or more employees with a majority owner, larger parent firm sizes result in an increase in the return of assets of spinoff firms. On the other hand, parent size has been shown to have a negative effect on spinoff performance. For spinoffs with a single owner or self-employed individuals, Sørensen and Phillips (2011) and Elfenbein et al. (2010) found evidence for this, which supports the view that smaller parent firms provide employees with a work environment that fosters entrepreneurial ability. Firms with employees of higher levels of education attainment are expected to have a higher chance of survival, as seen in the previous literature (Andersson and Klepper 2013; Eriksson and Kuhn 2006). Prior

Table 2 List of variables and descriptions

Variable	Description and anticipated effect on hazard rate
D_i	Distance of firm i from the cluster's centroid. Proxies for <i>unrealized</i> knowledge channels. Separated into three dummy variables; D_i^{0-1km} , D_i^{1-5km} and $D_i^{>5km}$, which take on the value of 1 if firm i is located within that category, or 0 otherwise. (+/-)
C_i	Closeness centrality of firm i , which proxies for <i>realized</i> knowledge channels, i.e. the efficiency of knowledge transfer between firms via genealogical links. (-)
\ln Founding Size $_i$	The size of firm i at the time of spinoff. (-)
\ln Parent Size $_i$	The size of the parent firm of firm i at the time of spinoff. (+/-)
Share Highly Educated $_i$	Fraction of firm i 's initial workers that have ≥ 3 years of university education. (-)
Share Male $_i$	Fraction of firm i 's initial workers that are male. (-)
Period n_i	Time period dummy, takes on the value of 1 if firm i was established in period n , and 0 otherwise. (+/-)
Sector Change $_i$	Dummy variable, takes on the value of 1 if firm i is in a different five-digit NACE industry as its parent firm, and 0 otherwise. (-)
Pulled Spinoff $_i$	Dummy variable, takes on the value of 1 if the firm i is a pulled spinoff, and 0 otherwise. (-)

Anticipated effects on hazard of firm exit in parentheses

research has found that spinoff firms (both pulled and pushed) have a lower hazard of exit with higher percentages of male employees (Rocha et al. 2018; Andersson and Klepper 2013; Eriksson and Kuhn 2006). This is largely a contextual factor, as male employees tend to exhibit longer tenure with their previous firm (Rocha et al. 2018). We employ time period dummies to account for any external effects pertaining to general macroeconomic performance. The effect of establishing a firm in a different sector of the parent firm is anticipated to be negative, as inheriting competencies from parent firms would not be as easily transferable due to a fall in cognitive proximity (Andersson and Klepper 2013). Furthermore, Sapienza et al. (2004) found that limited knowledge overlap between spinoff and parent firm hinders firm prospects due to limited knowledge search and assimilation. Finally, it is expected that pulled spinoffs have a lower rate of hazard compared to pushed spinoffs, as seen in the literature (Andersson and Klepper 2013; Eriksson and Kuhn 2006). We reason that this is due to underlying entrepreneurial incentives. While pushed spinoffs may be the result of a business opportunity perceived as profitable and novel, pushed spinoffs arise from a firm exit, which may imply the transfer of redundant products or inferior routines.

3.3.4 The sample

Sweden's Stockholm County contains a disproportionately high number of ICT firms, exhibiting regional specialization in the sector (Karlsson et al. 2004). We also take the view that economic geography as a system that operates in a continuum, with all firms subject to an economic center of gravity to some degree (even if minuscule). Given Stockholm County, which encompasses an area endowed with both rural and urban areas, we may compare firms at various levels of agglomeration that are subject to the same localization forces, rather than comparing firms that are subject to different localization forces with different economic centers of gravity.²¹ The region, and the city in particular, is also host to a number of universities and other institutes of higher education, all of which are centrally located. Audretsch and Lehmann (2005) found that firms located closer to universities are positively influenced by their research output, and Cassia et al. (2009) found the growth rates of young firms are positively linked to the knowledge output of universities. For empirical analysis, this paper includes all identified spinoffs, both pushed and pulled, in Stockholm county that entered from 1991-2010.²² These spinoffs must also have a parent firm in Stockholm county to be considered for analysis. The survival analysis used in this paper is not time-variant, and therefore only the year of entry, i.e. the first year of observation of each sample, is considered. Thus, the aim is to analyze the survival impact of the *initial* conditions of the spinoff firm. The descriptive statistics of this sample, which amount to 4,502 observations, may be seen in Table 3. We consider all

²¹If we were to expand our analysis outside the politically-defined boundaries of Stockholm County, for example, we would then begin including firms subject to the economic forces that pertain to two nearby and congruent cities; Uppsala and Nyköping. Restricting our analysis in geographic scope thus allows us to compare firms within a cluster, rather than firms across clusters (which are subject to alternative externalities).

²²This paper uses observations starting in 1990. However, given this range, 1991 is the earliest year one identify *spinoffs*, because to do so requires the identification of a parent firm.

Table 3 Descriptive statistics

Variable	All firms				Pulled spinoffs				Pushed spinoffs			
	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
D_i	8.611	8.066	26.43	84,676	8.661	7.958	26.43	82,193	8.410	8.487	87.47	84,676
$D_i^{0-1 km}$	0.0393	0.194	0	1	0.0400	0.196	0	1	0.0366	0.188	0	1
$D_i^{1-5 km}$	0.394	0.489	0	1	0.383	0.486	0	1	0.438	0.496	0	1
$D_i^{>5 km}$	0.567	0.496	0	1	0.577	0.494	0	1	0.525	0.500	0	1
C_i^+	0.0655	0.104	0	0.376	0.0748	0.109	0	0.376	0.0285	0.0697	0	0.306
In Founding Size _{<i>i</i>}	0.358	0.673	0	3.850	0.315	0.668	0	3.850	0.528	0.665	0	3.219
In Parent Size _{<i>i</i>}	3.108	1.697	0	7.554	3.490	1.548	0.693	7.554	1.582	1.379	0	6.816
Share Highly Educated _{<i>i</i>}	0.404	0.454	0	1	0.415	0.461	0	1	0.358	0.424	0	1
Share Male _{<i>i</i>}	0.850	0.309	0	1	0.856	0.310	0	1	0.826	0.300	0	1
Period 1 _{<i>i</i>} (1991-1992)	0.0258	0.158	0	1	0.0236	0.152	0	1	0.0344	0.182	0	1
Period 2 _{<i>i</i>} (1993-1995)	0.0506	0.219	0	1	0.0503	0.219	0	1	0.0522	0.222	0	1
Period 2 _{<i>i</i>} (1996-2000)	0.256	0.437	0	1	0.263	0.440	0	1	0.231	0.422	0	1
Period 2 _{<i>i</i>} (2001-2005)	0.310	0.463	0	1	0.298	0.457	0	1	0.358	0.480	0	1
Period 2 _{<i>i</i>} (2006-2010)	0.357	0.479	0	1	0.365	0.482	0	1	0.324	0.468	0	1
Sector Change _{<i>i</i>}	0.340	0.474	0	1	0.343	0.475	0	1	0.327	0.470	0	1
Pulled Spinoff _{<i>i</i>}	0.800	0.400	0	1								

Total observations: 4502 (3601 pulled spinoffs and 901 pushed spinoffs). Sample includes spinoff firms that also have a parent firm in Stockholm county in the year of entry. Years of observation: 1991-2010. [†]70.1% of total observations are not in the main component of the network, and therefore receive a closeness centrality score of zero.

firm sizes, including firms with an initial employee of 1. As we consider *initial* firm size, and not *current* firm size, we cannot discriminate between self-employed individuals and firms that will eventually expand. Removing firms with an initial size of 1 employee would thus be arbitrary.

There is considerable variation with firm distance from the cluster, with some spinoffs locating on the fringe of Stockholm county. More than half locate outside the 5-km core (at 56.7%), and a significant minority locate within 1 km of the core (only 3.9% of observations). This shows that there is a high firm density within the 5-km ring. Closeness centrality is, on average, relatively low but also somewhat dispersed and skewed. This is likely due to most observations (70.1%) automatically receiving a score of zero due to being outside the main component of the network. Both the founding size and parent size (which are logged) display somewhat normally distributed behavior. On average, 85% of employees in spinoffs are male, with less than half of employees possessing a university education of 3 years or more (40.4%). The first five years, which includes the time period where Sweden experienced a sharp recession, was a low point in terms of the number of new spinoff firms, with only 7.6% of firms in the sample established during that time. Overall, the number of new spinoff firms increased with each successive time period. Andersson and Klepper (2013) infer that this increase in spinoffs over time may be the result of increased incentives for pulled spinoffs during periods of high economic growth. In terms of industrial similarity, 34% of spinoffs in the sample switched to a different five-digit ICT industry to their parent, and 80% of observations in the sample are pulled spinoffs. When comparing pulled and pushed spinoffs, pulled spinoffs tend to exhibit a higher degree of closeness centrality on average, and with greater variation, than pushed spinoffs. This is despite following the same set of network rules as laid out in Table 1, indicating that pulled spinoffs are more likely to find themselves in a more complex sequence of spawning events. In terms of founding size, pushed firms tend to be larger, possibly an indication of the conservation of previous labor profiles of closed firms. Pulled spinoffs, on the other hand, tend to come from larger parent firms, which adds to the claim that entrepreneurs have a greater incentive to leave larger firms due to their inferior ability in spotting novel ideas due to increased bureaucracy (Hvide 2009). Pulled spinoffs tend to have a higher percentage of workers with a university education of 3 years or more compared with pushed spinoffs, which is a repeat of the pattern found by Andersson and Klepper (2013). In terms of distance from the cluster core, sector change, percentage of male employees, and time period of firm foundation; there is no significant difference between pulled and pushed spinoffs.

The correlation matrix (which forgoes distance and time dummies), may be seen in Table 4. A small yet positive relationship between increased distance from the cluster core and closeness centrality may reveal a substitutive relationship between these two measures, with spinoff firms locating further away from the core doing so only due to their genealogical relationships with other firms and the advantages this entails. There is a negative relationship with closeness centrality and founding size, revealing that network position favors larger initial firm sizes, and a positive correlation with centrality and parent sizes, inferring that larger firms (and therefore their subsequent spinoffs) tend to be part of the main component of the network. Larger spinoffs also tend to locate closer to the cluster than away from it.

Table 4 Pairwise correlation matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) D_i	1							
(2) C_i	0.0729***	1						
(3) \ln Founding Size $_i$	-0.1744***	-0.2351***	1					
(4) \ln Parent Size $_i$	0.0233	0.5588***	-0.0707***	1				
(5) Share Highly Educated $_i$	-0.0917***	0.1281***	-0.0382**	0.1079***	1			
(6) Share Male $_i$	0.0030	0.0010	-0.1267***	0.0185	0.0167	1		
(7) Sector Change $_i$	-0.0250*	-0.0437***	0.1277***	0.0082	-0.0214	0.0225	1	
(8) Pulled Spinoff $_i$	0.0124	0.1786***	-0.1264***	0.4499***	0.0507***	0.0388***	0.0131	1

* $p < 0.1$, ** $p < 0.5$, *** $p < 0.01$

Closeness centrality is positively correlated with pulled spinoffs, suggesting that entrepreneurs that establish firms due to Schumpeterian motivations tend to do so with a healthier attachment to the genealogical network of firms and hence have a greater potential of knowledge transfer from parent to spinoff. This implies that they may have a greater level of acquired industry-specific knowledge. This resonates with the theoretical assumptions of pulled spinoffs. Pulled spinoffs also tend to have an educated workforce, which is also positively correlated with closeness centrality. Higher centrality scores are negatively correlated with firms that switch their five-digit industry from that of their parents. Similarly, firms that locate closer to the core of the network are more likely to have the same five-digit industry code as their parent firms, and have more educated employees. Pulled spinoffs are also associated with larger parent firms. This resonates with Gompers et al. (2005), who found that, in absolute terms, the largest companies tend to spawn the most firms.²³ Of note is a negative correlation between founding size and parent size, as well as the negative correlation between founding size and the share of highly educated employees. Although empirical regularity would suggest otherwise (Andersson and Klepper 2013; Klepper 2009; Franco and Filson 2006), this is likely due to including firms with a founding size of one employee.

3.3.5 Censoring

Firm survival is not necessarily the opposite of firm exit. An entrepreneur may create a firm with the explicit aim of selling the firm to larger firms (Frenken et al. 2015; Andersson and Xiao 2016), and hence, exit is not always equivalent to firm failure. Indeed, De Vaan et al. (2013) found that the main factors that reduce the probability of exit by failure are the same as those that increase the probability of exit by acquisition. Hence, 'successful exits' are a possibility. We censor the survival data to consider successful exits, i.e. those firms that exited due to merger with other firms as well as those that split into new firms. Formally, exited firms with more than 50% of their employees moving to a new employer (indicating an ownership change) are treated as censored. We also censor firms surviving in 2010, which is the last year of observation.

4 Estimation and discussion of results

As 1,452 firms exited before the end of their first year, these observations are removed from any subsequent survival analysis. Table 5 shows the survival rates of firms in the sample, which may also be illustrated with the Kaplan-Meier survival curve²⁴ in

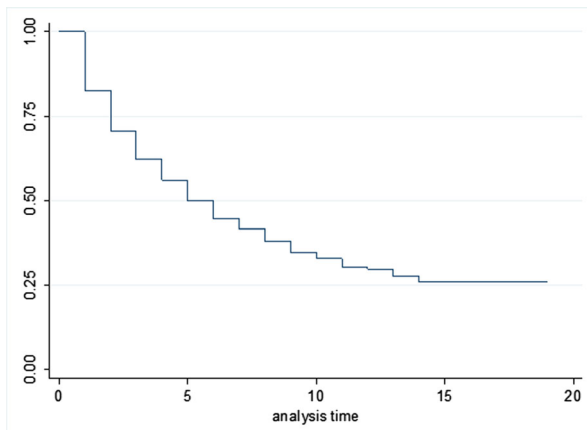
²³Gompers et al. (2005) stressed, however, when measured in terms of annual spawning per million of employees, smaller firms tend to spawn more frequently than larger firms. This was confirmed by Elfenbein et al. (2010) and Eriksson and Kuhn (2006).

²⁴This estimate, proposed by Kaplan and Meier (1958), is formally given by $\hat{S}(t) = \prod_{j|t_j \leq t} \left(\frac{n_j - d_j}{n_j} \right)$, where n_j is the number of firms at risk and d_j is the number of failures at time t_j .

Table 5 Survival rate of spinoff firms, 1991–2010

Year	Number of firms	Number of exits	Survival rate
1	3050	530	0.8262
2	2277	336	0.7043
3	1709	203	0.6206
4	1288	126	0.5599
5	988	109	0.4982
6	769	79	0.4470
7	608	44	0.4146
8	483	40	0.3803
9	387	35	0.3459
10	285	15	0.3277
11	207	17	0.3008
12	158	2	0.2970
13	109	8	0.2752
14	65	4	0.2582
15	43	0	0.2582
16	26	0	0.2582
17	18	0	0.2582
18	12	0	0.2582
19	4	0	0.2582

Figure 5, which illustrates the fraction of firms surviving to each age after censoring. This is a nonparametric estimate of the survivor function $S(t)$, i.e. the probability of survival (or failure) past time t . Roughly 50% of firms survive after five years after entry, and about 33% survive after 10 years. After 14 years, roughly 26% survive, which continues until the remaining year of the study period.

**Fig. 5** Kaplan-Meier survival estimate

We begin by estimating the effect of each of the main variables of interest separately on firm hazard, i.e. each of the three distance dummies as well as closeness centrality. We use the Cox (1972) proportional hazards model, which formally estimates the hazard rate for the j^{th} subject, $h(t | \mathbf{x}_j) = h_0(t) \exp(\mathbf{x}_j \beta_x)$, where \mathbf{x}_j are the explanatory variables, β_x their corresponding coefficients, and $h_0(t)$ the baseline hazard. The Cox model has several advantages over parametric methods, the most important of which is that there is no need to assume and hence specify a distribution for the baseline hazard $h_0(t)$ as it is left unestimated.²⁵ We consider the entire sample, and then repeat these four regressions for pulled and pushed spinoffs only. The results of these 12 estimations may be seen in Table 6. For the distance variables, a negative coefficient implies a lower hazard of exit compared to the other distance categories, and a positive coefficient implies a higher hazard. Similarly, for a continuous variable like closeness centrality, a negative coefficient implies that a higher value lowers the hazard of firm exit. When considering all types of spinoffs, the effect of increasing geographic distance on firm survival (or exit) is nonlinear. Spinoff firms located within the 1 km core of the cluster display a statistically significant and decreased hazard of exit compared to firms in other reference categories. Firms located in the middle category (1 to 5 km), on the other hand, display an *increased* hazard. This shows that there is an extremely rapid spatial decay in the benefits of locating close to the cluster's core, which resonates with the findings of Arzaghi and Henderson (2008). One may also envision that any localization economies gained from locating within a 1 km radius diminish as competition effects take over. This would coincide with the results by Sorenson and Audia (2000) and Stuart and Sorenson (2003) who found that firm performance within denser industrial clusters can be worse than firms in less concentrated regions due to competition effects. Combes et al. (2012), in a study of French regions, sought to distinguish between agglomeration economies (which may be enforced by localized natural advantage) and *firm selection*, building upon the work of Melitz (2003). Firm selection is the idea that larger markets attract more firms, which makes competition tougher. Thus, with stronger Darwinian selection, the less productive firms exit. Moreover, Combes et al. (2012) found that stronger firm selection in larger cities left-truncates a firm productivity distribution. Agglomeration effects, on the other hand, both dilates and shifts the firm productivity distribution to the right. In other words, while agglomeration effects have a positive impact on total firm productivity, firm selection effects create a lower-bound that favors more productive firms. Differences in firm productivity may serve to overcome localized costs such as the rents and wages associated with the urban core (such as those discussed by Glaeser and Maré (2001)) which do not widely differ on the neighborhood level. Thus, when one considers the intermediate radius of D_i^{1-5km} , localized costs increases selection pressure and thus the competition between firms. Although the cluster may still serve to promote entrepreneurship, it can worsen the performance of new firms. On the other hand, it may be inferred that the innermost radius, D_i^{0-1km} , populate with firms that are already equipped with the endowments that can cope with the higher localized costs. Hence, the hazard of

²⁵It is for this reason that the Cox model has no intercept (as it is subsumed into the baseline hazard).

exit is negative for firms in that category. Finally, firms that locate in the outermost radius, $D_i^{>5\text{km}}$, face lower costs yet benefit less from geographic proximity, and the resulting hazard ratio is negative but low in magnitude. Therefore the debilitating effect, if any, falls away for firms located outside the 5 km radius as the hazard rate for that category is notably smaller compared to other distance radii. Hence, although firms generally realize the benefits of co-location at this point, and even though those benefits are less than firms located within the 1 km radius, any supposed effects of competition and unintended knowledge spillovers have fallen away to an extent where the hazard of exit once again falls.

The hazard of exit with increased closeness centrality is statistically significant and below zero. This adheres with the theory that the closer a spinoff's genealogical ties to other firms, the greater the probability of survival. A hypothetical increase in one unit of the centrality measure would almost half the hazard rate of exit for when considering all types of firm. Thus, if we can proxy a spinoff's inherited knowledge network via closeness centrality, one could conclude that spinoff firms have a greater chance of survival when they are centrally-positioned in a chain of knowledge channels derived from parent-spinoff linkages. If these knowledge channels indeed imply the transmission of industry-specific knowledge and routines, we can infer that spinoffs gain a performance premium from such. This result is similar to that of Raz and Gloor (2007), who found that the long run survival for firms is correlated with being centrally placed in their founder's professional networks.

Repeating the above bivariate regressions but restricting ourselves to pulled and pushed spinoffs gives similar results with respect to the distance measures. However, they are not always statistically significant. The pattern repeats for pulled spinoffs although statistical significance falls away after 5 km. When comparing pushed spinoffs, on the other hand, there is no statistically significant effect with any of the distance measures. The magnitudes of the hazard ratios for closeness centrality are markedly different when considering pulled and pushed spinoffs separately. Although both display a decreasing hazard with respect to the measure, this hazard rate is notably higher for pulled spinoffs (compared to all firms), and considerably lower for pushed spinoffs. Hence, closeness centrality has a markedly different effect for different types of spinoff firms. Entrepreneurs that establish pushed spinoffs, when facing a different set of incentives (e.g. to escape unemployment), may carry a different set of attributes unrelated to those that establish pulled spinoffs. The Schumpeterian abilities and talents associated with entrepreneurs of pulled spinoffs are innate, and hence genealogical linkages, although still critical and an asset for survival, are not as vital. Entrepreneurs and founders of pushed spinoffs, on the other hand, who may not possess such Schumpeterian attributes, may rely on genealogical connections to an even greater extent, so much so that it can be an essential aspect of firm survivability. It can be said that pushed spinoffs profit more from the knowledge and experience acquired via firm genealogy than that of pulled spinoffs, who in-turn have more to lose by means of knowledge spillovers. This echoes the findings of Rigby and Brown (2015) in that firms with less developed capabilities may have more to gain from being in an industrial cluster.

Raz and Gloor (2007), despite finding evidence that firms benefit from being centrally-placed in social networks, did not find evidence of returns to location after

controlling for network effects. The next step is to thus investigate the effect of parameters on the survival of spinoff firms. Formally, we estimate the hazard rate for the j^{th} subject, $h(t | \mathbf{x}_j)$, as:

$$h(t | \mathbf{x}_j) = h_0(t) \exp(\mathbf{D}\beta_D + \mathbf{c}_j\beta_c + \mathbf{Z}\beta_Z) \tag{2}$$

$$h(t | \mathbf{x}_j) = h_0(t) \exp(\mathbf{c}_j\mathbf{D}\beta_{cD} + \mathbf{Z}\beta_Z) \tag{3}$$

where, in Eq. 2, $h_0(t)$ is the baseline hazard, \mathbf{D} represents the geographic distance dummy variables, \mathbf{c}_j is the closeness centrality measure and \mathbf{Z} is the vector of control variables, with β_D , β_c and β_Z their corresponding coefficients to be estimated. Equation 3 dispenses with the distance dummies and centrality variables and specifies them together as an interaction term, $\mathbf{c}_j\mathbf{D}$. As a robustness check we also use a parametric hazards model assuming a Gompertz distribution, which has a baseline hazard $h_0(t)$ of $\exp(\gamma t) \exp(\beta_0)$, where γ , an ancillary parameter, controls the shape of the baseline hazard function. The choice of the Gompertz distribution is based on that of previous research (De Vaan et al. 2013; Klepper 2002; Buenstorf and Klepper 2009) as well as the nonparametric results in this paper, that indicate that the hazard rate increases monotonically with duration. Results for multivariate analysis per the specifications in Eqs. 2 and 3 may be seen in Table 7, and the corresponding results for the parametric model robustness check may be found in Table 10 in the appendix, the results of which are very similar. Link tests are insignificant, inferring that the models are not misspecified.

Referring to the first column of results, the distance dummies show a reduced hazard of exit with respect to the reference category D_i^{1-5km} , mirroring the results of the bivariate regressions in Table 6. After controlling for other explanatory variables, therefore, we can see that after introducing geographic precision into the full model, the effects of being in a cluster depend on specific location. After including closeness centrality in the full specification, results slightly differ from the bivariate analysis with a small increase in magnitude. Furthermore, there is a statistically insignificant relationship with respect to closeness centrality when including only pulled firms in the specification listed in Eq. 2, as seen in the third column. One might say that there is little difference *between* pulled spinoffs in terms of historical firm relationships, and hence the impact of such relationships on firm survival. With reference to the bivariate results in Table 6, in which we see a statistical significant variable (but a coefficient comparable in magnitude), an increase in the standard error in the multivariate case indicates that other variables have an enduring role in explaining firm survival among pulled firms. Geographic distance and the included control variables thus have a greater influence when comparing firms in the group of pulled spinoff firms. When considering pushed spinoffs (fourth column of results), on the other hand, closeness centrality has a much greater influence when comparing survival prospects. This coefficient, which is statistically significant, negative and large in magnitude, serves as a prominent explanation for the exit rates among pushed spinoffs. Insignificant coefficients for the distance variables indicates that there is a notable difference among pulled and pushed spinoffs in terms of the role of different knowledge channels (Table 7).

Table 6 Estimations from the Cox proportional hazards model, bivariate analysis

All firms				
$D_i^{0-1 km}$	-0.3210**			
	(0.034)			
$D_i^{1-5 km}$		0.1427***		
		(0.006)		
$D_i^{>5 km}$			-0.0957*	
			(0.062)	
C_i				-0.5953**
				(0.028)
Pulled firms				
$D_i^{0-1 km}$	-0.3099**			
	(0.169)			
$D_i^{1-5 km}$		0.1367***		
		(0.058)		
$D_i^{>5 km}$			-0.0907	
			(0.058)	
C_i				-0.4830**
				(0.286)
Pushed firms				
$D_i^{0-1 km}$	-0.3632			
	(0.338)			
$D_i^{1-5 km}$		0.1653		
		(0.112)		
$D_i^{>5 km}$			-0.1148	
			(0.111)	
C_i				-2.0725***
				(1.008)

Standard errors in parentheses below coefficient estimates

$N = 3050$ (all firms), 2414 (pulled firms), 636 (pushed firms)

* $p < 0.1$, ** $p < 0.5$, *** $p < 0.01$

Regarding the control variables, larger founding sizes decrease hazard rates in all three samples. The size of the parent firm at the time of spinoff, while statistically significant, has a small yet contributory effect on hazard rates, mirroring the ‘small firm effect’ found by Elfenbein et al. (2010), who found that smaller firms on average tend to spawn more successful spinoffs. This may highlight a potential source of bias considering the high number of self-employed individuals included in the sample. Highly educated employees contribute to a decreased hazard rate, except in the case of the sample that includes only pushed spinoffs, where there is no statistically significant effect, which reflects the premise that such firms are established out of necessity and not out of any Schumpeterian incentives. For the pooled samples, as well as that of pulled spinoffs, a greater percentage of male employees is associated with decreased hazard rates, as found by Andersson and Klepper (2013). Little can be gauged from

Table 7 Estimations from the Cox proportional hazards model, multivariate analysis

	All		Pulled	Pushed
D_i^{0-1km}	-0.3528** (0.155)		-0.3502** (0.173)	-0.3319 (0.347)
$D_i^{>5km}$	-0.1766*** (0.054)		-0.1719*** (0.061)	-0.1819 (0.118)
C_i	-0.8044** (0.338)		-0.5342 (0.368)	-2.5739** (1.087)
$C_i \times D_i^{0-1km}$		-2.8056** (1.349)		
$C_i \times D_i^{1-5km}$		0.0817 (0.466)		
$C_i \times D_i^{>5km}$		-1.2293*** (0.388)		
In Founding Size _{<i>i</i>}	-0.2182*** (0.042)	-0.1954*** (0.041)	-0.1950*** (0.047)	-0.2715*** (0.095)
In Parent Size _{<i>i</i>}	0.0527*** (0.020)	0.0557*** (0.020)	0.0447* (0.023)	0.0742* (0.045)
Share Highly Educated _{<i>i</i>}	-0.1116* (0.058)	-0.1033* (0.057)	-0.1356** (0.064)	0.0369 (0.131)
Share Male _{<i>i</i>}	-0.1668** (0.083)	-0.1633** (0.083)	-0.1792* (0.092)	-0.1137 (0.194)
Period 1 _{<i>i</i>} (1991-1992)	0.0821 (0.152)	0.0608 (0.152)	0.2401 (0.171)	-0.3725 (0.336)
Period 3 _{<i>i</i>} (1996-2000)	0.1728* (0.097)	0.1610* (0.097)	0.1536 (0.110)	0.2684 (0.207)
Period 4 _{<i>i</i>} (2001-2005)	0.0655 (0.100)	0.0511 (0.100)	0.0732 (0.115)	0.0378 (0.208)
Period 5 _{<i>i</i>} (2006-2010)	-0.6510*** (0.121)	-0.6729*** (0.121)	-0.6588*** (0.136)	-0.5932** (0.269)
Sector Change _{<i>i</i>}	0.1119** (0.055)	0.1141** (0.055)	0.1680*** (0.062)	-0.0765 (0.129)
Pulled Spinoff _{<i>i</i>}	-0.0796 (0.071)	-0.0839 (0.071)		
<i>Link test</i>	-0.2129 (0.2026)	-0.2463 (0.2078)	-0.1533 (0.2278)	-0.4551 (0.3228)
<i>Number of observations</i>	3,050		2,414	636

Standard errors in parentheses below coefficient estimates

Variables D_i^{1-5km} and Period 2_{*i*} (1993-1995) act as reference categories

* $p < 0.1$, ** $p < 0.5$, *** $p < 0.01$

the time period dummies (with period 2 held as the reference group), with the exception of period 5 (2006-2010), where new firms established during that period leads to superior performance in terms of survival. This is true for all three samples, and

may reflect the fact that Sweden experienced an economic upturn during this time. Switching to a different industry than the parent firm, unsurprisingly, has an effect of increasing the hazard rate for pulled spinoff firms as well as the pooled sample. For the pushed spinoff samples, this relationship does not hold, however. One may envisage that this is again due to the reasons for firm establishment. What may increase the survival prospects among pulled spinoffs do not necessarily hold true for pushed spinoffs, as the incentives of the entrepreneur are different. In the pooled sample, being a pulled spinoff has no statistically significant effect on firm survival.

The second column displays the results reflecting the specification in Eq. 3, which includes both pulled and pushed firms, and reveals how closeness centrality affect hazard of exit in each separate geographic radii. The coefficients for control variables in this regression correspond to the pooled sample specified in Eq. 2, both in magnitude and statistical significance. Spinoff firms located within 1 km of the cluster core display strong returns to genealogical linkages and realized knowledge channels, with a negative coefficient high in magnitude. Spinoffs located within 1 to 5 km from the cluster core on the other hand exhibit no statistically significant returns associated with the implied transmission of industry knowledge associated with established networks. Thus firms located within the intermediate threshold, in this sample and specification, apparently *do* enjoy a net benefit from passive network effects. One can speculate that the added competition burden at this radii threshold, as gauged by the results in Table 6, diminish any returns that a increased closeness centrality may bring. Conforming to this pattern, firms located outside the 5 km ring benefit from increased network efficiency and position, but this magnitude is not as high as that for firms located in the innermost core. This reflects the findings of Stuart and Sorenson (2003), who found that leveraging existing social ties requires geographic proximity to those ties. Therefore, one may conclude that the mechanism of knowledge transfer in industrial clusters is based on a combination of both realized and unrealized information channels.

5 Conclusion

Prior studies have found that after controlling for the pre-entry experience of spinoff firms, the absence of cluster effects becomes clear. Incumbent firms serve as ‘training grounds’ for future firms that establish themselves in the same region (Buenstorf and Klepper 2009). Parent-spinoff linkages, acting as a conduit of industry-specific routines and knowledge, serve to describe the difference in survival prospects of new firms in industrial clusters. This paper adds to the existing research by identifying and measuring such a network, and to separate these passive and realized knowledge flows from those that derive from co-location alone. This passive network forms part of the mechanism of localization economies. Spinoffs gain a performance premium through the occupational background of the entrepreneur. Surviving firms then provide a source of potential knowledge for other firms in the same geographic area. The process perpetuates with each successive generation of spinoffs, and a feedback of realized and unrealized knowledge channels persists. Using a Cox proportional hazards model, it was found that firms with efficient realized knowledge channels,

derived via stronger multi-generational linkages, had a better chance of survival over the given study period. This represents the first conclusion of this paper. The survival prospects of new spinoff firms are heavily influenced by inherited organizational routines and knowledge, irrespective of the underlying motivations of the entrepreneur.

A second conclusion concerns the effect of geographic distance from the cluster core, and hence unrealized knowledge channels, from the cluster on firm survival. Previous studies, for the most part, overlooked any potential nonlinear relationships when considering a firm's geographic position in relation to an industrial cluster. The results point to an inverse U-shaped relationship. The results suggest the following patterns. At very short distances from the cluster's centroid, new firms gain an advantage. At intermediate distances, which, in this case, includes the densest portion of the cluster, any gains from co-location are lost, potentially due to competition effects. This negative externality dissipates with greater distances. The performance of new firms in industrial clusters are thus more complex and not a simple matter of comparing firm performance in and outside such clusters. The returns could be explained by the fact that returns from co-location may be exploited but come with costs that are associated with geography as well. The gains from locating within a cluster thus depend not only on specific location but on unique firm attributes and the firm's ability to absorb such costs. We may also extend this interpretation by addressing the interaction between realized and unrealized knowledge flows. Although we are able to separate the effects of these two information channels, geographic distance may serve as an enabler for firms that wish to exploit their existing inter-firm relationships.

A third conclusion pertains to the effects of firm background when comparing firms with the same entrepreneurial incentives and motivations. Heredity of knowledge and routines, for example, does not necessarily give pulled spinoffs an advantage over other pulled spinoffs. Distance from the cluster's core, however, does. When comparing pushed spinoffs, distance has little discernible effect, but heredity gives a considerable advantage over other pushed firms. Thus, unrealized knowledge channels give pulled firms an advantage among their fellow pulled peers; while among pushed spinoffs, realized knowledge channels take the prominent role. There is thus no one-size-fits-all model that can broaden our understanding, and returns to knowledge flows are not always discernible. Different entrepreneurial incentives and motivations give rise to different sets of variables that affect the survival prospects of the firm.

Several limitations in the paper could be explored with future research. Firstly, the nonlinear, inverse U-shaped returns to geographic proximity could be further scrutinized by taking into account localized costs such as rents and wages, as well as initial firm endowments such as value added. Doing so would give us further insight of the trade-off of firm selection and the returns from location. Secondly, given the current passive network, introducing a node weighting for firm size would depict a network that takes into account linkages to firm quality. Including alternative networks, for example an *active* network such as those that depict collaboration networks, could shed further light on the role of geographic proximity, i.e. if such active networks result from location. Finally, the analysis could be extended to include variables that are time-varying, as well as including the possibility of firm relocation.

Given the availability of appropriate data, it would not be unreasonable to expand this analysis to different industries and different regions. The factors that affect one industry (or region) may be markedly different to those that affect firms in the

Stockholm ICT sector in this study. Indeed, the effects of shared inputs, another factor of localization economies, may serve as an additional explanator, especially in more capital-intensive industries. However, we can conclude that this paper has shed some light on the separate roles of parent-spinoff networks and geographic proximity on firm survival in industrial clusters.

Funding This research has been supported by the project R-ARE: Entrepreneurship, innovation and development in regions of varying density; which in turn is funded by the Swedish Research Council [grant number FORMAS 2011-80].

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

Table 8 NACE Classification for the ICT Sector

NACE (2002)	NACE (1992)	Activity
72100	72100	Hardware consultancy
72210	72202	Software consultancy and supply
72220	72201/ 74841	Data processing
72300	72300	Other software consultancy and supply
72400	72400	Database activities
72500	72500	Maintenance and repair of office, accounting and computing machinery
72600	72600	Other computer related activities

Table 9 Network formation rules

Type of new firm	Link inheritance	Additional rules
Spinoffs (pulled and pushed)	Yes	No
Mergers	Yes	The nodes of merged establishments become one
Splits	Yes	Upon a split, a link is formed between each of the splits
New establishments (common firm)	No	New links with other establishments with a common firm
Other firms (<i>de novo</i> , etc.)	No	No links upon foundation

Note: classification of new firms may overlap, and therefore may follow multiple rules

Table 10 Estimations from the parametric (Gompertz) hazards model, multivariate analysis

	All		Pulled	Pushed
$D_i^{0-1 km}$	-0.3811** (0.155)		-0.3802** (0.174)	-0.3503 (0.347)
$D_i^{>5 km}$	-0.1934*** (0.054)		-0.1872*** (0.061)	-0.2013* (0.118)
C_i	-0.8271** (0.338)		-0.5509 (0.368)	-2.6638** (1.088)
$C_i \times D_i^{0-1 km}$		-2.9125** (1.349)		
$C_i \times D_i^{1-5 km}$		0.1049 (0.465)		
$C_i \times D_i^{>5 km}$		-1.2789*** (0.387)		
In Founding Size _{<i>i</i>}	-0.2243*** (0.042)	-0.1992*** (0.041)	-0.1997*** (0.047)	-0.2817*** (0.095)
In Parent Size _{<i>i</i>}	0.0545*** (0.020)	0.0580*** (0.020)	0.0462** (0.023)	0.0780* (0.045)
Share Highly Educated _{<i>i</i>}	-0.1228** (0.058)	-0.1132** (0.057)	-0.1465** (0.064)	0.0307 (0.131)
Share Male _{<i>i</i>}	-0.1824** (0.083)	-0.1781** (0.083)	-0.1925** (0.092)	-0.1309 (0.193)
Period 1 _{<i>i</i>} (1991-1992)	0.0548 (0.152)	0.0317 (0.152)	0.2441 (0.171)	-0.4898 (0.337)
Period 3 _{<i>i</i>} (1996-2000)	0.2271** (0.097)	0.2126** (0.097)	0.2064* (0.110)	0.3192 (0.206)
Period 4 _{<i>i</i>} (2001-2005)	0.1712* (0.101)	0.1530 (0.100)	0.1784 (0.115)	0.1382 (0.209)
Period 5 _{<i>i</i>} (2006-2010)	-0.4338*** (0.122)	-0.4592*** (0.122)	-0.4368*** (0.137)	-0.3844 (0.270)
Sector Change _{<i>i</i>}	0.1204** (0.055)	0.1231** (0.055)	0.1834*** (0.062)	-0.0926 (0.130)
Pulled Spinoff _{<i>i</i>}	-0.0797 (0.071)	-0.0845 (0.071)		
Constant	-1.6660*** (0.140)	-1.7929*** (0.135)	-1.7538*** (0.159)	-1.6833*** (0.290)
γ	-0.0589*** (0.009)	-0.0599*** (0.009)	-0.0573*** (0.011)	-0.0535** (0.021)
Number of observations	3,050		2,414	636

Standard errors in parentheses below coefficient estimates

Variables $D_i^{1-5 km}$ and Period 2_{*i*} (1993–1995) act as reference categories. Gamma is an ancillary parameter estimated by the data, with a negative value indicating that hazard function decreases with time

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

References

- Agarwal R (1998) Small firm survival and technological activity. *Small Bus Econ* 11(3):215–224
- Agarwal R, Echambi R, Fanco A, Sarkar M (2004) Knowledge transfer through inheritance: spinout generation, development, and survival. *Acad Manage J* 47(4):501–522
- Agrawal A, Cockburn I, McHale J (2006) Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. *J Econ Geogr* 6(5):571–591
- Alc cer J, Chung W (2014) Location strategies for agglomeration economies. *Strateg Manage J* 35(12):1749–1761
- Alchian A (1950) Uncertainty, evolution, and economic theory. *J Polit Econ* 58(3):211–221
- Andersson M, Klepper S (2013) Characteristics and performance of new firms and spinoffs in Sweden. *Ind Corp Change* 22(1):245–280
- Andersson M, Xiao J (2016) Acquisitions of start-ups by incumbent businesses: a market selection process of “high-quality” entrants? *Res Policy* 45(1):272–290
- Andersson M, Klaesson J, Larsson J (2014) The sources of the urban wage premium by worker skills: spatial sorting or agglomeration economies? *Pap Reg Sci* 93(4):727–747
- Andersson M, Klaesson J, Larsson J (2016) How local are spatial density externalities? Neighbourhood effects in agglomeration economies. *Reg Stud* 50(6):1082–1095
- Anton J, Yao D (1995) Start-ups, spin-offs, and internal projects. *J Law Econ Organ* 11(2):362–378
- Arzaghi M, Henderson J (2008) Networking off madison avenue. *Rev Economic Stud* 75(4):1011–1038
- Audretsch D, Agarwal R (2001) Does entry size matter? The impact of life-cycle and technology on survival. *J Ind Econ* 49(1):21–43
- Audretsch D, Lehmann E (2005) Does the knowledge spillover theory of entrepreneurship hold for regions? *Res Policy* 34(8):1191–1202
- Audretsch D, Houweling P, Thurik R (2000) Firm survival in the Netherlands. *Rev Ind Organ* 16(1):1–11
- Bagley M (2018) Small worlds, inheritance networks and industrial clusters. *Ind Innov (Advance online publication)*:1–28
- Bavelas A (1950) Communication patterns in task-oriented groups. *J Acoust Soc Am* 22(6):725–730
- Borgatti S (2006) Identifying sets of key players in a social network. *Comput Math Organiz Theor* 12(1):21–34
- Boschma R (2015) Do spinoff dynamics or agglomeration externalities drive industry clustering? A reappraisal of Steven Klepper’s work. *Ind Corp Change* 24(4):859–873
- Boschma R, Ter Wal A (2007) Knowledge networks and innovative performance in an industrial district: the case of a footwear district in the South of Italy. *Ind Innov* 14(2):177–199
- Boschma RA, Wenting R (2007) The spatial evolution of the British automobile industry. Does location matter? *Ind Corp Change* 16(2):213–238
- Breschi S, Lissoni F (2009) Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *J Econ Geogr* 9(4):439–468
- Bruneel J, de Velde E, Clarysse B (2013) Impact of the type of corporate spin-off on growth. *Entrep Theory Pract* 37(4):943–959
- Buenstorf G, Guenther C (2011) No place like home? Relocation, capabilities, and firm survival in the German machine tool industry after World War II. *Ind Corp Change* 20(1):1–28
- Buenstorf G, Klepper S (2009) Heritage and agglomeration: the Akron tyre cluster revisited. *Econ J* 119(537):705–733
- Cabral L, Wang Z (2009) Spin-offs: theory and evidence from the early U.S. automobile industry. Federal Reserve Bank of Kansas City Working Papers No 08–15
- Canter U, Graf H (2007) Growth, development and structural change of innovator networks – the case of Jena. *Jena Economic Research Papers* 2007–090
- Canter U, Wolf T (2016) On regional innovator networks as hubs for innovative ventures. *Jena Economic Research Papers* 2016–006
- Canter U, Wolf T (2018) Entrepreneurship and the industry life cycle. Springer, Cham, chap social structure and organizations, pp 175–204
- Carias C, Klepper S (2010) Entrepreneurship, the initial labor force, and the location of new firms. In: International schumpeter society conference, Aalborg
- Cassia L, Colombelli A, Paleari S (2009) Firms’ growth: Does the innovation system matter? *Struct Change Econ D* 20(3):211–220

- Cefis E, Marsili O (2015) A matter of life and death: innovation and firm survival. *Ind Corp Change* 14(6):1–26
- Chesbrough H (2003) *Open innovation: the new imperative for creating and profiting from technology*. Harvard Business School Press, Boston
- Combes PP, Duranton G, Gobillon L, Puga D, Roux S (2012) The productivity advantages of large cities: distinguishing agglomeration from firm selection. *Econometrica* 80(6):2543–2594
- Cox D (1972) Regression models and life-tables. *J Roy Stat Soc B Met* 34(2):187–220
- Da Silva D, McComb R (2012) Geographic concentration and high tech firm survival. *Reg Sci Urban Econ* 42(4):691–701
- Dahl M, Østergaard C, Dalum B (2010) *Handbook of evolutionary economic geography*, Edward Elgar, Cheltenham, chap emergence of regional clusters: the role of spinoffs in the early growth process, pp 205–220
- Davis GF, Yoo M, Baker W (2003) The small world of the American corporate elite, 1982–2001. *Strategic Organ* 1(3):301–326
- De Vaan M, Boschma R, Frenken K (2013) Clustering and firm performance in project-based industries: the case of the global video game industry, 1972–2007. *J Econ Geogr* 13(6):965–991
- Duranton G, Puga D (2001) Nursery cities: urban diversity, process innovation, and the life cycle of products. *Am Econ Rev* 91(5):1454–1477
- Durlauf S (2004) *Handbook of Regional and Urban Economics*, vol 4. Elsevier, North-Holland, chap Neighborhood effects, pp 2173–2242
- Elfenbein D, Hamilton B, Zenger T (2010) The small firm effect and the entrepreneurial spawning of scientists and engineers. *Manage Sci* 56(4):649–681
- Eriksson T, Kuhn J (2006) Firm spin-offs in Denmark 1981–2000 - patterns of entry and exit. *Int J Ind Organ* 24(5):1021–1040
- Fackler D, Schnabel C (2016) Spinoffs in Germany: characteristics, survival, and the role of their parents. *Small Bus Econ* 46(1):93–114
- Falck O (2007) Survival chances of new businesses: Do regional conditions matter? *Appl Econ* 39(16):2039–2048
- Fleming L, Frenken K (2007) The evolution of inventor networks in the Silicon Valley and Boston regions. *Adv Complex Syst* 19(1):53–71
- Franco A, Filson D (2006) Knowledge diffusion through employee mobility. *RAND J Econ* 37(4):841–860
- Frenken K, Van Oort F, Verburg T (2007) Related variety, unrelated variety and regional economic growth. *Reg Stud* 41(5):685–697
- Frenken K, Cefis E, Stam E (2015) Industrial dynamics and clusters: a survey. *Reg Stud* 49(1):10–27
- Fruchterman T, Reingold E (1991) Graph drawing by force-directed placement. *Software: Practice and Experience* 22(11):1129–1164
- Furlan A (2016) Who lives longer? Startups vs spinoffs founded as proprietorships. *Int J Entrep Behav Res* 22(3):416–435
- Geroski P (1995) What do we know about entry? *Int J Ind Organ* 13(4):421–440
- Giuliani E, Bell M (2005) The micro-determinants of Meso-level learning and innovation: evidence from a Chilean wine cluster. *Res Policy* 34(1):47–68
- Glaeser E (2000) The future of urban research: non-market interactions. *Brookings - Wharton papers on urban affairs* 2000:101–138
- Glaeser EL, Maré D (2001) Cities and skills. *J Labor Econ* 19(2):316–342
- Gompers P, Lerner J, Scharfstein D (2005) Entrepreneurial spawning: public corporations and the genesis of new ventures, 1986 to 1999. *J Financ* 60(2):557–614
- Gulati R, Sytch M, Tatarynowicz A (2012) The rise and fall of small worlds: Exploring the dynamics of social structure. *Organ Sci* 23(2):449–471
- Hall B (1987) The relationship between firm size and firm growth in the US manufacturing sector. *J Ind Econ* 35(4):583–606
- Heebels B, Boschma R (2011) Performing in Dutch book publishing 1880–2008: the importance of entrepreneurial experience and the Amsterdam cluster. *J Econ Geogr* 11(6):1007–1029
- Hitt M, Ireland R, Camp S, Sexton D (2001) Guest editors' introduction to the special issue strategic entrepreneurship: entrepreneurial strategies for wealth creation. *Strateg Manage J* 22(6–7):479–491
- Hvide H (2009) The quality of entrepreneurs. *Economic J* 119(539):1010–1035
- Jacobs J (1969) *The economies of cities*. Random House, New York

- Kaplan E, Meier P (1958) Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 53(282):457–481
- Karlsson C, Mellander C, Paulson T (2004) A spatial ict clusters in Sweden - an empirical method to identify necessary conditions for existence. CESIS working paper series in economics and institutions of innovation
- Klepper S (2001) Employee startups in high tech industries. *Ind Corp Change* 10(3):639–674
- Klepper S (2002) The capabilities of new firms and the evolution of the US automobile industry. *Ind Corp Change* 11(4):645–666
- Klepper S (2007a) Disagreements, spinoffs and the evolution of Detroit as the capital of the U.S. automobile industry. *Manage Sci* 53(4):616–631
- Klepper S (2007b) The geography of organizational knowledge. Mimeo
- Klepper S (2009) Spinoffs: a review and synthesis. *Eur Manag Rev* 6(3):159–171
- Klepper S (2010) The origin and growth of industry clusters: the making of Silicon Valley and Detroit. *J Urban Econ* 67(1):15–32
- Klepper S, Sleeper S (2005) Entry by spinoffs. *Manage Sci* 51(8):1291–1306
- Larsson J (2017) Non-routine activities and the within-city geography of jobs. *Urban Stud* 54(8):1808–1833
- Larsson JP (2014) The neighborhood or the region? Reassessing the density-wage relationship using geocoded data. *Ann Regional Sci* 52(2):367–384
- Maine E, Shapiro D, Vining A (2010) The role of clustering in the growth of new technology-based firms. *Small Bus Econ* 34(2):127–146
- Marchiori M, Latora V (2000) Harmony in the small-world. *Physica A* 285(3–4):539–546
- Marshall A (1890) Principles of economics. Macmillan, London
- Mata J, Portugal P (1994) Life duration of new firms. *J Ind Econ* 42(3):227–245
- Melitz M (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71(6):1695–1725
- Moore G, Davis K (2001) Learning the Silicon Valley way. Stanford institute for economic policy research, paper 00–45
- Nelson R, Winter S (1982) An evolutionary theory of economic change. Harvard University Press, Cambridge
- Nyström K (2007) An industry disaggregated analysis of the determinant of regional entry and exit. *Ann Regional Sci* 41(4):877–896
- Parhankangas A, Arenius P (2003) From a corporate venture to an independent company: a base for a taxonomy for corporate spin-off firms. *Res Policy* 32(3):463–481
- Phillips D (2002) A genealogical approach to organizational life chances: the parent-progeny transfer among Silicon Valley law firms, 1946–1996. *Admin Sci Quart* 47(3):474–506
- Ponds R, Van Oort F, Frenken K (2010) Innovation, spillovers, and university-industry collaboration: an extended knowledge production function approach. *J Econ Geogr* 10(2):231–225
- Porter P (1990) The competitive advantage of nations. Macmillan, London
- Powell W, Koput K, Smith-Doerr L (1996) Interorganizational collaboration and the locus of innovation: networks and learning in biotechnology. *Admin Sci Quart* 41(1):116–145
- Raz O, Gloor P (2007) Size really matters - new insights for start-ups' survival. *Manage Sci* 53(2):169–177
- Renski H (2011) External economies of localization, urbanization and industrial diversity and new firm survival. *Pap Reg Sci* 90(3):473–502
- Rigby D, Brown W (2015) Who benefits from agglomeration? *Reg Stud* 49(1):28–43
- Rocha V, Carneiro A, Varum C (2018) Leaving employment to entrepreneurship: the value of co-worker mobility in pushed and pulled-driven start-ups. *J Manage Stud* 55(1):60–85
- Rosenthal S, Strange W (2001) The determinants of agglomeration. *J Urban Econ* 50(2):191–229
- Rosenthal S, Strange W (2005) The geography of entrepreneurship in the New York metropolitan area. Federal Reserve Bank of New York Economic Policy Review 11(2):29–53
- Rosenthal S, Strange W (2008) The attenuation of human capital spillovers. *J Urban Econ* 64(2):373–389
- Sapienza H, Parhankangas A, Autio E (2004) Knowledge relatedness and post-spin-off growth. *J Bus Venturing* 19(6):809–829
- Schumpeter J (1934) The theory of economic development. Harvard University Press, Cambridge
- Shane S, Stuart T (2002) Organizational endowments and the performance of university start-ups. *Manage Sci* 48(1):154–170

- Sørensen J, Phillips D (2011) Competence and commitment: employer size and entrepreneurial endurance. *Ind Corp Change* 20(5):1277–1304
- Sorenson O, Audia P (2000) The social structure of entrepreneurial activity: geographic concentration of footwear production in the United States, 1940–1989. *Am J Sociol* 106(2):424–462
- Stinchcombe A (1965) *Handbook of organizations*, vol 7. Rand McNally & Company, Chicago, chap Social structure and organizations, pp 142–193
- Stuart T, Sorenson O (2003) The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms. *Res Policy* 32(2):229–253
- Suire R, Vicente J (2009) Why do some places succeed when others decline? a social interaction model of cluster viability. *J Econ Geogr* 9(3):381–404
- Teece D, Pisano G, Shuen A (1997) Dynamic capabilities and strategic management. *Strateg Manage J* 18(7):509–533
- Ter Wal A (2013) Cluster emergence and network evolution: a longitudinal analysis of the inventor network in Sophia-Antipolis. *Reg Stud* 47(5):651–668
- Uzzi B, Spiro J (2005) Collaboration and creativity: the small world problem. *Amer J Sociol* 111(2):447–504
- Van Oort F, Burger M, Knoblen J, Raspe O (2012) Multilevel approaches and the firm-agglomeration ambiguity in economic growth studies. *J Econ Surv* 26(3):468–491
- Wennberg K, Lindqvist G (2010) The effects of clusters of the survival and performance of new firms. *Small Bus Econ* 34(3):221–241
- Wenting R (2008) Spinoff dynamics and the spatial formation of the fashion design industry, 1858–2005. *J Econ Geogr* 8(5):593–614
- Wright M, Birley S, Mosey S (2004) Entrepreneurship and university technology transfer. *J Technol Transfer* 29(3–4):235–246

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.