UNIVERSITA’ COMMERCIALE “LUIGI BOCCONI”

PhD SCHOOL

PhD program in Business Administration and Management

Cycle: 33

Disciplinary Field (code): SECS-P/08

Three essays on the role of knowledge as a driver of entry and industry dynamics

Advisor: Franco MALERBA

PhD Thesis by

Sung Hoon LEE

ID number: 3053520

Year 2022

DISSERTATION ABSTRACT

This dissertation addresses the question of how the knowledge characteristics of firms and industries affect firm entry. Firm entry is considered an important aspect of any analysis in industrial economics or management, due to its direct implications on market structure and competition. Knowledge, in particular, is found to be a prominent driver of firm entry, and is important in explaining the differences across entrants in various industries. Moreover, the identification of the mechanism behind the differences in entry patterns can greatly increase our understanding of how industries evolve in the long run. This requires, however, a thorough examination of the interplay between the knowledge of potential entrants and the knowledge in industries. All these themes are at the base of this dissertation, and are addressed across three chapters from different perspectives.

The first chapter, titled “The technological regimes and barriers to entry”, focuses on how industries’ knowledge characteristics, described by their technological regimes, affect the entry of firms, each of which may react differently. The results show that there are significant interaction effects between the different elements of the technological regime. In addition, certain types of entrants seem be affected by the technological regime in significantly different ways, especially diversifiers that are innovative and are from related industries.

The second chapter, titled “Where do firms come from? Knowledge relatedness and firm entry”, looks at how the individual potential entrants, which may be diversifiers, spinoffs, or spinouts, make their decision to enter an industry based on how related their knowledge is to the knowledge of the target industry. The results suggest that the relationship takes an inverted U shape for spinouts. For diversifiers and spinoffs, the relationship is U-shaped instead, suggesting that the two effects identified in the theoretical framework vary in strength depending on the types of entrants.

The third chapter, titled “Who enters and when? Knowledge relatedness, entrant

heterogeneity, and industry characteristics", expands this perspective to include how the industry-level moderators may alter the relationship between knowledge relatedness and entry for different types of entrants. The results of this paper show that industry characteristics matter in moderating the relationship between knowledge relatedness and entry, but some are more relevant than others. Modularity and innovators concentration, in particular, seem to significantly affect the relationship.

By looking at the question with these different perspectives, the studies in the dissertation show that knowledge characteristics of both firms and industries have major impacts on entry, and by extension on market structure.

TABLE OF CONTENTS

Three essays on the role of knowledge as a driver of entry and industry dynamics

Sung Hoon Lee

OVERVIEW

Firm entry is an important aspect of any analysis in industrial economics or management, due to its direct implications on market structure and competition (Audretsch, 1991; Malerba & Orsenigo, 1997; Helfat & Lieberman, 2002; Agarwal & Shah, 2014; Adams, Fontana, & Malerba, 2015). One factor found to be a particularly prominent driver of firm entry is the knowledge characteristics of firms and industries. This can be related to the evolutionary perspective, where knowledge is seen as one of the major sources of competitive advantage, and a key determinant of the expected post-entry survivability and performance of a potential entrant.

Knowledge is also an important factor in the explanation of the vast differences across entrants in the various industries. The analysis of these differences has important implications for both policy and firm strategy. Moreover, the identification of the mechanism behind the differences in entry patterns can greatly increase our understanding of how industries evolve in the long run. This requires, however, a thorough examination of the interplay between the knowledge of potential entrants and the knowledge in industries. All these themes are at the base of this dissertation.

This dissertation is organized into three chapters, each looking at the question of how knowledge affects firm entry from different perspectives. The first chapter focuses on how industries' knowledge characteristics, described by their technological regimes, affect the entry of firms, each of which may react differently. Then, the second chapter looks at how the individual potential entrants, which may be diversifiers, spinoffs, or spinouts, make their decision to enter an industry based on how related their knowledge is to the knowledge of the target industry. Finally, the third chapter expands this perspective to include how the industry-level moderators may alter the relationship between knowledge relatedness and entry for different types of entrants.

CHAPTER 1- THE TECHNOLOGICAL REGIME AND BARRIERS TO ENTRY

The first chapter of the dissertation uses the technological regime framework to analyze firm entry across multiple manufacturing industries. The technological regime, defined in terms of appropriability, cumulativeness, and opportunity conditions, has often been thought to affect firms, especially their decision to enter. However, many cross-industry analyses consider only a few types of technological regimes, providing a limited understanding of how the technological regime affects firm entry.

Two issues, in particular, would benefit from further examination. First, the interaction effects of the dimensions of technological regimes deserve more detailed analyses, given the theoretical and empirical evidence that such interaction effects may be relevant. In addition, the disentangling entrant characteristics when discussing the impact of the technological regime may prove illuminating. This study thus aims to address this gap, which can have major implications for both firm strategy and policy.

The results show that there are significant interaction effects between the different elements of the technological regime. In addition, certain types of entrants seem be affected by the technological regime in significantly different ways, especially diversifiers that are innovative and are from related industries.

CHAPTER 2- WHERE DO FIRMS COME FROM? KNOWLEDGE RELATEDNESS AND FIRM ENTRY

The second chapter of the dissertation focuses on the role of knowledge relatedness between potential entrants and the target industry on the potential entrant's decision to enter the target industry. The role of knowledge, in particular its relatedness to that of the target industry, has often been emphasized as one of the main drivers in firms' entry decisions. However, empirical analyses seem to show that knowledge relatedness may not always have a positive impact on entry, because the relationship is potentially nonlinear and may vary by entrant type.

This study aims to explore these issues by proposing a perspective that may shed light on the mechanisms that govern the relationship between knowledge relatedness and firm entry. Specifically, it identifies two effects of knowledge relatedness on firm entry, dubbed the knowledge effect and the differentiation effect, which work in opposite directions. It then empirically tests these mechanisms by using a dataset constructed from the Western Electronics Manufacturer Association (WEMA) directory, spanning a time period from 1960 to 1990.

The results suggest that the relationship takes an inverted U shape for spinouts. For diversifiers and spinoffs, the relationship is U-shaped instead, suggesting that the two effects identified in the theoretical framework vary in strength depending on the types of entrants.

CHAPTER 3- WHO ENTERS AND WHEN? KNOWLEDGE RELATEDNESS, ENTRANT HETEROGENEITY, AND INDUSTRY CHARACTERISTICS

The third chapter of the dissertation looks at how the relationship between the knowledge relatedness of potential entrants to the target industry and their decision to enter may be moderated by industry-level factors, based on past evidence that shows that the relationship between relatedness and the probability to enter often varies across industry and time. This study provides a theoretical framework to identify these moderators, focusing on modularity, rate of change, technological discontinuities, and concentration across different types of entrants. It then empirically tests this framework using a dataset of firms active in the US electronics industry between 1960 and 1990.

The results of this paper show that industry characteristics matter in moderating the relationship between knowledge relatedness and entry, but some are more relevant than others. Modularity and innovators concentration, in particular, seem to significantly affect the relationship.

CONCLUDING REMARKS

By looking at the question with these different perspectives, the studies in the dissertation show that knowledge characteristics of both firms and industries have major impacts on entry, and by extension on market structure. For the prospective entrants to an industry, this means that they will be required to examine what their knowledge profile, and the strengths and weakness that come with it, are if they want to successfully enter. For the incumbents in an industry, knowing the knowledge characteristics of their industry will allow them to predict what types of entrants they can expect to face, which in turn determines the type of competition they may be facing. Finally, in terms of policy, identifying how knowledge characteristics of firms and industries govern firm entry will guide policy makers in assessing the consequences of their decisions concerning the competitive environment, which will have a direct effect on social welfare as well.

REFERENCES

Adams, P., Fontana, R., & Malerba, F. 2015. User-industry spinouts: Downstream industry knowledge as a source of new firm entry and survival. ***Organization Science***. 27 (1): 18-35.

Agarwal, R., & Shah, S. 2014. Knowledge sources of entrepreneurship: Firm formation by academic, user and employee innovators. ***Research Policy***. 43: 1109-1133.

Audretsch, D. 1991. New-firm survival and the technological regime. ***The Review of Economics and Statistics***. 73 (3): 441-450.

Helfat, C., & Lieberman, M. 2002. The birth of capabilities: market entry and the importance of pre-history. ***Industrial and Corporate Change***. 11 (4): 725-760.

Malerba, F., & Orsenigo, L. 1997, Technological regimes and sectoral patterns of innovative activities. ***Industrial and Corporate Change***. 6 (1): 83-117.

ACKNOWLEDGEMENTS

I would like to thank, first and foremost, my advisor Franco Malerba whose insights and feedback were invaluable for writing this dissertation. I could not have completed my PhD program without his support, and it was a privilege to be Franco's student. I would also like to thank Roberto Fontana, whose theoretical rigor and expertise on empirical methods were extremely helpful for me and my dissertation.

My family's support for my decision to pursue this PhD was also crucial, and for that I am very thankful. It is thanks to their support that I could persevere, and without their support this dissertation might never have been written.

I would also like to thank the faculty of Bocconi University, especially Alfonso, Arnaldo, Charlie, Dovev, Nel, and Sandeep for their feedback on the chapters during and after my presentation of the project. They have been extremely helpful in developing the chapters into works that can provide real contribution to the field of management.

I have also received valuable feedback on multiple chapters from outside Bocconi as well. I would especially like to mention Keun Lee from Seoul National University, who was previously my advisor when I was still a Masters student and continued to support me during my PhD program.

I would also like to thank my friends and colleagues in the Department of Business Administration and Management of Bocconi University. Can, as my officemate, has provided me with the support I needed to go through the PhD program, and I hope I was able to do the same for him. Hyoungwon helped me to think more critically about my work and develop it accordingly. Jens has been helpful in giving me valuable insight in formulating my theory and encouraging me and I hope I was helpful for him as well. Navya has been instrumental in creating an environment where we could all give each other valuable feedback for our work, for which I am grateful. Danilo and Margherita have also been very supportive in my PhD journey.

Finally, I would also like to thank the Strategy Research Foundation of the Strategic Management Society for their financial support, without which a significant part of this dissertation would not have been possible.

The technological regime and barriers to entry

ABSTRACT

The impact of the technological regime, defined in terms of appropriability, cumulativeness, and opportunity conditions, on firms has been examined in multiple studies. Within the analysis on technological regimes, firm entry in particular has been of consistent interest due to its implications on competition and on how firms may decide to enter an industry. However, where cross-industry differences are concerned, many analyses consider only a few types of technological regimes, providing only a limited understanding of how the technological regime affects firm entry. Two issues, in particular, have not been analyzed as thoroughly as they should have been. First, the interaction effects of the various dimensions of technological regimes have not necessarily been analyzed in detail in the literature, despite the theoretical and empirical evidence that such interaction effects may be relevant. In addition, the characteristics of the entrants has not been disentangled when discussing the impact of the technological regime. This study thus aims to address this gap, which can have major implications for both firm strategy and policy.

Keywords:

Empirical: Quantitative; Competitive Dynamics; Evolutionary Perspectives; Barriers to Entry; Technological Regime; Entrant Heterogeneity

The technological regime and barriers to entry

INTRODUCTION

Firm entry is considered an important aspect of firm behavior due to its lasting consequences on industry structure and competition (Agarwal & Gort, 1996; Klepper, 1996; Agarwal & Audretsch, 2001; Garavaglia, 2016). In analyzing firm entry, the technological environment of the industry is often emphasized as an important factor. In particular, the evolutionary perspective emphasizes the major role of the technological regime of the industry in affecting firm entry (Nelson & Winter, 1982; Malerba & Orsenigo, 1997; Klepper & Simons, 2000; Agarwal & Audretsch, 2001; Peneder, 2010; Agarwal & Shah, 2014; Capone, Malerba, & Orsenigo, 2019). However, while plenty of studies show that different technological regimes, characterized by different combinations of their defining elements, generate different rates of entry, there is still some disagreement on the empirical front as to how or why these differences occurs and are characterized (Breschi, Malerba, & Orsenigo, 2000; Lin & Huang, 2008; Peneder, 2010; Cockburn & MacGarvie, 2011; Agarwal & Shah, 2014). Moreover, the relationship is often not analyzed in depth, because the focus is on related aspects such as survival or the aggregate rate of entry. This paper aims to provide more direct evidence of the relationship between the technological regime and firm entry, focusing on possible sources of nonlinearity.

The first source of nonlinearity is the possible existence of interaction effects between the various elements of the technological regime. This stems from two observations. First, the available empirical evidence on how the technological regime affects entry tends to be somewhat contradictory. Some studies suggest, for example, that higher appropriability in the industry is associated with lower rates of entry (Breschi et al., 2000; Cockburn & MacGarvie, 2011), while some show that the opposite may be the case (Peneder, 2010). Second, multiple streams of literature suggest not only that interaction effects exist, but also that they would be significant enough to dramatically change the effect of the individual elements of the technological regime. One example of such stream of literature is the one on patent policy and innovation (Green & Scotchmer, 1995; Bessen & Maskin, 2009; Norman, Pepall, & Richards, 2016; Parra, 2019), where it is implied that the effect of appropriability can change depending on the level of cumulativeness, even to the point that the direction of the effect may turn opposite. Such interaction effects are also alluded to in earlier works on the technological regime as well, such as Malerba & Orsenigo (1997), but they are indicated mostly as possibilities for future research.

The second source of nonlinearity is the heterogeneity among entrants. While heterogeneity among entrants is increasingly analyzed in the literature on firm entry (Helfat & Lieberman, 2002; Bayus & Agarwal, 2007; Agarwal & Shah, 2014), it is not considered in the technological regime literature, except when changes in the technological regime are used as external shocks (Chen, Williams, & Agarwal, 2012). Given that one of the principal differences between entrants is often framed in terms of differences in their knowledge, it is plausible to assume that an industry with a specific technological regime may appear more attractive to one type of entrant than to other types.

The importance of both these issues arises from their implications for policy and firm strategy. Policy-wise, the significance is related to how changes in the technological environment may affect competition in the industry. This in turn may even change the technological trajectory of the industry, because changes in competition could lead to a selection of firms within the industry, which may influence how technology changes in the future. In terms of firm strategy, the importance is related to the entrants, in that the analysis may highlight what kind of technological barriers they can expect to find when deciding whether to enter an industry. In addition, the analysis may also show how entrants should allocate their resources once they decide to enter, since different types of entrants can face different costs. Finally, the analysis also has useful implications for incumbents, since it allows them to predict the level and type of competition they may face in a specific technological environment.

This paper thus aims to provide an analysis of how the technological regime may affect firm entry by emphasizing the interaction effects between the elements of the technological regime and the role of entrant heterogeneity. The paper is organized in the following way. First, a literature review to highlight these issues and the possible reasons why these issues have not been analyzed in depth is provided. Then, the theoretical framework on which the empirical analysis will be based follows. Then, the procedures and the main results are presented and discussed. Finally, the potential implications of this study are examined.

LITERATURE REVIEW

Barriers to entry and the technological regime

The concept of barriers to entry has been a classic component of any analysis involving competition in industries or markets. The attention on the subject is based on many reasons, most of which revolve around its impact on the competitive environment of the industry. The basic idea is that because higher barriers to entry lead to lower rate of entry overall, this can have a major impact in the number of firms competing in the industry and their profits. In addition, their impact on firms' decision to enter and the subsequent performance of entrants has also been analyzed fairly extensively (Acs & Audretsch, 1989; Geroski, 1995; Agarwal & Gort, 1996; Agarwal & Audretsch, 2001; Cockburn & MacGarvie, 2011; Garavaglia, 2016).

The importance attributed to this concept also means that various factors have been proposed as contributing to the deterrence of entry. One of the core mechanics through which barriers to entry play a role is based on their ability to raise the costs of entry, as is evident in the definition proposed by the literature as listed in McAfee, Mialon, & Williams (2004). Thus, many of these factors are explained as those that confer an advantage or disadvantage specifically to entrants. Some of the common factors that are consistently cited based on this logic are scale economies, capital intensity, minimum efficient scale, industry size, etc. as noted by Geroski (1995).

In particular, the technological characteristics of the industry have been consistently emphasized as a major contributor to entry deterrence. This can be traced at least as far back as Schumpeter's discussions on whether the small entrepreneurial firms or large incumbents are the main innovators, later dubbed each as Schumpeter Mark I (Schumpeter, 1934) and Schumpeter Mark II (Schumpeter, 1942). The idea, essentially, is that the characteristics of the technology in the industry determines whether incumbents or entrants can actually thrive. This is evident in studies such as Nelson & Winter (1982), Klepper & Simons (2000), or Agarwal & Audretsch (2001), where technological change is linked to how the industry evolves.

However, the role of technology has not always been discussed in a systematic manner, particularly with regards to entry. Klepper & Simons (2000), for example, simply mentions the role of technical change, while Agarwal & Audretsch (2001) only looks at technological intensity, which essentially measures the rate at which innovations occur. While these are useful in highlighting the significance of the impact technology has in general, they do not quite explain how or why that impact can be important to an industry. In particular, they do not really show how the qualitative aspects of technology in one industry can be different from others, which in turn led to the unfortunate oversight on how changes that are seemingly similar might lead to different outcomes. In other words, they simply focus on the impact of specific technology-related phenomena rather than the actual mechanics through which technology can affect firms.

This is not to say that there have not been studies that analyze the role of technology beyond its quantitative aspects. The technological regime, a concept suggested by Nelson & Winter (1982), is one way with which this issue has been addressed. This can be thought of as a framework with which such technological characteristics can be organized.

In many cases, studies using the technological regime to study entry seem to rely heavily the distinction between routinized and entrepreneurial regimes, or some other taxonomy such as that suggested by Pavitt (1984). This is often true with older studies such as Audretsch (1997), but even the more recent studies such as Chen et al. (2012), show a similar tendency as well.

There have also been attempts to identify the key elements of the technological regime that has a significant impact on firms and industries, which can be classified into appropriability, cumulativeness, opportunity, and knowledge base according to Malerba & Orsenigo (1997). Appropriability refers to the extent to which firms can capture the value of their own innovations. Cumulativeness characterizes how important previous innovations are for successful innovation in the future. Opportunity condition determines the ease with which a firm can innovate. The last, knowledge base, usually refers to tacitness, basicness, complexity, etc.

The usefulness of this framework has been shown in several studies, such as, Breschi, Malerba, & Orsenigo (2000) or Park & Lee (2006), albeit with somewhat differing implications from each of the elements. Breschi et al. (2000) shows that appropriability and cumulativeness tend to deter entry while opportunity encourages it. On the other hand, Park & Lee (2006) is more ambiguous on the subject, partly because the study does not address entry directly, but it seems to suggest that at least for some types of firms the impact of each element may be very different.

It is worth noting that a sizable proportion of studies in the literature tend to treat these elements, at least appropriability and cumulativeness, as something strategic in nature (Ceccagnoli, 2009; Ahuja, Lampert, & Novelli, 2013; Huang, Ceccagnoli, Forman, & Wu, 2013; Gans & Stern, 2017). While there is no denying that this is the case, these also tend to refer to the firms' choice of mechanisms in order to enforce appropriability. On the other hand, in the context of this study, appropriability is merely a description of the industries' technological environment. The manner in which a certain level of appropriability or other technological characteristics manifest in the industry may vary, but as far as the entrants are concerned, it is not clear that this affects their decisions to enter an industry.

Technological regime and possible interaction effects

What is notably absent in both approaches that use the technological regime is that the possibility of interaction effects between elements of the technological regime is often ignored. This is despite the fact that a fair number of analyses acknowledge their existence and their potential importance, albeit without offering a very detailed analysis of this point. Examples of this include studies such as Malerba & Orsenigo (1997), which discusses the possibility of nonlinearity of the interplay between the different elements.

The theoretical analysis of patent policy and innovation arguably sheds more light on this topic (Mansfield, Schwarz, & Wagner, 1981; Green & Scotchmer, 1995; Bessen & Maskin, 2009; Norman et al., 2016; Parra, 2019). The motivation for this stream of literature was originally centered around the issue of whether the use of patents was actually optimal in promoting innovation. While common sense tends to imply that patents would lead to more innovations, one of the major questions in the literature is if and when patents may actually have an adverse effect. Much of the discussion started out from a setting where two firms were competing to obtain patents for some set of innovations, which then led to conclusions on whether either the leader or follower would invest in innovation *ex ante*.

One of the central developments in the process was the concept of sequential innovations. This refers to a situation where innovations are achieved based on previous innovations. The idea was that while patent protection may be useful for providing incentives to innovate in certain situations, this may actually be detrimental for the overall rate of innovation if innovations were sequential because it made it difficult for other firms to innovate further. In other words, in exchange for letting innovators keep the value of their innovations for themselves, there was the side effect that others' ability to innovate in the first place was undermined if innovations were sequential in nature.

One may note here that many of the constructs discussed in this literature, as described so far, can arguably be translated to elements of the technological regime. The use of patent policy can be thought to be analogous to an increase in appropriability. This is especially true because while the context is the changes in patent policy, these are generally conceptualized simply as changes in the costs of imitating existing innovations. In other words, the model in these studies could be applied to the context where secrecy is the main appropriability mechanism used with very few modifications. On the other hand, the comparison of cases where innovation is sequential with the static case can be interpreted as changes in cumulativeness. This part is immediately clear from the definition of sequential innovation, which according to Bessen & Maskin (2009) refers to a situation where "each successive innovation builds on the preceding one".

Some may argue, of course, that the comparison is of somewhat limited use, since the literature takes a very narrow view of how any level of appropriability is achieved. This is certainly a valid concern, in that it is true that this literature does not even attempt to address other ways in which firms might capture the value of their innovations, such as secrecy. Rather paradoxically, however, this may actually be why studies on patent policy and innovation have gone in a direction that allowed the comparison to be more useful. Since they only focus on patents, the theoretical representation of how they affect firms is simply modelled as increase in imitation costs, i.e., costs for other firms to learn or 'invent around' the existing innovations, without going into the details of how patent protection works[1]. One may note that this is precisely the definition of appropriability as given in studies such as Malerba & Orsenigo (1997), especially the one based on protecting from imitation. The same goes for the cumulativeness aspect, which is not defined in a way specific to patents in the first place. In other words, while the discussion focuses solely on patents, its logic is constructed in such a way that it could be translated into the more general analysis of the technological regime without too much modification.

Entrant type

Another important topic in the discussion of entry is the heterogeneity between entrants. As emphasized in studies such as Helfat & Lieberman (2002), Bayus & Agarwal (2007), or Agarwal & Shah (2014), the eventual performance of the entrants can depend heavily on

[1] The only detail contested in this literature is the extent to which patents are modelled as actually protecting inventions or innovations. While some studies make the assumption that patents provide perfect protection, others introduce the possibility of learning and 'inventing around', essentially suggesting that patents merely increase the costs of innovation by some finite amount for latecomers. Bessen & Maskin (2009) bases its analysis on the former setting, assuming that if one firm manages to innovate when patents are available, the other is unable to even attempt achieving the next possible innovation. On the other hand, Green & Scotchmer (1995) is closer to the latter in its assumptions, where variations are considered in how easily new patents are accepted as non-infringing.

what resources and knowledge they hold by the time of entry. This is particularly emphasized when comparing diversifying entrants with de novo entrants, where one of the main differences is the access, or lack thereof, to a relatively large stock of related knowledge or assets.

Of the diversifying entrants, a particular group received special attention. This consisted of diversifiers that came from not just any industry, but a related industry, and they were often thought to perform better than other firms (Markides & Williamson, 1994; Helfat & Lieberman, 2002). In fact, studies on related diversification form their own stream of literature, generally focusing on the question of whether firms should aim for related diversification as opposed to the unrelated kind in order to achieve optimal levels of performance. Some examples include Chang (1996), where the advantages of related diversification are highlighted, also citing other works such as Singh & Montgomery (1987). The general idea is that related diversifiers have an advantage compared to other types of firms due to their possession of knowledge that is applicable in the industry or sector.

The contrast between innovative and non-innovative (which is usually referred to as 'imitative' in the literature) entry is also considered important in the literature on entrant types[2], partly because this is a good indicator of the entrant's technological capabilities, which is often thought to be a major contributor to performance. In addition, much of the patent policy and innovation literature is centered around the question of patent's role in limiting imitation by latecomers and how that affects overall rate of innovation (Bessen &

[2] Here, the definitions are based on Acemoglu & Cao (2015), where growth from innovations by incumbents and entrants are analyzed. Innovative entry essentially refers to a situation where the entrant enters the market with an innovation, whereas non-innovative entry refers to a case where the entrants simply imitate the existing innovations.

Maskin, 2009; Norman et al., 2016). It is also a question of strategy, as shown by Ethiraj & Zhu (2008), in that this may be linked to first-mover advantage.

In many studies, the technological regime is not necessarily a point of focus in this context, with perhaps some exceptions like Chen et al. (2012). However, there are enough analyses to indicate that there may be some relationship. Agarwal & Audretsch (2001), for example, shows that there is some interaction between the size of entry and the technological environment, which may hint at a similar relationship between types of entrants and the technological regime.

THEORETICAL FRAMEWORK

Appropriability, cumulativeness, and opportunity

The technological regime can be thought of as being composed of four major components. The first is appropriability, which refers to the extent to which firms can capture the value of their own innovations (Malerba & Orsenigo, 1997). One particularly useful way to frame this is to link it to imitation costs (Cohen & Levin, 1989; Mansfield et al., 1981), which refers to the ease with which competitors can imitate and thus take away the value of the innovation for themselves. This essentially emphasizes appropriability as a measure of preventing imitation, which can also translate into an incentive to innovate. Cumulativeness is the second, which characterizes how important previous innovations are for the success of innovating in the future. The third is the opportunity condition, which determines the ease with which a firm can innovate. The last is the knowledge base, which usually refers to tacitness, codifiability, complexity, specificity, etc.

In this paper, the knowledge base is excluded from the analysis, as seems to be the case in much of the literature, which tends to focus on the first three[3]. This is because the knowledge base tends to be a rather vague and multifaceted concept. While this in itself may not be enough justification, this also leads to a more serious problem that measuring it becomes extremely difficult. Not helping matters is the issue that it is generally harder to distinguish from the other elements. For example, whether the knowledge is tacit or codified is often associated with how easy it is to learn, which coincides somewhat closely with the concept of appropriability.

Appropriability and entry. Higher level of appropriability can essentially be seen as higher imitation costs, which suggests that incumbent firms will tend to be able to benefit more from their innovations if they have any. Conversely, entrants will tend to find it more difficult to imitate or absorb the existing knowledge in the industry, which in turn would have a negative impact on entry. This is because firms will need to at least use the existing innovations with some proficiency in order to compete effectively, which an entrant will typically be unable to do without actually learning them in the first place. In addition, since technological capabilities, and the firm's success in innovation by extension, can be highly dependent on the absorption of external knowledge (Cohen & Levinthal, 1989), the increase in imitation costs can be detrimental for the entrants' long-term performance, which may also hurt entry.

On the other hand, because high appropriability means more of the benefits stay with the innovator, it could also motivate potential entrants if they believe they can innovate before or soon after entry. This has, interestingly, also been demonstrated in some studies such as Bessen & Maskin (2009), where the results imply that innovative entry is actually more likely

[3] Some studies seem to ignore it altogether, examples of which include Castellacci (2008) and Peneder (2010). Unlike the latter, which does not even mention the knowledge base as an element of the technological regime, the former does reference it. However, it also ignores the knowledge base in its actual analysis.

under higher appropriability. It is possible that this is due to the rather extreme assumption that imitation is costless without patents and the setting for that particular result where no sequential innovations are considered, but it nevertheless highlights the positive impact appropriability can have on entry.

Cumulativeness and entry. Next, if cumulativeness of technology in the industry is high, it would become more critical for firms to absorb existing knowledge in order to successfully innovate. This is also highlighted in the discussion of sequential innovation in the patent policy literature (Green & Scotchmer, 1995; Bessen & Maskin, 2009; Norman et al., 2016; Parra, 2019). Since absorbing knowledge can rarely, if ever, be accomplished without considerable cost and effort, higher cumulativeness will tend to make it harder for entrants to enter. The key difference here in relation to the situation with appropriability is that this affects whether the imitation costs are actually incurred, rather than the level of the imitation costs itself.

However, high cumulativeness can also mean that any successful innovation can potentially lead to a stream of benefits, which may encourage entry instead. This is because the original innovator will tend to have an edge over the competitors in creating a new innovation that builds on an innovation it owns, due to the differences in the level of knowledge on that particular innovation[4]. In this case, as with appropriability, an entrant may actually be more likely to enter, especially if it believes that it can successfully innovate before or soon after entry.

[4] Note, however, that this does not preclude the situation where the competitor ends up building on the innovation instead. Changes in cumulativeness at the technology level does not affect the extent to which the original innovator has an advantage over the competitors. Instead, it simply means that an innovation can potentially provide allow the firm to reap benefits beyond the value of the innovation itself.

Opportunity and entry. Finally, opportunity can have a positive effect on entry, especially when the entrant is innovative. Higher level of opportunity means that with the same level of investment, an entrant has a better chance of successfully innovating (Malerba & Orsenigo, 1997). In particular, one of the major ways to describe this is essentially how many innovations are left to "discover" or "fish" from as Dosi, Marengo, & Pasquali (2006) put it[5]. One might note, however, that while this will likely be an advantage for innovative entrants, the benefit for non-innovative entrants is less clear, a problem that will be addressed in one of the later hypotheses. One may also note that this is an element emphasized in the industry life cycle model, where industries in the early stages are generally characterized as having a high level of technological opportunity, along with higher rates of entry, compared to the later stages (Agarwal & Gort, 2002).

However, there is also the problem that the benefits of higher opportunity also apply to the incumbents. This means that even if the level of opportunity is high, there is the possibility that they are mostly taken up by the incumbents, which would 'crowd out' the potential entrants (Marsili, 2002). In this case, higher opportunity could actually be detrimental to entry, as the higher level of opportunity increases the gap between entrants and incumbents. This could apply to even non-innovative entrants, because they would be outperformed by both incumbents and other potential entrants who are diversifiers, innovative, or both.

Table 1 summarizes the effects of each element of the technological regime.

[INSERT TABLE 1 ABOUT HERE]

[5] Similar interpretations can be found in Hall, Helmers, & von Graevenitz (2015), where greater technological opportunity is associated with the number of ways in which a technology may be applied. Others interpret this as firms having the ability to choose the innovation possibility that is best suited for them (Song, Hooshangi, Zhao, & Halman, 2014; Lee, Park, & Bae, 2017)

Interaction effects

One issue with the discussion in the previous section is that the impact from each element of the technological regime is seen as being quite separate, like in some studies that focus on a single industry (Agarwal & Gort, 1996; Hall & Ziedonis, 2001; Klepper & Simons, 2000). This view, however, contradicts the theoretical results from the patent policy literature and some of the empirical results from the technological regime literature. The former shows that an increase in the level of appropriability can have very different impacts depending on how cumulative the innovation process is, although this generally focuses on incentives to innovate rather than firm entry. The latter also shows something similar, although in a somewhat rough manner, only implying that the interaction effect might exist by dividing the industries into a small number of types.

Regardless, the studies show that at least some interaction effect may exist between the elements of the technological regime when it comes to how they can impact firms (Green & Scotchmer, 1995; Bessen & Maskin, 2009; Norman et al., 2016; Parra, 2019). In addition, while the focus tends to be on the incentive to innovate, they still provide sufficient indication of how entry might be affected as well. Bessen & Maskin (2009), for example, shows that in a setting where innovation is cumulative in nature, the patent system will tend to discourage the latecomer firm, which may be interpreted as an entrant, from deciding to invest in innovation when another firm is already doing so.

Appropriability and cumulativeness. If one sees the investment in innovation as being analogous to the decision to enter, the results from the patent policy literature shows that the increase in appropriability tends to have a more negative impact on firms' entry decisions under higher cumulativeness. The reason why this might be the case is that under high cumulativeness, the need for entrants to absorb the existing innovations will be more critical. In this condition, an increase in appropriability can have a larger impact because higher cumulativeness implies that the necessity of imitating previous innovations is higher. In other words, the need to actually incur the costs of absorbing existing knowledge would increase, which serves to amplify the effects of imitation costs.

While it is also possible that the benefits will be amplified instead, this requires that the entrant either manages to innovate before entry or soon after. In addition, the entrant may not necessarily succeed in creating new innovations afterwards, which increases the risk of entry.

Appropriability and opportunity. As explained previously, higher appropriability can be seen as equivalent to higher imitation costs, which means that while entrants will find it difficult to absorb and use existing knowledge, they will also benefit more from their own innovations. If opportunity is high in this case, two mechanisms can come into effect. First, because this implies that innovation involves lower costs, entrants can now expect higher returns if they manage to successfully innovate. In other words, higher appropriability can increase the positive impact of higher opportunity.

On the flip side, it also means the entrants are even worse off if much of the opportunity is taken up by the incumbents. In this case, the higher level of opportunity simply means that the incumbents get higher net benefits than before, which can lead to an increased gap in performance with the entrants. This would make entry even more difficult, particularly since this also makes it less likely that they can survive afterwards. The less myopic the entrants are, the stronger this effect will be.

Overall, while exactly which effect will be dominant is somewhat ambiguous, in this paper the hypothesis will be stated under the assumption that the first effect will be stronger. This is because the cost side can be argued to be of more immediate concern than the eventual dominance by the incumbents, since the latter will likely only impact entrants over the long term.

Cumulativeness and opportunity. The impact of cumulativeness can also be affected by the level of opportunity, though in a slightly different manner. The benefits of high cumulativeness to the initial innovator mainly comes from the fact that they have an edge over their competitors in building on their own innovations, the effects of which would amplify under high cumulativeness. If there is a high level of opportunity, however, it would be easier for the competitors to build on that innovation instead. This would apply to entrants as well, assuming they are not crowded out by the incumbents. In addition, the main costs of cumulativeness, raising the need to absorb existing knowledge, can be less of a concern if the level of opportunity is high due to the decrease in costs of innovating.

Determining the direction of effect is harder than with the case with appropriability and opportunity because both mechanisms require a long-term view to be relevant for entry decisions.

Based on the analysis so far, the following hypotheses can be stated.

Hypothesis 1a. The effect of appropriability on rate of entry becomes more negative as the level of cumulativeness increases.

Hypothesis 1b. The effect of appropriability on rate of entry becomes more positive as the level of opportunity increases.

Hypothesis 1c. The effect of cumulativeness on rate of entry becomes more positive as the level of opportunity increases.

Entrant type

One may also notice that much of the discussion in the first section tends to ignore the distinction between different types of entrants. However, previous studies such as Bayus & Agarwal (2007) show that different types of entrants have different capabilities, thus facing different advantages and disadvantages. One way in which entrants can be classified is whether the entrants are innovative or not. Here, innovative entrants refer to those that enter the industry with an innovation. Non-innovative entrants, on the other hand, essentially refer to those that imitate existing innovations (Ethiraj & Zhu, 2008). The importance of imitation, and by extension non-innovative entrants, is also evident from much of the literature on innovation and patent policy, which is also noted in the previous sections.

Keeping this in mind, one may note that the impact of appropriability can be quite different depending on whether the entrant is innovative or not. For the former type, while high appropriability can be problematic for innovating in the future, it can also allow them to benefit more from any innovations they own. The latter type, on the other hand, would not have an innovation from which they can benefit. While the net effect is not entirely clear, this still shows that any negative impact from high appropriability should be smaller for innovative entrants compared to the non-innovative ones.

As for cumulativeness, the effect should be somewhat similar. Higher cumulativeness implies that previous innovations have a larger impact on successfully innovating in the future, so if a firm owns some of those, it would be facing much lower costs. Again, this would be of little value to non-innovative entrants because they would not have an innovation to rely on by definition, but this could be very useful for innovative entrants.

Opportunity can also have very different effects. By definition, opportunity represents the ease with which a firm can innovate, which by extension should make innovative entry easier. For non-innovative entry, however, opportunity would have relatively little impact, because no innovative effort is involved. Thus, the following hypotheses emerge.

Hypothesis 2a. For innovative entrants, relative to non-innovating entrants, the effect of appropriability on rate of entry is more positive.

Hypothesis 2b. For innovative entrants, relative to non-innovating entrants, the effect of cumulativeness on rate of entry is more positive.

Hypothesis 2c. For innovative entrants, relative to non-innovating entrants, the effect of opportunity on rate of entry is more positive.

Another popular categorization of entrants is whether they are diversifying or de novo (Bayus & Agarwal, 2007). One of the main differences is that diversifying entrants in general tend to be thought of as having access to more resources due to its parent firm. Due to this advantage, the impact of appropriability or cumulativeness is likely be smaller on diversifiers than de novo entrants, since the cost in absorbing existing technologies can be mitigated by the resources they can draw on.

Of particular interest is the comparison between diversifying entrants that come from related industries with the other types. This is due to the issue of whether a firm already has knowledge that may be applied in the industry. This would generally be the case with diversifying entrants from related industries, which would mean that they have an advantage in absorbing technologies that are useful in the industry, assuming they do not already have such technologies. Thus, such entrants would face lower costs overall in absorbing new knowledge, which can lead to the increase in appropriability or cumulativeness having a smaller impact on them. In fact, the effects would be even smaller than those for other diversifying entrants because related diversifiers would also have access to more resources than de novo entrants. Thus, the following hypotheses can be stated.

Hypothesis 3a. For diversifying entrants, relative to de novo entrants, the effect of appropriability on rate of entry is more positive.

Hypothesis 3b. For diversifying entrants, relative to de novo entrants, the effect of cumulativeness on rate of entry is more positive.

Hypothesis 4a. For diversifying entrants from related industries, relative to other diversifying entrants, the effect of appropriability on rate of entry is more positive.

Hypothesis 4b. For diversifying entrants from related industries, relative to other diversifying entrants, the effect of cumulativeness on rate of entry is more positive.

One may note that these hypotheses do not mention opportunity. One reason for this is because there is little reason to predict that the effect of opportunity is different unless one type tends towards a higher rate of innovative entry or has a better chance of later successfully innovating than the other. Changes in opportunity simply lead to changes in the rate of successful innovation given the same overall level of investment, thus both types would be affected to the same extent. For example, when comparing diversifying entrants from related industries that are also (non-)innovative with other (non-)innovative entrants, the differences may not necessarily be significant. Of course, if these assumptions are not true, then the impact of opportunity would also change depending on the entrant type.

DATA AND EMPIRICAL STRATEGY

Data

The study will rely on two main sources of data. The first is ORBIS, which contains information on all firms. This includes not only financial data, but also the date of incorporation and what subsidiaries the firms have. The second is Patentsview, which concerns the patents registered in the USPTO (United States Patent and Trademark Office). Here, data on the application and grant date of patents, their assignees, their abstracts, etc. are listed, which should be useful in measuring the elements of the technological regime.

The sample will consist of US firms that were founded between 2010 and 2016. Effectively, the sample will include all US firms that entered some industry during that period. However, those that are not entrants will also be used as a means of selecting industries, among the manufacturing industries (i.e., industries which have 31, 32, or 33 as the first two digits of NAICS) to include in the sample. In addition, these will also be used to calculate variables such as the Herfindahl index. It should also be noted that the period analyzed is only from 2010 to 2014, even though the data for the later years are also used. This is mainly because this paper will take a similar approach to Cockburn & MacGarvie (2011), in that firms that enter later will be used as proxies for potential entrants.

As for actually constructing the sample, the first step will be to merge the two datasets together. This will be done primarily through the matching of the names of the assignees in Patentsview with the names of the companies as listed in ORBIS. In many cases, however, the names listed in the former dataset and the latter tend to have many differences, such as whether the letters have been capitalized, the corporate identifier (Co., Ltd., etc.), or even whether commas are included or not. To address these problems, all names will be converted into lowercase letters, and all punctuations have been removed as well. The corporate identifiers have also been removed.

All firms on ORBIS that are listed as either a branch or a subsidiary in the same industry as the parent will be removed as well. With the former case, the justification is that a branch of a company is little more than an expansion into a different geographic region, and is not even legally independent. Subsidiaries in the same industry, on the other hand, will be removed partly because such companies tend to have essentially the same name as the parent. The more important reason, however, is that for the purposes of this paper such subsidiaries do not contribute significantly, since they are in most cases more or less controlled by the parent anyway. They would be more useful if one could assume that such subsidiaries are essentially the parent's attempt to expand into certain segments within the industry, but this is unfortunately outside the scope of this analysis.

Finally, after filtering for entrants in the relevant time period, only the five-digit NAICS industries that seem to actually use patents as an appropriability mechanism will be used. This is based on the concern that, as Cohen, Nelson, & Walsh (2000) shows, patents are not always used in some industries due to doubts on their perceived effectiveness. While one might argue that this is not necessarily a problem for the measures for the technological regime, this is valid only if the knowledge embodied in the patents are representative of the industry as a whole, which may or may not be true. This is particularly an issue because, as pointed out in many of the studies on the subject, patents are not the only appropriability mechanism available to firms. This means that in industries where patents are not used very often, patents are less likely to be useful as measures for the technological regime, since it is less likely that they will provide sufficient information on the overall characteristics of the technology present there. In other words, many of the measures relying on patents will be rather unreliable in such industries. In this study, this problem will be addressed by restricting the sample to industries that either have an above-median number of patents per employee, or above-mean number of patents per firm, based on the criterion used in the report by Economics and Statistics Administration & United States Patent and Trademark Office (2012).

The reason why five digits are to be used instead of, say, four digits is primarily empirical, in that this is in response to a need to maximize the number of industries available for analysis. However, there is also a more conceptual reason as well, which is that using the four-digit level essentially assumes that the industries within that group are more or less homogenous in their technological regimes. While there may be some merit to this, it is also likely that this is only true for some elements like appropriability, since at least some factors like availability of patents are likely to be similar within the same group of industries. For other elements such as opportunity or cumulativeness, the situation could be more complex. Opportunity, in particular, may be subject to more variation depending on whether there was a recent change in the technological trajectory. For example, the development of the electric car may have changed the opportunity conditions in the industry for manufacturing engines, but likely had less of an effect on that related to steering or brakes. While the use of five digits is not necessarily the best way to address this possibility, it is likely one of the better options nonetheless.

Six digits are not used either, because in some cases, the disaggregation is somewhat excessive, especially in the context of competition. For example, NAICS 311811 refers to retail bakeries while 311812 refers to commercial bakeries. While the two are evidently distinct, most of that difference seems to be in terms of size rather than any technological or market difference. While the distinction would certainly be useful in some contexts, this is somewhat superfluous in this analysis. It may even detrimental, because all this would do is create industries with very few firms, which can affect the reliability of the industry-level measures.

Variables

Dependent variable. The event of interest is the entry of the firm in the industry, denoted by a binary variable. This will be based on the year of incorporation as listed in ORBIS, such that on the year of incorporation, the variable will be equal to 1, and 0 on all years before that. Note that since a firm will be dropped from the sample once entry has occurred, any eventual exit of the entrants is not considered.

Explanatory variables. The first independent variable of interest is the level of *appropriability* in the industry. Here, this is measured by the industry-wide R&D elasticity[6] of market share based on the method used in Hall (1993) and Knott (2008) is also used to measure appropriability at the industry level. This is obtained based on the following equation[7]:

$$MS_i = AK_i^{\alpha_j} L_i^{\beta_j} R_i^{\gamma_j} e^{u_i} \qquad (1)$$

Here, MS_i is the market share of the firm i in industry j as measured by its sales, K_i is the capital as measured by total assets, L_i is the labor input as measured by the number of employees, and R_i is the R&D expenses of the firm. The last term is essentially the error associated with the equation. This can then be transformed into:

$$\log(MS_i) = \log(A) + \alpha_j \log(K_i) + \beta_j \log(L_i) + \gamma_j \log(R_i) + u_i \qquad (2)$$

This will then be estimated by industry and year. The idea here is that if appropriability is high, then by definition the returns to R&D should be high as well, Thus, the coefficient γ_j as estimated can be used as a measure of appropriability, since it is essentially the average R&D elasticity of industry j.

[6] This is also called the "research quotient" in Knott (2008).

[7] This is a slightly modified version of equation (1) from Hall, Mairesse, & Mohnen (2009), which is also used in a similar form in Knott (2008) and Hall (1993). Specifically, external knowledge is not included here, because this is estimated separately for each industry rather across several industries at once, as is the case in the studies on which this is based. Since this is only estimated for one industry at a time, the "external knowledge" may be considered to be perfectly and negatively correlated with the firm's own knowledge depending on the operationalization. The estimation is done separately for each industry because the value of R&D elasticity for each industry is required.

This is chosen over measures that use the forward self-citation[8] ratio of the patents in the industry, which are based on Park and Lee (2006) and Trajtenberg, Henderson, and Jaffe (1997), due to several concerns. Some may find, for example, that this measure is rather questionable given that it could easily be linked to cumulativeness at the firm level. While conceptually this is not necessarily an issue, it still indicates that it would be advisable to use an alternative measure to avoid any problems. In addition, collinearity with the measure for cumulativeness, which is based on similarity of the abstracts of the patents, may also be an issue. If, for example, patents filed by the same company tend to be worded similarly regardless of the actual knowledge involved, then high appropriability as measured by self-citation ratios may systematically lead to high cumulativeness. While it is possible that some correlation does in fact exist between the two elements, it would be more prudent to try to make sure that the relationship does not simply arise out of how the measures are constructed.

The second, *cumulativeness* at the industry level, is measured as the average of the similarity of the abstracts between the focal patents in the industry and their cited patents[9]. The measure is constructed based on the methodology in Kelly, Papanikolaou, Seru, & Taddy (2018), where the approach of using all patents filed before the focal one is adapted to using only those that have been cited by the focal one. The idea here is that a high degree of similarity implies heavier dependence of the focal patent on the knowledge of cited patents. In other words, a high value in this measure should represent a more significant impact of the previous patents in the successful filing of the focal patent.

[8] Forward citations refer to instances where future patents cite the focal patent. For example, if patent A was issued in 2010 and this was cited by patent B and C that were filed later, then patent A has two forward citations. In addition, if patent B was filed by the same firm as A but not C, then the forward self-citation ratio would be 0.5.

[9] While other measures exist, such as self-citations, they are generally constructed under the premise that cumulativeness refers to firms building on their own knowledge. In this study however, following the conceptualization in the literature on patents and innovation, cumulativeness refers to building on past knowledge regardless of who had created it. This makes the use of self-citations somewhat inappropriate.

The third is the *opportunity* at the industry level, which is measured as the increase in the number of patents in the industry over 3 years, normalized by the increase in the R&D expenditure of the industry over the corresponding period. This is based on the measure used in Park & Lee (2006), the difference being the limit of 3 years and the normalization. The first difference is due to the concern that while the technological regime at the industry level is likely rather stable, it would not necessarily be completely static either. In other words, this is a way to allow for some variation without letting that variation be too extreme. In addition, this may help mitigate the issue of the time lag between R&D and patenting without arbitrarily deciding what that time lag actually is. The second feature, normalization, is included in order to fit the measure better with the definition of the concept. Opportunity is related to the probability of successful innovation given the costs, and including the R&D expenses in the measure makes it possible interpret it such that an industry can be seen as having a high opportunity if the value is higher than 1 and vice versa.

As for the types of entrants, an *innovative entrant* is denoted by whether the firm has at least one patent application by the time of entry. Patent applications will be used instead of the granted patents simply because not having a patent is not necessarily indicative of not having an innovation the firm can use, especially since some lag exists between applying for a patent and actually being issued a patent[10].

In addition, an entrant is classified as being a *diversifier* if it is a subsidiary of a firm in another industry, a *related diversifier* if the parent firm is in a related industry, and *de novo* otherwise. This is determined by the use of subsidiary data in ORBIS, which lists what subsidiaries each firm has. Whether the diversifier is from a related industry will be

[10] On the other hand, the technological regime variables will rely on patents that have actually been issued even if they are organized by application date rather than grant date. This is due to the limit that published applications do not tend to have citations listed in them.

determined by the degree of overlap in the CPC (Cooperative Patent Classification) group (e.g. B01L, which refers to “chemical or physical laboratory apparatus for general use”) between the industries, partly inspired by the measure for knowledge-relatedness from Breschi, Lissoni, & Malerba (2003). Essentially, for each patent p issued in the years 2010-2014, a vector $\boldsymbol{F^p}$ indicating whether the patent is assigned the CPC class is defined, so that for each class i,

$$F_i^p = \begin{cases} 1, & patent\ p\ is\ assigned\ class\ i \\ 0, & otherwise \end{cases} \quad (3)$$

Then the sum of the vectors is taken for each industry j, giving a vector $\boldsymbol{C^j}$ of the number of times each class is assigned to patents in the industry. Then, to determine the degree of relatedness, S_{jk} between a pair of industries j and k, the cosine similarity for the two vectors is calculated, leading to:

$$S_{jk} = \frac{\boldsymbol{C}^j \cdot \boldsymbol{C}^k}{\|\boldsymbol{C}^j\| \cdot \|\boldsymbol{C}^k\|} \quad (4)$$

Using this measure, an industry is considered related to the other if the value of the measure is larger than or equal to 0.5.[11]

Finally, two industry-level control variables are included in the analysis as well. The first is the Herfindahl Index, which will be included because industry concentration is often cited as an important factor in competition. The second is the average size of the firms in the industry as the number of employees, which is used to represent barriers to entry in the more traditional sense.

[11] Note that this is an industry-level variable and not a firm-level one, which means that whether each individual firm has patents does not factor in here. For example, suppose Firm 0 in Industry A establishes a subsidiary Firm 1 in Industry B. If Industry A and B are related, based on the patents in the two industries, then the value of the related diversifier dummy will be 1. If the two industries are not related, the value will be 0. Neither case needs to consider whether Firm 0 or 1 has patents, so long as the industries involved have patents.

The descriptive statistics of the dataset used for the analysis are presented in Table 2. Table 3 shows the correlation table.

[INSERT TABLES 2 AND 3 ABOUT HERE]

Estimation method

The analysis will use the Cox model, extended for use with time-varying covariates and stratified by industry-year. The Cox model is often used for survival analysis, which may make this model seem like an odd choice in this setting. However, it is also true that the coefficients for the explanatory variables represent their effect on the hazard rate of an event occurring given that the event has not occurred up to a point in time, the event being the entry into an industry in this study. Combined with the manner in which the data is arranged, where each firm is treated as having been a potential entrant since 2010[12], this would make it possible to draw conclusions on how the covariates affect the rate of entry.

The Cox model is chosen over a probit or logit model because in comparing the entrants of that year with the other potential entrants, it may be safer to account for possible firm-year specific effects. This is also similar to Cockburn & MacGarvie (2011), where the authors use a discrete hazard model. There is also the added advantage that the interpretation of interaction terms makes better sense, while this is not the case with probit or logit models (Ai & Norton, 2003). This is because the form of the regression can be treated as being log linear, where the effect of the covariates on the log hazard rate is estimated (Cader & Leatherman,

[12] More precisely, the data is arranged according to the manner described in Zhang et al. (2018), where each firm is given a row for each year up until it entered or the year 2014, and the entry variable is defined as 1 for the year it actually entered and 0 for all previous years.

2011). In other words, interpreting the interaction terms is as simple as in the case of the usual linear regressions.

In addition, the Cox model is chosen over other survival models for two reasons. The first is that the Cox model does not require assumptions on the baseline hazard function. This makes it easier to use than most parametric hazard models, and also more reliable. The second reason is that it is easier to include multiple covariates, which tends to be a problem with nonparametric models. This is particularly important in this context because the analysis requires the use of interaction terms between most of the major independent variables as well.

The main focus of the estimation will be centered around appropriability, cumulativeness, and opportunity conditions of the industry, along with whether the entrant is innovative and whether it is a related diversifier, diversifiers that are not from a related industry, or de novo. Testing Hypotheses 1a, 1b, and 1c will involve testing whether the coefficient for the interaction between appropriability and cumulativeness is negative, and whether those for the interactions involving opportunity are positive. For the remaining hypotheses, the aim will be to test the interaction between the three technological regime variables with the dummies for innovative entrants, related diversifiers, and other diversifiers. According to the hypotheses, the coefficients should be positive for all these interactions, with the ones for related diversifiers being the largest. In addition, the magnitude of the coefficient for interactions involving related diversifiers should be larger than that for other diversifiers.

RESULTS

Initial results

The initial results of the analysis are shown in Table 4. Models 1 through 4 use the full sample available for the analysis. Model 1 is the benchmark model with no interaction effects included. Model 2 includes the interaction effects within the technological regime variables, while Model 3 includes only the interaction effects with entrant types. Model 4 includes both types of interaction effects. Model 5 focuses only on the industry-years that have at least 50 entrants. Finally, Model 6 only includes entrants for whom the data are available up to the year they actually entered. In other words, entrants who got right censored due to lack of data in any of the independent variables, despite entering on or before 2014, were removed from the analysis. Models 5 and 6 are meant to be robustness checks, to see if there are issues with the sample that may be affecting the results.

All results seem to agree that Hypothesis 1a is supported, while all other hypotheses are not. It is particularly worth noting the coefficient for the interaction between appropriability and opportunity is significant at the 5% level, but in a direction opposite to that predicted in Hypothesis 1c.

[INSERT TABLE 4 ABOUT HERE]

One should also note that the main effects of the three elements of the technological regime being analyzed are not as expected. Opportunity, in particular, is interesting in that it actually has a negative effect on entry. This suggests that, as mentioned in the analysis, the fact that the increase in opportunity also affects incumbents leads to the entrants being discouraged from entry instead.

This may also explain why the coefficient of the interaction between appropriability and opportunity become negative. If opportunities are more beneficial to incumbents, then one might suppose that the imitation cost aspect of appropriability becomes more pronounced than the potential benefits of innovating after entry. In such a case, potential entrants would be further discouraged from entering the industry.

The results also do not support the hypotheses related to entrant types, which seems somewhat difficult to reconcile theoretically. At face value, they show that the impact of the technological regime is more or less homogenous regardless of entrant type. At the same time, however, one might also argue that there are two problems with the model. The first is that the model does not consider the two typologies of entrants jointly. For example, one might suppose that innovativeness can lead to big differences in behavior among diversifiers but not necessarily among de novo entrants especially if innovativeness only leads to meaningful changes when backed by access to large amounts of resources.

Extended results

The reason why the interaction of entrant types with technological regime variables do not yield any significant results may be due to oversights on two issues. The first is that among diversifiers or related diversifiers, one may also suppose that some are innovative while others are not. In other words, instead of treating the two ways of categorizing entrants separately, perhaps it makes more sense to consider them jointly. The results of using this new specification are summarized in Table 5 and 6. The full results are shown in Table A1, found in the appendix[13].

[INSERT TABLES 5 AND 6 ABOUT HERE]

Interestingly, the results in Table 5 and 6 now partly support the hypotheses on entrant types. Specifically, one can see that the impact of all three technological regime variables is more positive for innovative, related diversifiers. This is somewhat consistent with the hypotheses, in that being both innovative and a related diversifier causes the impact of all

[13] Models 1, 2, and 3 in this table are the same as Models 4, 5, and 6 in Table 3. This is also true for Table A2.

three technological regime variables to be more positive. It also suggests, however, that neither property is necessarily enough to lead to changes in entry decisions.

This suggests that innovative related diversifiers are indeed somehow different from non-innovative related diversifiers or other types of innovative entrants. One possibility is that differences in knowledge stock can change the value of innovations, and thus change the impact of innovativeness as well. One of the key characteristics of related diversifiers is that their knowledge is already useful in some way in the industry. If so, innovativeness could be more beneficial for them than others because their innovations may be more valuable.

Another point that is worth addressing is the fact that the interaction effects of the technological regime variables are assumed to be homogenous for all entrant types, despite the fact that this is theorized to be untrue for the individual effects. One of the main ways in which the impact of an element changes for different types of entrants is that the magnitude of the costs and benefits change based on their characteristics. Since this is also true of how interaction effects may arise, it is possible that interaction effects are different for each entrant type.

Table 7 shows the results of the analysis that verifies whether this is indeed the case, the full version of which can be found in the appendix, in Table A2.

[INSERT TABLE 7 ABOUT HERE]

The results in Table 7 have several implications. First, the impact of the technological regime on unrelated diversifiers are indeed different from that of others, although in a more subtle way than the hypotheses imply. In particular, the impact of appropriability, which is the one directly referenced in Hypothesis 3a, only seems to change for diversifiers when also coupled with high levels of opportunity. This implies that the diversifier's advantage in resources only translates to increased entry when it believes there to be a high enough benefit from continuing to operate and innovate in the industry.

The results for related diversifiers are even more interesting. Appropriability seems to have a more negative impact on the entry of related diversifiers, but this changes dramatically when coupled with high cumulativeness. Normally, the interaction effect of these two elements tends to be negative because the increased imitation costs from high appropriability are amplified under high cumulativeness by the fact that more imitation needs to occur in the first place. However, the results for related diversifiers are different, which may be because they tend to be less affected by costs in the first place, while the benefits may be amplified because they are more likely to successfully innovate than others.

DISCUSSION AND CONCLUSION

The results from this analysis provide a better understanding of how the technological regime can affect firm entry. In particular, they show that interaction effects are significant enough that they need to be considered thoroughly when changing policies affecting the technological environment or when making entry decisions. In addition, considering the different types of entrants is found to be an important aspect for such decisions, although the exact mechanisms seem to be more complicated than was initially assumed.

These results may also be useful for shedding light on how industries might evolve over time. For example, one might infer from the results in Table 6 that in a Schumpeter Mark II industry, related diversifiers are more likely to enter than others. This may mean that the industry would tend towards oligopoly, which may further increase barriers to entry. In other words, the consequences of changes in the technological regime may be longer lasting than one might assume.

There are still many issues that will need to be addressed in future studies on the subject. The first is the problem of industry choice. In this study, one of the main concerns was the fact that some industries had to be dropped due to the availability and reliability of measures. This does, however, mean that it is unclear whether the results in this paper can also be applied across different industries that do not rely on patents as much. This is especially worth addressing because the technological regime framework need not be restricted to patent-intensive or technology-intensive industries. Finding measures that do not rely on patents may be one way to deal with this, but this runs into feasibility issues.

Another problem with the analysis is that it is currently unable to distinguish between true de novo entrants and spinoffs or spinouts. Spinoffs are firms established by a previous employee of a firm in the same industry, while spinouts are similar except that the firm the founder comes from is in a different industry. Their unique advantages against other types of entrants have been highlighted in several studies such as Klepper & Sleeper (2005) or Adams, Fontana, & Malerba (2015). Unfortunately, it is very difficult to see if an entrant firm is a spinout or spinoff in ORBIS, mainly because it does not provide detailed information on a company's founder.

The somewhat inconsistent level of disaggregation in NAICS is also an issue that should be addressed, which unfortunately cannot be solved by simply taking all six digits. This comes into light from the problem of the selection of firms, particularly due to the existence of subsidiaries that are in the same industry as the focal firm. While these were ignored in this particular analysis, this shows that there may be segments within the industry that induced the firm to expand by creating a subsidiary. For example, a firm in the automobile manufacturing (336111) that decided to establish a subsidiary in that same industry may have done so because it wanted to address a different market segment. It could also represent differences in opportunity conditions, which would not be captured when the usual methods of classifying industries such as NAICS is used. The opposite may also be true, in that in some cases, industries that are not necessarily different may have different NAICS codes, which can also confound the analysis. In either case, gaining data at the product market level or so may be desirable.

REFERENCES

Acemoglu, D., & Cao, D. 2015. Innovation by entrants and incumbents. ***Journal of Economic Theory***. 157: 255-294.

Acs, Z., & Audretsch, D. 1989. Small-firm entry in US manufacturing. ***Economics, New Series***. 56 (222): 255-265.

Adams, P., Fontana, R., & Malerba, F. 2015. User-industry spinouts: Downstream industry knowledge as a source of new firm entry and survival. ***Organization Science***. 27 (1): 18-35.

Agarwal, R., & Audretsch, D. 2001. Does entry size matter? The impact of the life cycle and technology on firm survival. ***The Journal of Industrial Economics***. 46 (1): 21-43.

Agarwal, R., & Gort, M. 1996. The evolution of markets and entry, exit and survival of firms. ***The Review of Economics and Statistics***. 78 (3): 489-498.

Agarwal, R., & Gort, M. 2002. Firm and product life cycles and firm survival. ***American Economic Review***. 92 (2): 184-190.

Agarwal, R., & Shah, S. 2014. Knowledge sources of entrepreneurship: Firm formation by academic, user and employee innovators. ***Research Policy***. 43: 1109-1133.

Ai, C., & Norton, E. 2003. Interaction terms in logit and probit models. ***Economics Letters***. 80 (1): 123-129.

Audretsch, D. 1991. New-firm survival and the technological regime. ***The Review of Economics and Statistics***. 73 (3): 441-450.

Audretsch, D. 1997. Technological regimes, industrial demography and the evolution of industrial structures. ***Industrial and Corporate Change***. 6 (1): 49-82.

Bayus, B., & Agarwal, R. 2007. The role of pre-entry experience, entry timing, and product technology strategies in explaining firm survival. ***Management Science***. 53 (12): 1887-1902.

Bessen, J., & Maskin, E. 2009. Sequential innovation, patents, and imitation. ***The RAND Journal of Economics***. 40 (4): 611-635.

Breschi, S., Lissoni, F., & Malerba, F. 2003. Knowledge-relatedness in firm technological diversification. ***Research Policy***. 32: 69-87.

Breschi, S., Malerba, F., & Orsenigo, L. 2000. Technological regimes and Schumpeterian patterns of innovation. ***The Economic Journal***. 110: 388-410.

Cader, H., & Leatherman, J. 2011. Small business survival and sample selection bias. ***Small Business Economics***. 37: 155-165.

Capone, G., Malerba, F., & Orsenigo, L. 2019. Spinoffs in context: Entry and performance across different industries. ***Industrial and Corporate Change***. 28 (2): 259-282.

Chang, S.J. 1996. An evolutionary perspective on diversification and corporate restructuring: Entry, exit, and economic performance during 1981-1989. ***Strategic Management Journal***. 17: 587-611

Chen, P., Williams, C., & Agarwal, R. 2012. Growing pains: Pre-entry experience and the challenge of transition to incumbency. ***Strategic Management Journal***. 33: 252-276.

Cockburn, I., & MacGarvie, M. 2011. Entry and patenting in the software industry. ***Management Science***. 57 (5): 915-933.

Cohen, W., & Levin, R. 1989. Empirical studies of innovation and market structure. ***Handbook of Industrial Organization***. 2: 1059-1107.

Cohen, W., & Levinthal, D. 1989. Innovation and learning: The two faces of R&D. ***The Economic Journal***. 99 (397): 569-596.

Cohen, W., Nelson, R., & Walsh, J. 2000. ***Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not)***. (NBER working paper no. 7552). Cambridge, MA: National Bureau of Economic Research.

Dosi, G., Marengo, L., & Pasquali, C. 2006. How much should society fuel the greed of innovators? On the relations between appropriability, opportunities and rates of innovation. ***Research Policy***. 35: 1110-1121.

Economics and Statistics Administration, & United States Patent and Trademark Office. 2012. ***Intellectual property and the U.S. economy: Industries in focus***.

Ethiraj, S., & Zhu, D. 2008. Performance effects of imitative entry. ***Strategic Management Journal***. 29: 797-817.

Garavaglia, C. 2016. Thinking about entry of firms: A theoretical discussion. ***International Business Research***. 9 (2): 46-73.

Geroski, P. 1995. What do we know about entry?. ***International Journal of Industrial Organization***. 13: 421-440.

Green, J., & Scotchmer, S. 1995. On the division of profit in sequential innovation. ***The RAND Journal of Economics***. 26 (1): 20-33.

Hall, B. (1993). The stock market's valuation of R&D investment during the 1980's. ***The American Economic Review***. 83 (2): 259-264.

Hall, B., & Ziedonis, R. 2001. The patent paradox revisited: an empirical study of patenting in the US semiconductor industry, 1979-1995. ***The RAND Journal of Economics***. 32 (1): 101-128.

Hall, B., Mairesse, J., & Mohnen, P. 2009. ***Measuring returns to R&D*** (NBER working paper no. 15622). Cambridge, MA: National Bureau of Economic Research.

Hall, B., Helmers, C., & von Graevenitz, G. 2015. ***Technology entry in the presence of patent thickets*** (NBER working paper no. 21455). Cambridge, MA: National Bureau of Economic Research.

Helfat, C., & Lieberman, M. 2002. The birth of capabilities: market entry and the importance of pre-history. ***Industrial and Corporate Change***. 11 (4): 725-760.

Jaffe, A., Trajtenberg, M., & Henderson, R. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. ***The Quarterly Journal of Economics***. 108 (3): 577-598.

Kelly, B., Papanikolaou, D., Seru, A., & Taddy, M. 2018. ***Measuring technological innovation over the long run*** (NBER Working Paper no. 25266). Cambridge, MA: National Bureau of Economic Research.

Klepper, S. 1996. Entry, exit, growth, and innovation over the product life cycle. ***The American Economic Review***. 86 (3): 562-583.

Klepper, S., & Simons, K. 2000. The making of an oligopoly: Firm survival and technological change in the evolution of the U.S. tire industry. ***Journal of Political Economy***. 108 (4): 728-760.

Klepper, S., & Sleeper, S. 2005. Entry by spinoffs. ***Management Science***. 51 (8): 1291-1306.

Knott, A.M. 2008. R&D/Returns causality: Absorptive capacity or organizational IQ. ***Management Science***. 54 (12): 2054-2067.

Lee, J., Park, J., Bae, Z. 2017. The effects of licensing-in on innovative performance in different technological regimes. ***Research Policy***. 46: 485-496.

Lin, P., & Huang, D. 2008. Technological regimes and firm survival: Evidence across sectors and over time. ***Small Business Economics***. 30: 175-186.

Malerba, F., & Orsenigo, L. 1997. Technological regimes and sectoral patterns of innovation activities. ***Industrial and Corporate Change***. 6 (1): 83-118.

Mansfield, E., Schwartz, M., & Wagner, S. 1981. Imitation costs and patents: an empirical study. ***The Economics Journal***. 91: 907-918.

Markides, C., & Williamson, P. 1994. Related diversification, core competences and corporate performance. ***Strategic Management Journal***. 15: 149-165.

Marsili, O. 2002. Technological regimes and sources of entrepreneurship. ***Small Business Economics***. 19 (3): 217-231.

McAfee, R., Mialon, H., & Williams, M. 2004. What is a barrier to entry?. ***American Economic Review***. 94 (2): 461-465.

Nelson, R., & Winter, S. 1982. ***An Evolutionary Theory of Economic Change***, Cambridge, MA: Harvard University Press.

Norman, G., Pepall, L., & Richards, D. 2016. Sequential product innovation, competition and patent policy. ***Review of Industrial Organization.*** 48: 289-306.

Park, K.H., & Lee, K. 2006. Linking the technological regime to the technological catch-up: analyzing Korea and Taiwan using the US patent data. ***Industrial and Corporate Change***. 15 (4): 715-753.

Parra, A. 2019. Sequential innovation, patent policy, and the dynamics of the replacement effect. ***RAND Journal of Economics***. 50 (3): 568-590.

Pavitt, K. 1984. Sectoral patterns of technical change: towards a taxonomy and a theory. ***Research Policy***. 13: 343-373.

Peneder, M. 2010. Technological regimes and the variety of innovation behaviour: Creating integrated taxonomies of firms and sectors. ***Research Policy***. 39: 323-334.

Schumpeter, J. 1934. ***The Theory of Economic Development***, Cambridge, MA: Harvard Economic Studies.

Schumpeter, J. 1942. ***Capitalism, Socialism, and Democracy***, New York: Harper and Brothers.

Singh, H., & Montgomery, C. 1987. Corporate acquisition strategies and economic performance. ***Strategic Management Journal***. 8: 377-386.

Song, M., Hooshangi, S., Zhao, Y., & Halman, J. 2014. How does technological regime affect performance of technology development projects?. ***Journal of Product Innovation Management***. 31 (S1): 60-74.

Trajtenberg, M., Henderson, R., & Jaffe, A. 1997. University versus corporate patents: A window on the basicness of invention. ***Economics of Innovation and New Technology***. 5 (1): 19-50.

Zhang, Z., Reinikainen, J., Adeleke, K., Pieterse, M., & Groothuis-Oudshoorn, C. 2018. Time-varying covariates and coefficients in Cox regression models. ***Annals of Translational Medicine***. 6 (7).

LIST OF TABLES

TABLE 1
Summary of effects of technological regime on firm entry

	Positive	Negative
Appropriability	More benefits from innovation	Higher imitation costs
Cumulativeness	Stream of benefits from innovation	Incur imitation costs more often
Opportunity	Lower costs of innovation due to more lines of innovations to pursue	Benefits may be focused on incumbents, not entrants

TABLE 2
Descriptive statistics

	N	Min	Max	Median	Mean	S.D.
Entrant	111363	0	1	0.00	0.32	0.47
Appropriability	111363	-2.24	5.99	0.00	0.08	0.47
Cumulativeness	111363	0.04	0.34	0.15	0.14	0.02
Opportunity	111363	0.00	14.93	0.68	1.15	1.82
Nonrelated Diversifier	111363	0	1	0.00	0.20	0.40
Related Diversifier	111363	0	1	0.00	0.02	0.15
Innovative Entrant	111363	0	1	0.00	0.02	0.14
Employees per firm	111363	0.12	9327.50	6.92	17.26	53.26
HHI	111363	0.05	0.97	0.19	0.28	0.22

TABLE 3
Correlation table

	1	2	3	4	5	6	7	8
1. Appropriability	1.00							
2. Cumulativeness	0.38	1.00						
3. Opportunity	-0.08	-0.18	1.00					
4. Nonrelated Diversifier	-0.07	-0.11	0.02	1.00				
5. Related Diversifier	-0.02	-0.03	0.03	-0.08	1.00			
6. Innovative Entrant	-0.02	-0.05	0.00	0.02	-0.00	1.00		
7. Employees per firm	-0.05	-0.01	-0.12	0.00	-0.01	0.06	1.00	
8. HHI	0.15	0.20	0.32	-0.06	-0.00	-0.02	-0.05	1.00

TABLE 4
Initial results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Appropriability	-0.03*	0.35**	-0.04*	0.28	0.29	0.19
	(0.02)	(0.17)	(0.02)	(0.18)	(0.19)	(0.21)
Cumulativeness	0.68	-0.18	0.64	-0.22	-0.38	-0.09
	(0.59)	(0.83)	(0.69)	(0.89)	(0.91)	(0.99)
Opportunity	-0.01**	-0.10*	-0.02**	-0.11*	-0.13**	-0.16**
	(0.01)	(0.06)	(0.01)	(0.06)	(0.06)	(0.07)
Nonrelated Diversifier	-0.43***	-0.44***	-0.45*	-0.44*	-0.44*	-0.52**
	(0.03)	(0.03)	(0.24)	(0.24)	(0.25)	(0.25)
Related Diversifier	-0.56***	-0.56***	-1.17**	-1.13**	-1.14**	-1.35***
	(0.08)	(0.08)	(0.55)	(0.55)	(0.54)	(0.52)
Innovative Entrant	0.22***	0.22***	0.31	0.34	0.40	0.38
	(0.05)	(0.05)	(0.27)	(0.27)	(0.27)	(0.27)
Appropriability x Cumulativeness		-1.89**		-1.57*	-1.59*	1.21
		(0.82)		(0.90)	(0.93)	(1.07)
Appropriability x Opportunity		-0.05**		-0.05**	-0.06**	-0.04*
		(0.02)		(0.02)	(0.03)	(0.03)
Cumulativeness x Opportunity		0.62		0.65	0.80*	1.03*
		(0.42)		(0.41)	(0.43)	(0.55)
Appropriability x Nonrelated Diversifier			0.07	0.06	0.03	0.05
			(0.08)	(0.10)	(0.11)	(0.11)
Appropriability x Related Diversifier			0.18	0.20	0.44	0.19
			(0.16)	(0.17)	(0.27)	(0.17)
Appropriability x Innovative Entrant			-0.13	-0.11	-0.09	-0.09
			(0.13)	(0.12)	(0.12)	(0.13)
Cumulativeness x Nonrelated Diversifier			-0.03	-0.16	-0.14	0.44
			(1.69)	(1.71)	(1.74)	(1.75)
Cumulativeness x Related Diversifier			3.99	3.68	3.53	5.55
			(3.88)	(3.90)	(3.78)	(3.61)
Cumulativeness x Innovative Entrant			-0.77	-0.96	-1.32	-1.38
			(1.96)	(1.96)	(1.98)	(1.94)
Opportunity x Nonrelated Diversifier			0.02	0.02	0.02	0.02
			(0.01)	(0.02)	(0.02)	(0.02)
Opportunity x Related Diversifier			0.03	0.04	0.05	0.04
			(0.03)	(0.03)	(0.03)	(0.03)
Opportunity x Innovative Entrant			0.01	0.01	0.01	0.01
			(0.02)	(0.02)	(0.02)	(0.02)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Wald Test	228.98***	250.24***	296.98***	312.73***	315.57***	291.20***

Standard errors are clustered by industry-year. This applies to all results in this study.
* $p < 0.1$
** $p < 0.05$
*** $p < 0.01$

TABLE 5
Summary of Table A1, Model 1, interaction effects

	Coefficient
Appropriability	0.28
	(0.18)
Cumulativeness	-0.24
	(0.89)
Opportunity	-0.11*
	(0.06)
Appropriability x Cumulativeness	-1.55*
	(0.90)
Appropriability x Opportunity	-0.05**
	(0.02)
Cumulativeness x Opportunity	0.65
	(0.41)

** p < 0.1*
*** p < 0.05*
**** p < 0.01*

TABLE 6
Summary of Table A1, Model 1, by entrant type

		Change in Effect For				
	Noninnovative De Novo	Innovative De Novo	Innovative Nonrelated Diversifier	Noninnovative Nonrelated Diversifier	Innovative Related Diversifier	Noninnovative Related Diversifier
Appropriability	0.28	-0.20	0.52	0.06	7.40***	0.20
Cumulativeness	-0.24	-0.10	-6.13	-0.07	43.32***	3.30
Opportunity	-0.11*	0.03	-0.06	0.02	0.36***	0.03

** p < 0.1*
*** p < 0.05*
**** p < 0.01*

TABLE 7

Summary of Table A2, Model 1

		Change in Effect For		
	Noninnovative De Novo	Nonrelated Diversifier	Related Diversifier	Innovative Entrant
Appropriability	0.45**	-0.96	-4.72***	0.22
Cumulativeness	-0.47	1.82	-2.24	-3.09
Opportunity	-0.14**	0.22	-0.35	-0.20
Appropriability x Cumulativeness	-2.39**	5.46	33.57***	-1.57
Appropriability x Opportunity	-0.08***	0.13**	0.09	-0.02
Cumulativeness x Opportunity	0.85*	-1.43	3.03	1.63

** p < 0.1*
*** p < 0.05*
**** p < 0.01*

APPENDIX
TABLE A1
Full results for finer categories of entrant types

	Model 1	**Model 2**	**Model 3**
Appropriability	0.28	0.29	0.19
	(0.18)	(0.19)	(0.22)
Cumulativeness	-0.24	-0.40	-0.11
	(0.89)	(0.91)	(0.99)
Opportunity	-0.11*	-0.13**	-0.16**
	(0.06)	(0.06)	(0.07)
Innovative Nonrelated Diversifier	0.75	0.84	0.64
	(0.82)	(0.82)	(0.82)
Noninnovative Nonrelated Diversifier	-0.45*	-0.46*	-0.54**
	(0.24)	(0.25)	(0.25)
Innovative Related Diversifier	-6.23***	-6.24***	-6.35***
	(1.50)	(1.50)	(1.51)
Noninnovative Related Diversifier	-1.07*	-1.08**	-1.29**
	(0.56)	(0.55)	(0.53)
Innovative De Novo Entrant	0.18	0.23	0.23
	(0.31)	(0.31)	(0.31)
Appropriability x Cumulativeness	-1.55*	-1.56*	-1.19
	(0.90)	(0.94)	(1.07)
Appropriability x Opportunity	-0.05**	-0.06**	-0.04*
	(0.02)	(0.03)	(0.03)
Cumulativeness x Opportunity	0.65	0.81*	1.03*
	(0.41)	(0.43)	(0.55)
Appropriability x Innovative Nonrelated Diversifier	0.52	0.52	0.53
	(0.52)	(0.52)	(0.52)
Appropriability x Noninnovative Nonrelated Diversifier	0.06	0.03	0.04
	(0.10)	(0.11)	(0.11)
Appropriability x Innovative Related Diversifier	7.40***	7.40***	7.54***
	(2.64)	(2.63)	(2.64)
Appropriability x Noninnovative Related Diversifier	0.20	0.43	0.19
	(0.17)	(0.27)	(0.18)
Appropriability x Innovative De Novo Entrant	-0.20	-0.17	-0.17
	(0.15)	(0.14)	(0.15)
Cumulativeness x Innovative Nonrelated Diversifier	-6.13	-6.86	-5.41
	(6.06)	(6.14)	(6.11)
Cumulativeness x Noninnovative Nonrelated Diversifier	-0.07	-0.04	0.53
	(1.71)	(1.75)	(1.76)
Cumulativeness x Innovative Related Diversifier	43.32***	43.49***	44.21***
	(8.97)	(8.98)	(8.97)
Cumulativeness x Noninnovative Related Diversifier	3.30	3.15	5.19
	(3.94)	(3.81)	(3.66)
Cumulativeness x Innovative De Novo Entrant	-0.10	-0.41	-0.57
	(2.21)	(2.23)	(2.21)
Opportunity x Innovative Nonrelated Diversifier	-0.06	-0.06	-0.06
	(0.05)	(0.05)	(0.05)
Opportunity x Noninnovative Nonrelated Diversifier	0.02	0.02	0.02
	(0.01)	(0.01)	(0.01)
Opportunity x Innovative Related Diversifier	0.36***	0.36***	0.36***
	(0.10)	(0.10)	(0.10)
Opportunity x Noninnovative Related Diversifier	0.03	0.04	0.03
	(0.04)	(0.04)	(0.04)
Opportunity x Innovative De Novo Entrant	0.03	0.03	0.03
	(0.03)	(0.03)	(0.02)
Controls	Yes	Yes	Yes
Wald Test	786.10***	804.17***	772.61***

** p < 0.1*
*** p < 0.05*
**** p < 0.01*

TABLE A2
Full results for interaction of technological regime variables with entrant type

	Model 1	Model 2	Model 3
Appropriability	0.45**	0.46**	0.35
	(0.19)	(0.20)	(0.23)
Cumulativeness	-0.47	-0.63	-0.37
	(0.94)	(0.95)	(1.04)
Opportunity	-0.14**	-0.16**	-0.19**
	(0.07)	(0.07)	(0.09)
Nonrelated Diversifier	-0.72**	-0.74**	-0.85***
	(0.30)	(0.31)	(0.30)
Related Diversifier	-0.40	-0.48	-0.76
	(0.63)	(0.64)	(0.63)
Innovative Entrant	0.63	0.78*	0.59
	(0.44)	(0.46)	(0.44)
Appropriability x Cumulativeness	-2.39**	-2.42**	-1.94*
	(0.96)	(0.97)	(1.14)
Appropriability x Opportunity	-0.08***	-0.09**	-0.07**
	(0.03)	(0.04)	(0.03)
Cumulativeness x Opportunity	0.85*	1.01*	1.26*
	(0.51)	(0.52)	(0.66)
Appropriability x Nonrelated Diversifier	-0.96	-1.00	-0.51
	(0.63)	(0.66)	(0.68)
Appropriability x Related Diversifier	-4.72***	-5.30***	-3.93***
	(1.13)	(0.98)	(1.04)
Appropriability x Innovative Entrant	0.22	0.64	0.22
	(1.24)	(1.28)	(1.28)
Cumulativeness x Nonrelated Diversifier	1.82	1.92	2.81
	(2.16)	(2.24)	(2.17)
Cumulativeness x Related Diversifier	-2.24	-1.94	0.71
	(4.51)	(4.61)	(4.50)
Cumulativeness x Innovative Entrant	-3.09	-4.06	-2.93
	(3.10)	(3.25)	(3.12)
Opportunity x Nonrelated Diversifier	0.22	0.25	0.26*
	(0.14)	(0.15)	(0.14)
Opportunity x Related Diversifier	-0.35	-0.24	-0.26
	(0.27)	(0.25)	(0.14)
Opportunity x Innovative Entrant	-0.20	-0.24	-0.14
	(0.14)	(0.15)	(0.15)
Appropriability x Cumulativeness x Nonrelated Diversifier	5.46	5.57	2.18
	(4.05)	(4.26)	(4.36)
Appropriability x Cumulativeness x Related Diversifier	33.57***	37.60***	27.79***
	(7.38)	(6.24)	(6.64)
Appropriability x Cumulativeness x Innovative Entrant	-1.57	-3.63	-1.55
	(6.07)	(6.25)	(6.25)
Appropriability x Opportunity x Nonrelated Diversifier	0.13**	0.15**	0.13**
	(0.05)	(0.07)	(0.06)
Appropriability x Opportunity x Related Diversifier	0.09	0.26***	0.09
	(0.08)	(0.07)	(0.09)
Appropriability x Opportunity x Innovative Entrant	-0.02	-0.04	-0.02
	(0.08)	(0.08)	(0.08)
Cumulativeness x Opportunity x Nonrelated Diversifier	-1.43	-1.60	-1.70
	(1.11)	(1.17)	(1.10)
Cumulativeness x Opportunity x Related Diversifier	3.03	2.41	2.39
	(2.06)	(1.92)	(1.96)
Cumulativeness x Opportunity x Innovative Entrant	1.63	1.94*	1.17
	(1.07)	(1.10)	(0.06)
Controls	Yes	Yes	Yes
Wald Test	572.40***	668.20***	634.53***

** p < 0.1*
*** p < 0.05*
**** p < 0.01*

Where do firms come from? Knowledge relatedness and firm entry

ABSTRACT

The role of knowledge, in particular its relatedness to that of the target industry, is often emphasized as one of the main drivers in firms' entry decisions. However, empirical analyses seem to show that knowledge relatedness may not always have a positive impact on entry, because the relationship is potentially nonlinear and may vary by entrant type. This study aims to explore these issues by proposing a perspective that may shed light on the mechanisms that govern the relationship between knowledge relatedness and firm entry. It then empirically tests these mechanisms by using a dataset constructed from the Western Electronics Manufacturer Association (WEMA) directory, spanning a time period from 1960 to 1990. The results suggest that the relationship takes an inverted U shape, which however can change drastically depending on the type of entrant.

Keywords:

Firm Entry; Knowledge Relatedness; Entrant Heterogeneity

Where do firms come from? Knowledge relatedness and firm entry

INTRODUCTION

There is strong evidence that knowledge plays a major role in firms' decision to enter an industry, whether that firm is a spinout, spinoff, or diversifier (Helfat & Lieberman, 2002; Agarwal & Shah, 2014; Adams, Fontana, & Malerba, 2015). In particular, knowledge relatedness between the firm and the industry is generally thought be a key driver of firm entry, the idea being that higher knowledge relatedness means that the potential entrant has a better match with the target industry which in turn leads to better expected performance (Helfat & Lieberman, 2002; Helfat & Eisenhardt, 2004; Uzunca, 2011). However, empirical evidence calls this into question, suggesting that the mechanisms according to which knowledge relatedness affects firm entry seem more nuanced than previously assumed, especially when different types of entrants are considered (Klepper & Sleeper, 2005; Adams, Fontana, & Malerba, 2015; Sakakibara & Balasubramanian, 2020). This study provides a theoretical framework to explain how higher relatedness may not necessarily translate into a higher rate of entry and examines this relationship by type of entrant. By doing so, it is able to reconcile some of the empirical contradictions. This theoretical framework is tested by using a dataset that tracks the US electronics industry from 1960 to 1990.

The general consensus on how knowledge relatedness affects the rate of entry is that higher knowledge relatedness leads to a higher rate of entry, a point which is proposed in earlier studies (Helfat & Lieberman, 2002; Klepper & Sleeper, 2005; Qian, Agarwal, & Hoetker, 2012). The reasoning underlying this consensus is based on the idea that entry is mainly driven by the proximity between the knowledge base of the potential entrant and the knowledge required to successfully compete in the target industry. Higher knowledge relatedness means that potential entrants have more useful knowledge that can be utilized in the target industry. This would imply that they can generally expect better profitability, which is one of the key criteria for firms' entry decisions (Chatterjee & Wernerfelt, 1991; Helfat & Lieberman, 2002; Klepper & Sleeper, 2005).

However, the actual empirical evidence is surprisingly mixed. There are certainly studies that support the idea that higher knowledge relatedness leads to higher rates of firm entry, such as Chang (1996) or Klepper & Sleeper (2005). However, there is also other evidence, usually from more recent studies that look at the entry of spinoffs and spinouts[14] rather than diversifiers, that do not support this idea (Sapienza, Parhankangas, & Autio, 2004; Clarysse, Wright, & Van de Velde, 2011; Sakakibara & Balasubramanian, 2020). Some of them, such as Adams et al. (2015), imply that higher relatedness may not necessarily lead to higher rate of entry, or at least not monotonically so. One common thread in these analyses is that heterogeneity across entrant types may be such that different types of entrants (i.e., diversifiers, spinoffs, spinouts, and startups) do not necessarily respond to knowledge relatedness in the same way. Explaining how these differences arise is key to identifying how knowledge relatedness actually affects entry. This, in turn, would also have repercussions for what kind of competition incumbents may ultimately expect from entrants.

The main aim of this study is to provide a theoretical framework to explain how knowledge relatedness affects the entry of firms (which can be spinouts, spinoffs, or diversifiers) and to test it empirically. The theoretical framework starts by defining relatedness in terms of two components: industry-specific and general-purpose knowledge. These two components represent the building blocks on which the mechanisms through which knowledge relatedness influences firms' entry decisions will be built. The first mechanism describes the knowledge effect, which says that higher knowledge relatedness implies a larger stock of industry-

[14] In this study, spinoffs refer to those entrants that are also called within-industry spinouts, while spinouts refer to those that are also called out-of-industry spinouts.

specific knowledge that would tend to encourage entry. The other is the differentiation effect, which says that high knowledge relatedness also means that the potential entrant will tend to have trouble differentiating itself with the incumbents in the target industry, tending to discourage entry instead. We then rely on the same framework to show how these mechanisms can also explain how relatedness may affect the various types of entrants differently. We finally test these theoretical assertions by using a dataset derived from the Western Electronics Manufacturers Association (WEMA) directory, which provides information on firms in the US electronics industry from 1960 to 1990.

The paper is structured as follows. We first review the literature on the role of knowledge in firm entry. Then, we theorize on the relationship between knowledge relatedness and the rate of entry. We focus, in particular, on how the relationship may be nonlinear and can vary across different types of entrants. Next, we describe the dataset used and the empirical method through which we verify the relationship and test our hypotheses. Finally, we discuss the implications from the results.

LITERATURE REVIEW

Firm entry

Firm entry is one of the main factors affecting market structure, which in turn affects the performance of firms (Christensen & Montgomery, 1981; Gort & Klepper, 1982; Dunne, Roberts, & Samuelson, 1988; Klepper & Simons, 2000; Agarwal, Sarkar, & Echambadi, 2002; Geroski, Mata, & Portugal, 2010; Pe'er, Vertinsky, & Keil, 2016). It is therefore unsurprising that there has been consistent interest in trying to determine the patterns of entry and the mechanisms behind firms' decisions to enter an industry, particularly from an evolutionary perspective (Klepper, 1996; Klepper & Sleeper, 2005; Dencker, Gruber, & Shah, 2009; Cattani & Malerba, 2021).

Early contributions in the field started out by comparing patterns of entry over time and across industries, in an effort to understand what factors may affect firm entry (Gort & Klepper, 1982; Klepper & Graddy, 1990). The identification of a general pattern was later developed into the concept of the product (Abernathy & Utterback, 1978), and then of the industry life cycle (Agarwal & Gort, 1996).

Later works hinged on the point that that not all entrants are similar in nature (Klepper & Simons, 2000; Helfat & Lieberman, 2002; Uzunca, 2011; Benner & Tripsas, 2012). This heterogeneity, in turn, is often thought to arise from the variety of sources of entry, based on the idea that entrants do not simply appear out of thin air. The origin of the potential entrant is thought to determine the knowledge it has, which would then determine its entry mode and its expected post-entry performance.

Role of knowledge in firm entry

Pre-entry knowledge was identified as one of the key drivers of entry, but it is also true that not all knowledge might be valuable for potential entrants (Klepper & Simons, 2000; Helfat & Lieberman, 2002; Bayus & Agarwal, 2007). While some types of knowledge are thought to be more general in their usefulness than others (Helfat & Lieberman, 2002; Pisano, 2017), some industries demand very specific knowledge, or possibly combinations of different types of knowledge. In other words, possessing a stock of knowledge is important, but having the relevant type of knowledge is even more so.

Knowledge relatedness is the concept that has been employed to capture the relevance of knowledge (Sapienza et al., 2004; Tanriverdi & Venkatraman, 2005; Benner & Tripsas, 2012; Chang, Eggers, & Keum, 2021). In its simplest form, high knowledge relatedness essentially means that the potential entrant has a high stock of knowledge that is relevant to the industry in question. Despite the apparent simplicity of the concept, however, it turned out that defining exactly what "relevant" knowledge means is harder than expected.

The straightforward approach to the implementation of the concept is in terms of the overlap or similarity of the knowledge between firms and industries. This has indeed been the most common approach followed, regardless of the context (Ahuja & Katila, 2001). While this approach is simple and intuitive, however, some scholars have argued that a different aspect, complementarity of knowledge, should be considered, since in some cases, knowledge only has value when combined with others (Makri, Hitt, & Lane, 2010). Even more importantly, even if a potential entrant does not possess a specific piece of knowledge, some of its knowledge may still be applicable, such as its knowledge in effectively managing R&D (Pisano, 2007).

The overarching theme in all these conceptualizations of knowledge relatedness seems to be that relatedness measures the extent to which a firm's knowledge might be useful in one way or another to compete in an industry. In other words, higher knowledge relatedness should imply that firms have more knowledge that is applicable in the industry in question. Some studies state this explicitly, noting that higher knowledge relatedness implies that the knowledge resources are more applicable in the industry (Chatterjee & Wernerfelt, 1991; Helfat & Lieberman, 2002; Klepper & Sleeper, 2005).

Another interesting aspect to knowledge relatedness is that while its importance in entry tends to be undisputed, exactly how it is significant remains to be explored. One's immediate answer might be that higher relatedness should lead to higher rate of entry, if only because the potential entrant can expect to have more knowledge that it can use when operating in the industry in question. This, indeed, is supported by a substantial number of studies in the literature, such as Klepper & Sleeper (2005) or Adams et al. (2015).

At the same time, however, an opposite tendency is identified, or at least implied, in empirical analyses, particularly in the context of entry by spinouts (Clarysse, Wright, & Van de Velde, 2011; Sakakibara & Balasubramanian, 2020). Sakakibara & Balasubramanian (2020), for example, shows that under certain conditions, spinouts distance themselves from their parents, in an effort to shield themselves from competitive pressure. In other words, despite having, presumably, very high knowledge relatedness with the industry, the threat of competition could actually lead firms to avoid entering such industries. While these studies focus on the parent firms rather than the target industry, they still provide valuable insights into how knowledge relatedness may affect the entry decision of potential entrants. In particular, they suggest the possibility that not all types of entrants necessarily respond to knowledge relatedness in the same way, emphasizing the need to consider different entrant types and knowledge relatedness.

Knowledge and entrant types

The discussion above also highlights a particularly interesting aspect of entrants: how diverse they actually are. In an evolutionary perspective, extant studies show the necessity of considering these differences more in depth, especially given the importance of pre-entry knowledge (Helfat & Lieberman, 2002; Klepper & Sleeper, 2005; Gruber, MacMillan, & Thompson, 2013). One of the first distinguishing factors to be examined has been whether entrants were actually new firms (i.e., independent ‘de novo’ entities) or previously established firms coming from a different industry (i.e., diversifiers). While the two types of entrants were often analyzed separately, more recent studies tend to compare the two, emphasizing that they mainly differ in terms of the stock of knowledge each type tends to have.

The differences among the de novo entrants, in particular, are gaining more and more attention (Helfat & Lieberman, 2002; Adams, Fontana, & Malerba, 2015). Inter-industry spinouts refer to new firms established by founders who were formerly employees of an existing company in a different industry[15]. Intra-industry spinoffs, on the other hand, are new firms established by founders coming from firms in the same industry[16]. In line with the concept of pre-entry knowledge, the idea is that such entrants tend to perform better than other de novo entrants (i.e., startups without any industry-specific knowledge).

Organic growth through diversification can also be seen as a form of entry, by firms that have their own set of pre-entry knowledge. Their importance is highlighted not only in the diversification literature, but also the entry literature in general (Klepper & Simons, 2000). Diversifiers are generally thought to have an inherent advantage over de-novo entrants, in that they have the stock of knowledge they have already accumulated over time. However, even within diversifiers, the firms' origin matters, as evidenced by the analyses on related vs. unrelated diversification (Chang, 1996; Benner & Tripsas, 2012; Qian et al., 2012).

THEORETICAL FRAMEWORK

Relatedness and entry

While the actual definition of knowledge relatedness varies across studies, one common aspect is the understanding of relatedness in terms of the utility of the knowledge the firm has in relation to the industry. In this regard, applying the categorization of knowledge proposed by Pisano (2017) to classify knowledge into industry-specific and general-purpose knowledge provides a useful framework. The former refers to knowledge that is applicable in a specific

---

[15] These firms are also called within-industry spinouts in some studies (e.g. Sakakibara & Balasubramanian, 2020)

[16] These firms are also called out-of-industry spinouts (e.g. Sakakibara & Balasubramanian, 2020)

industry, while the latter is applicable in a wider range of activities (Pisano, 2017). The latter, in particular, shares some similarities with the notion of general-purpose technology[17] (Rosenberg, 1963; Bresnahan & Trajtenberg, 1995; Bresnahan & Gambardella, 1998; Conti, Gambardella, & Novelli, 2019), except that 'general-purpose' in the GPT context refers to its applicability to several industries and/or sectors. General-purpose knowledge here, on the other hand, is related to the usefulness of knowledge in tackling distinct problems in different contexts.

It should be noted that the distinction between industry-specific and general-purpose knowledge is not a discrete one. In other words, the two types of knowledge are representations of two extremes of a continuum. The reality is that most knowledge will not actually be of either extreme but rather fall somewhere in between, generally not even staying constant over time. For example, the knowledge of the manufacturing process of a computer would obviously be the most applicable in the computer manufacturing industry, yet this could also be useful in other industries such as pre-packaged software as different hardware architectures can potentially affect the optimal configurations for different software.

The second point to note, which arises from the first, is that the framework works best when applied in reference to a specific industry. For knowledge to be industry-specific or general-purpose, it needs to be relevant in the industry in the first place. The only difference is the degree to which the knowledge is also applicable in other industries, not whether the knowledge is useful at all.

In this paper, we build upon this distinction to define knowledge relatedness as the average similarity of knowledge between the potential entrant i (which may be a diversifier, spinout, or spinoff) and the target industry j that they may enter, weighted by the degree of specificity

[17] This will be referred to as GPT hereafter.

to the target industry (see equation 1 below). This conceptualization is based on the notion that given a similar level of similarity, the knowledge that is more specific to the target industry is a better indication of the knowledge relatedness between the potential entrant and the industry, which is often implied in previous studies on knowledge relatedness (Chatterjee & Wernerfelt, 1991; Helfat & Lieberman, 2002). Note that the term "industry specificity" here refers to the target industry (i.e., the industry the potential entrant might enter), and not to the industry of origin (i.e., the industry the potential entrant originates from, which for spinouts and spinoffs means the industry of the firm the founder originally worked in and for diversifiers means the industry their parent company is operating in).

$$Knowledge\ Relatedness_{ij} = \frac{1}{K}\sum_{k} S_j(Knowledge_i) \cdot \left(Industry\ Specificity_j\right) \qquad (1)$$

Here, $S_j(Knowledge_i)$ represents the similarity of the unit of knowledge held by potential entrant i to the knowledge of target industry j, $Industry\ Specificity_j$ refers to the degree at which the knowledge in question is specific to the target industry j, and $k = 1, \ldots, K$ is the stock of knowledge the potential entrant has. In summary, the equation defines knowledge relatedness as the average similarity of the knowledge stock of the potential entrant relative to the target industry it may enter, but with the knowledge that is specific to that target industry contributing more significantly.

Essentially, knowledge relatedness as defined by equation 1 is comprised of two main parts. The first is the similarity of knowledge held by the potential entrant relative to the target industry it may enter. This is the part that is more consistent with what is usually understood as knowledge relatedness, in that many studies often consider the knowledge relatedness to be synonymous to knowledge similarity. The second is the element of specificity, which draws on the industry-specific and general-purpose knowledge concept based on Pisano (2017). The idea is that while sharing knowledge that is more general-purpose still counts toward higher similarity in knowledge to the target industry, they are not as significant in the firms' entry decisions because they can also be applied in other industries, giving them the choice to not enter the target industry in question.

Some readers may notice that this definition is somewhat different from the usual definitions of knowledge relatedness, especially in the sense that industry specificity is involved. This definition is chosen in this study for three reasons. First, it helps make the concept of knowledge relatedness applicable across different types of entrants, in this case spinouts, spinoffs, and diversifiers. In many studies, knowledge relatedness is applied in the context of diversification, where the main benefit of entering industries with which the diversifier has high relatedness, is thought to be the potential for economies of scope. This is a particularly common approach in the context of diversification, where the measure is often based on firms' diversification patterns, Bryce & Winter (2009) being one of the best examples. However, this would be less useful for spinouts, since economies of scope would not apply to spinouts given that they usually do not start out in multiple industries. The current definition is more generally applicable, especially since the main mechanisms associated with the two types of knowledge do not need to exclusively rely on the economies of scope dimension.

Second, it is still sufficiently consistent with the previous definitions of knowledge relatedness in previous studies. One common conceptualization of knowledge relatedness is the similarity in knowledge (Sapienza et al., 2004), which in this study would be between the potential entrant (which can be a diversifier, spinout, or spinoff), and the industry that they are about to enter. The definition in Equation 1 is essentially based on the similarity of knowledge as well, with an additional layer suggesting that some types of knowledge contribute more to the similarity in knowledge than others. This is partly inspired by

Chatterjee & Wernerfelt (1991), where firms' tendency for related diversification is linked to the specificity of resources.

Finally, it is also original, in that by incorporating the framework in Pisano (2017), the definition allows us to identify two main mechanisms through which knowledge relatedness may affect firms' entry decisions into a specific target industry. As outlined in Pisano (2017), the two types of knowledge play a very different role in firms' performance, which in turn will be the main factor in firms' decision to enter a focal industry. The main idea is that both types of knowledge contribute to the competitive advantage of the potential entrant, but in very different ways.

The knowledge effect. On the one hand, higher knowledge relatedness implies a higher share of (target) industry-specific knowledge relative to general-purpose knowledge. This means that the potential entrant has more knowledge that can be immediately deployed in the industry it is targeting[18]. In turn, this implies that a potential entrant can expect to survive better and be more profitable. Therefore, higher relatedness encourages entry (Helfat & Lieberman, 2002; Bayus & Agarwal, 2007). In addition, potential entrants with a higher stock of industry-specific knowledge may exhibit overconfidence in their ability to perform well, especially in relation to other industries that they might consider. This is partly supported by the GPT literature, which shows that the more pervasive and 'general' the GPT is, the more likely it is that general specialization will occur (Bresnahan & Gambardella, 1998). In our context, the interpretation may be that the more 'general' the knowledge of the firm is, the more deployable it becomes, which implies that the spectrum of industries the firm might

[18] This does not mean that general-purpose knowledge cannot be deployed in the target industry. It means, however, that in some cases generality may come at the cost of specialization, hence limiting its value in specific contexts (Bresnahan & Gambardella, 1998).

potentially enter is wider (i.e., the less likely it is that the firm will enter the specific target industry), as it will aim to keep the knowledge relevant to multiple industries at once.

The differentiation effect. On the other hand, higher knowledge relatedness also implies that, in our framework, the potential entrants have relatively little general-purpose knowledge. While higher knowledge similarity in the traditional sense tends to be harmful for the entrants' ability to innovate and differentiate themselves form the incumbents (Katila, 2002), this problem is exacerbated when they also do not have a sufficient level of general-purpose knowledge. This is because general-purpose knowledge would be helpful in the entrant's innovation and product differentiation, since general-purpose knowledge would tend to lead to a broader knowledge domain for the entrant, thus helping it to innovate better and in a more radical fashion.

The literature on the relationship between parents and their spin-offs or spinouts has already highlighted this implication (Sapienza et al., 2004; Clarysse, Wright, & Van de Velde, 2011; Sakakibara & Balasubramanian, 2020). The framework of industry-specific and general-purpose knowledge helps us extend this argument in a more cohesive manner. Specifically, the higher knowledge relatedness implies that the applicability of the potential entrant's knowledge is more limited to the target industry, thus making it difficult for it to innovate except in a very incremental manner, in which the incumbents are usually much more adept at. In other words, entrants with higher knowledge relatedness are more vulnerable to competition from incumbents, since they will find it more difficult to innovate in a way that allows them to significantly differentiate their products from the existing ones in the target industry. This would have negative consequences for profitability, thus discouraging entry.[19]

[19] This reasoning also stems from the innovation complementarities, as emphasized in the GPT literature (Bresnahan & Trajtenberg, 1995; Bresnahan & Gambardella, 1998), based on the following parallel that can be drawn. While advances in GPT trigger advances in individual application sectors, possessing general-purpose knowledge should help the single potential entrant better use its industry-specific knowledge and innovate. One

The two effects combined. Before going into the discussion of the changes in the relationship between knowledge relatedness and the rate of entry, along with discussions on the two effects, Figure 1 provides a visualization of some of the descriptions of these changes. This figure should be useful in understanding the discussion.

[INSERT FIGURE 1 ABOUT HERE]

As the two mechanisms work in opposite directions, their combined effect on the rate of entry becomes an empirical matter, in the sense that how the marginal effect of each mechanism changes with relatedness becomes a key factor in the actual relationship between knowledge relatedness and firm entry. Table 1 below, and the discussion in the Appendix, summarizes the possible outcomes, assuming a continuous and monotonic trend for the total and marginal effects of both mechanisms, and the existence of some level of relatedness at which the marginal effects cancel out[20].

[INSERT TABLE 1 ABOUT HERE]

While, in principle, all outcomes summarized in Table 1 are possible, some of them are more likely than others. Specifically, for the marginal knowledge effect to increase with relatedness, it is required that there exists a high degree of complementarity among the units of knowledge. While this may indeed be the case in industries in which knowledge is highly integrated, in most industries there is some degree of modularity in knowledge.

This leads to the initial hypothesis that *the relationship between relatedness and entry should display an inverted-U shape*. Thus, while Table 1 theoretically allows for a U-shaped

---

limit of this parallel is that the GPT framework focuses on technological change at the sector level, while our focus is at the firm level. Therefore, if the knowledge relatedness of the potential entrant with respect to the target industry is too high, the potential entrant will end up relying on incremental innovations at best, which is generally something incumbents are better at.

[20] In other words, the magnitude of the positive effect of having more deployable knowledge and negative effect of the difficulty in differentiation is always assumed to be nondecreasing with knowledge relatedness.

relationship to emerge, an inverted-U shape seems to be the more likely outcome as long as the marginal knowledge effect is not increasing with relatedness. For a more detailed discussion on how the overall impact of rate on entry can take an inverted U-shape based on the two mechanisms, see the Appendix.

Relatedness and entrant types

We focus on three types of entrants: diversifiers, spinouts, and spinoffs. Diversifiers are incumbents of another industry who decide to expand into the target industry (Helfat & Lieberman, 2002). One of the major characteristics of diversifiers compared to the other types of entrants is that they tend to be rich in knowledge resources (Helfat & Lieberman, 2002; Agarwal, Echambadi, Franco, & Sarkar, 2004; Bayus & Agarwal, 2007; Ganco & Agarwal, 2009; Chen, Williams, & Agarwal, 2012). This means that, given the same level of relatedness, diversifiers tend to have higher stocks of both general-purpose and industry-specific knowledge compared to the other types of entrants.

Spinouts are entrants whose founders previously worked in a firm in another industry before establishing their own independent firm.[21] They share similarities with diversifiers in that they need to "bridge" knowledge across industry boundaries in order to enter the target industry. They are also very similar to de novo entrants, in that they are new firms whose defining characteristic is the fact that they have some stock of knowledge inherited from the parent via their founders (Adams, Fontana, & Malerba, 2015).

[21] In the literature, the term 'spinout' generally refers to 'employee spinout'. Examples of such usage can be found in past papers such as Agarwal et al. (2004), Agarwal & Shah (2009), and Sakakibara & Balasubramanian (2020).

Spinoffs are similar to spinouts, except that the founder comes from a company that is active in the same industry as the one they enter. They are, in some sense, a special case of spinouts, but with a higher relatedness to the industry they enter than a typical spinout.

The three types of entrants differ in terms of the stock and combination of industry-specific and general-purpose knowledge, as summarized in Table 2 below. Each cell denotes the range which the respective type of knowledge can take, including theoretical extremes. For example, diversifiers could theoretically have any level of relatedness between 0 and 1, which is why the range of the general-purpose knowledge is taken to be from 0 to 100. Of course, it is unlikely that a diversifier will have a very low value for general-purpose knowledge. What the values here indicate is that diversifiers can potentially have a low level of relatedness with the target industry.

[INSERT TABLE 2 ABOUT HERE]

Diversifiers vs spinouts. One of the principal differences between diversifiers and the other types of entrants would be that diversifiers tend to have a larger stock of knowledge than spinouts (Helfat & Lieberman, 2002; Bayus & Agarwal, 2007; Ganco & Agarwal, 2009; Chen et al., 2012). This would mean that, given the same level of relatedness, the diversifier would tend to have a larger stock of both industry-specific and general-purpose knowledge. If this is indeed the case, the interplay between the knowledge effect and the differentiation effect would work as follows.

First, diversifiers may have a larger stock of industry-specific knowledge compared to a spinout with the same level of relatedness (Helfat and Lieberman, 2002). In other words, they may have more knowledge that can be deployed in the industry, which would help them achieve better profitability compared to other types of entrants, such as spinouts or spinoffs. This means that the knowledge effect will tend to be larger for diversifiers, which will in turn lead diversifiers to show a higher rate of entry compared to spinouts or spinoffs with the same level of relatedness.

Second, the larger stock of general-purpose knowledge can help diversifiers differentiate better from incumbents in the target industry. In other words, the marginal differentiation effect would tend to be smaller for diversifiers. As noted in previous studies such as Pisano (2017), general-purpose knowledge can be a valuable complement to industry-specific knowledge and allows firms to combine them in various ways. Given this potentially greater degree of variety, diversifiers may tend to find it easier to differentiate their products, and thus be more protected from competition. This implies that the marginal differentiation effect will tend to vary less compared to spinouts, causing the relationship between the differentiation effect and relatedness to become more linear overall.

In summary, compared to spinouts, the rate of entry by diversifiers would tend to be higher, leading to a higher proportion of diversifiers among entrants. In addition, the relationship between relatedness and rate of entry is expected to be flatter for diversifiers.

Spinoffs vs spinouts. The differences between spinoffs and spinouts, conceptually, are harder to define, and there is relatively little insight on the subject in the literature aside from a few exceptions. This is partly because most of the literature tends to focus more on intra-industry spinoffs, although some do make the distinction between "insiders" and "outsiders" (Agarwal & Shah, 2014; Adams et al., 2015; Sakakibara & Balasubramanian, 2020). Those that do, however, provide some reasons why potential entrants that ultimately become spinouts may be different from those that become spinoffs.

One source of diversity may be the extent to which potential entrants actually benefit from the level of knowledge relatedness with a potential target industry. The marginal benefit of having industry-specific knowledge may tend to be larger for spinoffs than spinouts, if they are better able to exploit any existing complementarity in knowledge.[22] More precisely, we may suppose that spinoffs are those that actually can benefit more from any existing complementarity of the knowledge compared to spinouts, which would tend to cause the relationship between the knowledge effect and relatedness to become more convex.

Another possibility is that spinoffs might have better knowledge to serve possible technological niches within the industry, or at least believe they do. This is in line with the results from Sakakibara & Balasubramanian (2020), where the tendency to generate spinoffs over spinouts are found to be dependent on the human capital the founders have. In other words, compared to spinouts, the differentiation effect may tend to be smaller when entering their industry of origin. At the same time, the same impact may tend to be perceived as being relatively larger with respect to other possible target industries, causing spinoffs to avoid entering less related industries more than spinouts.

As a result, the relationship between knowledge relatedness and the rate of entry is expected to be flatter for spinoffs.

EMPIRICAL ANALYSIS

The creation and processing of the database

We test the implications of our theory on a large sample of firms in the US electronics industry that are listed in a series of directories published annually by the Western Electronics Manufacturers Association (WEMA), from 1960 to 1990. The dataset has been purposedly built for the objective of this thesis from these directories, whose publisher, WEMA, is an organization that collects the membership of the companies that operated in the industry since

[22] This is similar to how spinouts in general are thought to be better at searching for necessary complementary knowledge as described in Agarwal et al. (2004).

its inception.[23] The original data is only available in paper form, so the initial steps of data preparation involved scanning the directories page by page, followed by an OCR (Optical Character Recognition) process to extract the text into an Excel format. This process allowed us access to information such as company name, location, year of establishment, number of employees, companies' products, as well as information on executive and board members.

In some cases, the OCR process returned incomplete results and/or results riddled with typos, partly because not all scans were ideal for OCR. For example, a common occurrence was a situation where an "I" might be read as "/" by the OCR. Thus, manual cleaning was performed, for which we generally took the approach of assigning IDs for disambiguation purposes based on name, state, and town of each observation. Even if the companies appeared to have different names, they were assigned the same ID if the names were only different due to typos from OCR, formatting issues, etc. This cleaning was partly assisted by an automated process, where companies whose names were sufficiently close in terms of the Levenshtein distance[24] or in which one company's name was wholly contained in another were initially considered to be the same companies, subject to manual scrutiny later.

The WEMA directory was also employed to initially identify the type of entrant. Specifically, the directory explicitly states whether a company is a subsidiary or division of an existing company, often by listing the company as "XXX, subsidiary of YYY". However, the directory is also somewhat inconsistent in how it chooses to disclose this information, as sometimes the subsidiaries and divisions are simply listed immediately after the company to which they are associated, usually in a smaller font. Since this aspect tended to be lost in the

[23] We thank Leslie Berlin, the Project Historian for the Silicon Valley Archives at Stanford University, for pointing to our attention the existence of this invaluable source.

[24] This is defined as the minimum number of replacements, removals, or additions of letters required to transform one word into another. For the purposes of our study, we divide this value by the larger of the number of letters in the two words involved.

OCR process, further human intervention was required to confirm whether the company is a division or subsidiary of another.

In order to obtain the information on the companies' founders and the associated parent companies, some extensive searching on the web was required. We started off with the company name and the top executives' names to identify the founders by searching for them across various sources. This involved extensive web search, which led to certain sources such as newspapers.com, where founders were sometimes mentioned in association with the company in newspaper articles, obituaries, etc. This also provided information on where the founders originally worked, which helped us identify the parent company. Sources such as ORBIS were also extensively used in order to determine the SIC of the companies and the associated parent companies[25].

Information on the SIC codes are particularly important for our purpose, as determining the knowledge relatedness between the company and a target industry, in addition to whether the company actually entered that industry, relies heavily on identifying the SIC of both the company and the parent company. However, the identification of a unique SIC code was not always available, especially when the companies in question happened to be divisions. To counter this issue, information on products, of which the directory is very rich in, was used extensively. Since there are around 50,000 different products listed in the directory, the products were classified into a smaller number of categories to make the data more manageable through the implementation of a semi-supervised learning approach based on a kNN classifier relying on the Levenshtein distance between product names and the rate of co-occurrence in patents. The approach involves applying the classifier on the set of products using a predetermined training set, where only a subset of the products is actually classified

[25] Parent companies are companies from which the founder originated. For example, if a founder used to work in Company X before founding Company Y, then Company X is the parent company of Company Y.

based on threshold requirements designed to keep only the most reliable classifications, such as minimum required values for similarity. The classified products are then added into the training set, based on which the classifier attempts to classify the remaining products, until no more products meet the requirements.

These classifications were then associated with suitable SIC codes, based on which relatedness was calculated. The classification itself is a variation of the framework proposed by Saviotti & Metcalfe (1984), assigning to each product a set of three categories: technical, service, and market.

Method

One of the main issues in empirically analyzing firms' entry decisions is that we can only observe the firms that entered ex-post, as opposed to firms that could have in principle entered ex-ante but did not. This issue often makes regressions that take entry as the dependent variable uninformative, as all firms in the sample have entered by definition. Because of this problem, past empirical analyses of firm entry tended to end up being nothing more than comparisons of overall patterns, for example in terms of number of entrants, across industries and/or years.

Given the importance of entry decisions, however, there have been attempts to address this particular issue. Studies such as Cockburn & MacGarvie (2011), for example, assumed that all firms in the sample were potential entrants to a segment over the sample period, then used a logit hazard model to see when these firms actually entered. Essentially, the question of 'whether' the firms entered was reframed as 'when' the firms entered.

In our study, we follow a slightly different approach and choose to treat all firms established in each year as potential entrants into an industry. Essentially, the empirical analysis is framed in terms of determining whether firms decide to enter the chosen target industry over a different industry, given their level of knowledge relatedness to the target industry as determined by their industry of origin. Thus, the actual dependent variable becomes a binary variable that equals 1 if the firm is established in the chosen target industry and 0 if it is established elsewhere. Thus, the dependent variable and the measure of relatedness, our most important explanatory variable, vary depending on which target industry the analysis focuses on. We then use the logit model to estimate the relationship between knowledge relatedness and the rate of entry.

This paper runs the same regressions with respect to different industries, in an effort to estimate and compare the relationship between knowledge relatedness and firm entry. Specifically, the relationship is estimated for SICs 3674 (semiconductors), 3571 (computers), 7372 (pre-packaged software), and 3663 (radio and TV broadcasting and communication equipment).

The choice of these industries is motivated by a number of reasons. First, an inspection of the overall dataset revealed that the majority of products listed in the directory belonged to these industries. In addition, these industries tend to be considered as being heavily related to each other, especially 3674, 3571, and 7372. This would mean that comparing the patterns of entry may also lead to drawing some implications on the pattern of co-evolution of these industries, an interesting aspect that may be worth investigating further in the future.

Dependent variable

The main dependent variable is the entry into the target industry. In this study, four target industries are considered, which are SICs of 3674, 3571, 7372, and 3663. The value of this variable is equal to 1 if the company has commercialized a product that is associated with the industry in question. As such, there are four dependent variables, for which the logit regressions are run separately. As explained above, whether a product is associated with the industry is defined on the basis of the classification results by a semi-supervised learning algorithm, which assigned technical, service, and market categories to each product.

Explanatory variables

Our main explanatory variable is the measure of knowledge relatedness of the potential entrant (which may be a spinout, spinoff, or a diversifying firm) to the target industry. The measure is based on the indicator proposed by Chang et al. (2021) for technological relatedness which in turn is based on the measure pioneered by Breschi, Lissoni, & Malerba (2003). There are some differences, however. Unlike the original measure, which is defined at the firm level, we compute the measure at the industry level, since we determine the potential entrants' pre-entry knowledge on the basis of their 'industry of origin'.[26] More precisely, the measure is calculated for the potential entrant's industry of origin relative to the target industry. This choice circumvents a potential problem with the original measure, namely that firms tend to get dropped from the sample very easily. This was present because it was constructed on the basis of patents, which not all companies possess enough of to calculate a viable value of the measure for. We are also reminded of the usual caveat that patents are not always the most representative of the technological knowledge in the industry.

The measure in this study, like the original one in Chang et al. (2021), is calculated in two steps[27]. First, we determine the relatedness between dyads of CPCs (Cooperative Patent

[26] In our study, the industry of origin refers to the industry of the parent company. For diversifiers, subsidiaries, or divisions, this is the industry in which the firm were originally active. In the case of spinoffs or spinouts, it is the industry where the founder was previously employed.

[27] See Chang et al. (2021), especially the online appendix, for a more thorough description of how the original measure is calculated.

Classification), based on the similarity of citation patterns of the patents within each CPC class, as in Breschi et al. (2003). This is calculated annually, using patents granted between 1 to 3 years before.

Then, after determining the CPCs of patents in each SIC industry, the knowledge relatedness between a pair of SICs is determined by aggregating these similarities. This is where our measure diverges from the original one in Chang et al. (2021), in that instead of taking the maximum similarity among the possible combinations of CPCs across the two industries, we compute the relatedness between industries i and j as follows.

$$R_{ijt} = \begin{cases} \frac{1}{n(C)} \sum_{c \in C} \max_{\mathrm{d} \in \mathrm{D}} (S_{cdt}) \cdot (1 - Generality_{ct}), & i \neq j \\ 1, & i = j \end{cases} \qquad (2)$$

Here, S_{cdt} refers to the similarity of citation patterns between CPC $c \in D$ and $d \in D$, where C and D refer to the set of CPCs associated with industry i and j respectively, evaluated in year t. Note that relatedness simply equals 1 if it is calculated for the same industries.

Generality is calculated as suggested by Hall, Jaffe, & Trajtenberg (2001), but extended so that the generality of the CPC is calculated instead of the individual patent. The exact calculation is based on the following formula.

$$Generality_{ct} = 1 - \sum_{i} s_{cdt}^2 \qquad (3)$$

Here, s_{cit} refers to all forward citations made to any patent in CPC c up until 5 years after year t by all patents belonging to CPC d. Thus, generality is essentially the reverse of the Herfindahl index of CPCs of citing patents, such that the higher this value, the higher the generality of the CPC. The similarity of the CPC with higher generality is given a lower weight in the calculation of relatedness.

The reasons behind this approach are twofold. First, we ultimately want to develop an indicator that matches the theoretical framework as closely as possible. One of the key features in the definition of relatedness is that it depends heavily on there being general-purpose and industry-specific knowledge, which the measure treats as a continuous concept.

Second, too often the use of the original measure (which relies on maximum CPC-level similarity) has the tendency to label many pairs of SICs to be highly related to each other, especially since sharing even one CPC class automatically causes the value to be equal to 1. This would tend to artificially reduce the variance in the measure, making it difficult to interpret the results in the context of the theoretical framework.

It should also be noted that in many cases, the relatedness measure simply could not be defined. There were two situations in which this might occur. The first occurred when patents for the industries simply did not exist in the specific time period. The second occurred when patents existed, but the generality measure could not be defined because the patents in question were not cited within the 5-year time frame. In the main analysis, both are treated as cases with zero relatedness, since both represent situations where no useful knowledge can be brought to the target industry from the industry of origin.

The other two main explanatory variables are dummies that identify the potential entrant as a diversifier or a spinoff. The diversifier dummy equals 1 if the potential entrant is a subsidiary or division of another company, regardless of which industry that company is in. The spinoff dummy is equal to 1 if the company is not a diversifier or a plain start-up, and if the SIC of the parent company of the founder is associated with any of the products produced by the company, regardless of what that SIC is. In other words, the spinoff dummy simply refers to potential entrants that are not diversifiers or plain start-ups that eventually decided to enter the same industry as their parent company. For example, for a company founded by an employee working in the SIC industry of 3999 and entered 3999, the value of the spinoff dummy is equal to 1.

Note that, unlike the spinoff dummy, the relatedness is calculated between the parent company's SIC and the target industry SIC. So, for the same company described previously, the value of the relatedness to the target industry, for example SIC 3674, has nothing to do with the fact that the company is a spinoff in SIC 3999. The same goes for the diversifier case, where the relatedness between the parent and the actual industry it entered is not a factor.

We then control for other variables such as the founding year, location, log of the number of employees of the company, the diversity of the SICs associated with the products produced by the company, and the age of the parent company.

Descriptive statistics

The entire dataset includes 14,766 unique firms. 7,650 firms are non-startups (i.e., spinouts, spinoffs, and diversifiers), which can be broken down into 4,539 divisions or subsidiaries and 3,111 confirmed spinouts or spinoffs (See Table 3).

[INSERT TABLE 3 ABOUT HERE]

The proportion of each type of entrant varies by year, as can be seen from Figures 2 and 3. In both figures, each bar represents the count over 5-year periods such as 1961-1965 or 1986-1990. Figure 2 shows that the total number of entrants tend to show an increase over time, while Figure 3 shows that the proportions of each type of entrants fluctuate.

[INSERT FIGURES 2 AND 3 HERE]

Figure 4 shows the average level of relatedness for actual entrants and non-entrants to the chosen target industries (i.e., SICs 3674, 3571, 7372, and 3663). The entrants are defined as companies that produce any product that corresponds to the industry. The trend is similar, especially for 3674 and 3571. One interesting aspect is that there is a spike in the relatedness of non-entrants around 1980. A comparable spike occurs for 3571 while this is not true for 3674, which presents an interesting contrast. This might be related to the commercialization of the personal computer around 1975, suggesting that this may be an indication of some kind of discontinuity. The difference in trends across industries and across time also suggests that there may be some industry-level heterogeneity.

[INSERT FIGURE 4 ABOUT HERE]

Table 4 reports the descriptive statistics of the variables used in the analysis. Some additional firms were dropped from Table 3 due to missing data on the industry of origin.

[INSERT TABLE 4 ABOUT HERE]

Results

Table 5 shows the results of the analysis of entry into semiconductors (SIC 3674), based on the estimation of a logit model.

[INSERT TABLE 5 ABOUT HERE]

The linear model, Model 2, shows that there is a negative and significant relationship between relatedness and the rate of entry for spinouts, the baseline for the regressions. At the same time, the relationship is shown to become more positive for diversifiers and spinoffs. In particular, the effect is dramatic enough for spinoffs that the relationship no longer remains negative.

Looking at Model 3 (the full model) gives a more complete story, indicating that there is an inverse U-shaped (IUS) relationship between relatedness and the rate of entry for spinouts, the baseline for the regressions. It should be noted, however, that there is more of the decreasing segment than the increasing segment within the range of values relatedness can take, as the results essentially show that the peak of the curve should occur at a relatedness of $\frac{2.5255}{2\times 6.6739} = 0.189\ (3\ d.p.)$.[28] This also explains why Model 2 also shows such highly significant negative coefficient for relatedness. Interestingly, this relationship essentially flips for diversifiers and spinoffs, to the point that they appear to show a U-shaped relationship instead. This suggests that, for diversifiers and spinoffs, the knowledge effect, differentiation effect, or both change drastically enough such that the relationship changes form.

It is also worth noting that while the theoretical framework does not highlight this aspect, some horizontal translations of the curves also occur. The coefficients suggest that while the peak of the IUS of the spinouts sits at around 0.19 relatedness, the trough for the diversifiers and spinoffs are, respectively, 0.39 and 0.20. In other words, the peaks or troughs tend to occur at higher levels of relatedness for diversifiers, while for spinoffs this happens at lower levels of relatedness. This suggests that while the knowledge effect eventually overwhelms the differentiation effect, this happens much quicker for spinoffs than it does for diversifiers.

Table 6, 7, and 8 show similar results for SICs 3571, 7372, and 3663 respectively.

[INSERT TABLES 6, 7, AND 8 ABOUT HERE]

The results tend to be quite similar for all industries, but some key differences emerge. For example, the entry of spinouts in 3571 (the computer industry) seems to show a much flatter relationship overall, although this may partly be due to the relatively lower number of

[28] This was done in exactly the same way as if one were to find the peak or trough of a quadratic curve.

entrants in the first place. At the same time, 7372 (the software industry) shows a dramatic difference between the different types of entrants, while 3663 (communication equipment) shows the interesting property that spinouts show a consistently lower rate of entry compared to the other types. This is more evident when the predictions are visualized, as in Figure 5 and Figure 6.

[INSERT FIGURES 5 AND 6 ABOUT HERE]

Figure 5 is a visualization of the predicted probabilities of the logit regressions, while Figure 6 is a visualization of the log odds. Both versions show that there is generally a significant and consistent difference in the relationship between relatedness and rate of entry for different types of entrants. Figure 6 is presented alongside Figure 5 because this better represents what the regression equations show. Figure 5, on the other hand, is shown to confirm that the regression does in fact imply that there is a statistically significant difference across the different types of entrants, due to the nonlinear nature of logit regression.

Robustness checks[29]

To see how sensitive our results are to the choices made in the construction of the dataset and/or in the definition of the variables, we carry out some robustness checks. First, we address the issue of whether imputing a value of zero for relatedness when it cannot be defined affects the results. One alternative is to employ the annual average value of relatedness instead. Another alternative is to simply remove observations that have missing values for relatedness. The results of these regressions for SIC 3674 are shown in Table 9, and they seem to be consistent.

[29] The results for the robustness checks are only provided for 3674 due to presentation issues. The results for other industries are available upon request.

[INSERT TABLE 9 ABOUT HERE]

Another issue worth investigating concerns the way entrant types are divided. One possibility may be to compare spinoffs with diversifiers and spinouts. This comparison could be relevant, given that spinoffs have the special property of entering the same industry as their parent company. Spinouts, by definition, enter a different industry and diversifiers could enter either. With this in mind, Table 10 analyzes different comparisons, first between spinoffs and the others, then between spinoffs and diversifiers entering their own industry of origin against the others, for the semiconductor industry.

[INSERT TABLE 10 ABOUT HERE]

Finally, there is the concrete possibility that the relationship between relatedness and entry varies depending on the timeframe of industry evolution. This would actually be somewhat expected given the literature review and Figure 3. While the original analysis included year effects to address this problem, it is also possible that the relationship itself may change over time. For example, the curvature of the knowledge effect may change if the products in the industry tend to become more modular over time, which in this case may cause the relationship to become a steeper IUS. Table 11 examines this possibility by breaking down the sample into different decades within the sample for the semiconductors industry (SIC 3674). The general trend seems to be that while all decades point to a downward curve, the relationship becomes an actual IUS for the range of relatedness that matters starting from the decade 1971-1980. In other words, there seems to be some right shifting over time, although not necessarily in a monotonic manner.

[INSERT TABLE 11 ABOUT HERE]

Discussion

All in all, the results support the claim that there is an IUS relationship between knowledge relatedness and the rate of entry. It also shows that the nature of the relationship changes drastically depending on the type of entrant. This indicates that for each type of entrant, the common knowledge composition will tend to vary.

This is particularly interesting in that it suggests that spinouts are more different from diversifiers and spinoffs than the latter two are to each other. While it is always the case that a higher level of knowledge relatedness is usually better for diversifiers and spinoffs, for spinouts the peak tends to occur at a somewhat lower level, often below 0.5 according to our measure. The theoretical framework would suggest that this is mostly because the differentiation effect is a greater concern for spinouts than for the other types of entrants, because spinouts are less able to identify profitable niches or to sufficiently differentiate themselves from the incumbents, compared to the other types of entrants. In addition, it is unlikely that the knowledge effect will be any larger for spinouts.

Another explanation hinges on the possibility that, instead of the differentiation effect actually being larger, spinouts simply overestimate its magnitude compared to diversifiers or spinoffs. The theoretical framework, in its current form, makes no distinction on whether the two mechanisms are based on perception or objective valuations, but it may be an interesting aspect to explore in further research.

In addition, some considerations on the effects of industry evolution on the probability to enter can be made. The industry evolution literature suggests that an industry can undergo several phases during its evolution, which may also affect the relationship between knowledge relatedness and rate of entry. For example, if knowledge tends to become more modular in nature over time, the marginal impact of the knowledge effect is likely to decrease with relatedness, since the individual units of knowledge would by definition be less complementary to each other. At the same time, differentiation would tend to be easier as firms could simply swap out individual components to make any changes.

This might also explain why the peak seems to vary across the four industries, as seen in Figure 4, and why the trends in Figure 3 are so different from each other. This suggests that there are some industry-level factors that cause the relative strengths of the effects to vary, leading to a change in the overall relationship between relatedness and rate of entry.

This can be linked to one of the major implications of the results, which is that new entrants, whether they are diversifiers, spinouts, or spinoffs, will tend to come from industries that are related to some degree. More importantly, entrants from more related industries will tend to be diversifiers or spinoffs, suggesting that for the incumbents, their most direct competitors in the industry among new firms are either those that have access to a large stock of resources or those that have in-depth knowledge of the industry. On the other hand, the spinouts may be aiming for smaller niches within the industry, choosing to target segments that are not addressed by the incumbents. This may be an interesting avenue for further study, especially in the context of industries where the market is more segmented.

CONCLUSION

The analysis in the study suggests that knowledge relatedness has a nonlinear relationship with the rate of entry, the relationship varying across the different types of entrants. This indicates that there may indeed be two different mechanisms that explain how knowledge relatedness affects firms' decision to enter an industry.

In general, for management and strategy, the results of this paper suggest that incumbents should indeed expect entry from related industries, but that entry has to be disaggregated and examined in terms of the types of entrants that tend to enter the industry. In fact, the composition of entrants greatly affects the market structure of an industry and the types of competition that takes place between incumbents and newcomers.

The variance in the results across industries indicates that industry differences matter. This suggests that there may be industry-level moderators that alter the relationship. It would thus be interesting to investigate in further research what these moderators may be, since this would also alter the types of entrants the incumbents would expect in the first place.

REFERENCES

Abernathy, W., & Utterback, J. 1978. Patterns of industrial innovation. ***Technology Review***. 80 (7): 3-9.

Adams, P., Fontana, R., & Malerba, F. 2015. User-industry spinouts: Downstream industry knowledge as a source of new firm entry and survival. ***Organization Science***. 27 (1): 18-35.

Agarwal, R., & Gort, M. 1996. The evolution of markets and entry, exit and survival of firms. ***The Review of Economics and Statistics***. 78 (3): 489-498.

Agarwal, R., Echambadi, R., Franco, A., & Sarkar, M. 2004. Knowledge transfer through inheritance: Spin-out generation, development, and survival. ***Academy of Management Journal***. 47 (4): 501-522.

Agarwal, R., Sarkar, M., & Echambadi, R. 2002. The conditioning effect of time on firm survival: An industry life cycle approach. ***Academy of Management Journal***. 45 (5): 971-994.

Agarwal, R., & Shah, S. 2014. Knowledge sources of entrepreneurship: Firm formation by academic, user and employee innovators. ***Research Policy***. 43: 1109-1133.

Ahuja, G., & Katila, R. 2001. Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study. ***Strategic Management***. 22: 197-220.

Bayus, B., & Agarwal, R. 2007. The role of pre-entry experience, entry timing, and product technology strategies in explaining firm survival. ***Management Science***. 53 (12): 1887-1902.

Benner, M., & Tripsas, M. 2012. The influence of prior industry affiliation on framing in nascent industries: The evolution of digital cameras. ***Strategic Management Journal***. 33: 277-302.

Bettis, R. 1981. Performance differences in related and unrelated diversified firms. ***Strategic Management Journal***. 2: 379-393.

Breschi, S., Lissoni, F., & Malerba, F. 2003. Knowledge-relatedness in firm technological diversification. ***Research Policy***. 32: 69-87.

Bresnahan, T., & Gambardella, A. 1998. The division of inventive labor and the extent of the market. In: Helpman, E. (Ed.) General purpose technologies and economic growth. MIT Press, Cambridge, MA, 253-281.

Bresnahan, T., & Trajtenberg, M. 1995. General purpose technologies 'Engines of growth'? ***Journal of Econometrics***, 65: 83-108.

Cattani G., & Malerba F. 2021. Evolutionary approaches to innovation, the firm, and the dynamics of industries. ***Strategy Science***. 6 (4) : 265-289.

Chang, S.J. 1996. An evolutionary perspective on diversification and corporate restructuring: Entry, exit, and economic performance during 1981-89. ***Strategic Management Journal***. 17: 587-611.

Chang, S., Eggers, J.P., & Keum, D. 2021. Bottleneck resources, market relatedness, and the dynamics of organizational growth. ***Organization Science***, Articles in Advance, 1-19.

Chatterjee, S., & Wernerfelt, B. 1991. The link between resources and type of diversification: Theory and evidence. ***Strategic Management Journal***. 12 (1): 33-48.

Chen, P., Williams, C., & Agarwal, R. 2012. Growing pains: Pre-entry experience and the challenge of transition to incumbency. ***Strategic Management Journal***. 33: 252-276.

Christensen, H., & Montgomery, C. 1981. Corporate economic performance: Diversification strategy versus market structure. ***Strategic Management Journal***. 2 (4): 327-343.

Clarysse, B., Wright, M., & Van de Velde, E. 2011. Entrepreneurial origin, technological knowledge, and the growth of spin-off companies. ***Journal of Management Studies.*** 48: 1420-1442.

Conti, R., Gambardella, A., & Novelli, E. 2019. Specializing in generality: Firm strategies when intermediate markets work. ***Organization Science***, 30 (1): 126-150.

Dunne, T., Roberts, M., & Samuelson, L. 1988. Patterns of firm entry and exit in U.S. manufacturing industries. ***The RAND Journal of Economics***. 19 (4): 495-515.

Ganco, M., & Agarwal, R.2009. Performance differentials between diversifying entrants and entrepreneurial startups: A complexity approach. ***The Academy of Management Review***. 34 (2): 228-252.

Geroski, P., Mata, J., & Portugal, P. 2010. Founding conditions and the survival of new firms. ***Strategic Management Journal***. 31: 510-529.

Gort, M., & Klepper, S. 1982. Time paths in the diffusion of product innovations. ***The Economic Journal***. 92 (367): 630-653.

Gruber, M., MacMillan, I., & Thompson, J. 2013. Escaping the prior knowledge corridor: What shapes the number and variety of market opportunities identified before market entry of technology start-ups?. ***Organization Science***. 24 (1): 280-300.

Hall, B., Jaffe, A., & Trajtenberg, M. 2001. ***The NBER patent citations data file: Lessons, insights and methodological tools***. (NBER working paper no.8498). Cambridge, MA: National Bureau of Economic Research.

Helfat, C., & Eisenhardt, K. 2004. Intertemporal economies of scope, organizational modularity, and the dynamics of diversification. ***Strategic Management Journal***. 25: 1217-1232.

Helfat, C., & Lieberman, M. 2002. The birth of capabilities: market entry and the importance of pre-history. ***Industrial and Corporate Change***. 11 (4): 725-760.

Helfat, C., & Raubitschek, R. 2000. Product sequencing: Co-evolution of knowledge, capabilities, and products. ***Strategic Management Journal***. 21: 961-979.

Katila, R. 2002. New product search over time: Past ideas in their prime?. ***The Academy of Management Journal***. 45 (5): 995-1010.

Klepper, S. 1996. Entry, exit, growth, and innovation over the product life cycle. ***The American Economic Review***. 86 (3): 562-583.

Klepper, S., & Graddy, E. 1990. The evolution of new industries and the determinants of market structure. ***The RAND Journal of Economics***. 21 (1): 27-44.

Klepper, S., & Sleeper, S. 2005. Entry by spinoffs. ***Management Science***. 51 (8): 1291-1306.

Kumar, M.V. 2013. The costs of related diversification: The impact of the core business on the productivity of related segments. ***Organization Science***. 24 (6): 1827-1846.

Makri, M., Hitt, M., & Lane, P. 2010. Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions. ***Strategic Management Journal***. 31: 602-628.

Pe'er, A., Vertinsky, I., & Keil, T. 2016. Growth and survival: The moderating effects of local agglomeration and local market structure. ***Strategic Management Journal***. 37: 541-564.

Pisano, G. 2017. Toward the prescriptive theory of dynamic capabilities: connecting strategic choice, learning, and competition. ***Industrial and Corporate Change***. 26 (5): 747-762.

Qian, L., Agarwal, R., & Hoetker, G. 2012. Configuration of value chain activities: The effect of pre-entry capabilities, transaction hazards, and industry evolution on decisions to internalize. ***Organization Science***. 23 (5): 1330-1349.

Rosenberg, N. 1963. Technological change in the machine tool industry, 1840-1910. ***The Journal of Economic History***. 23 (4): 414-443.

Sakakibara, M., & Balasubramanian, N. 2020. Human capital, parent size, and the destination industry of spinouts. ***Strategic Management Journal***. 41: 815-840.

Sapienza, H., Parhankangas, A., & Autio, E. 2004. Knowledge relatedness and post-spin-off growth. ***Journal of Business Venturing***. 19: 809-829.

Saviotti, P., & Metcalfe, J. 1984. A theoretical approach to the construction of technological output indicators. ***Research Policy***. 13: 141-151.

Tanriverdi, H., & Venkatraman, N. 2005. Knowledge relatedness and the performance of multibusiness firms. ***Strategic Management Journal.*** 26: 97-119.

Uzunca, B. 2011. Comparative advantages of spinoff firms: An evolutionary perspective. ***Journal of Technology Management and Innovation***. 6 (4): 80-92.

Uzunca, B. 2018. A competence-based view of industry evolution: The impact of submarket convergence on incumbent-entrant dynamics. ***Academy of Management Journal***. 61 (2): 738-768.

LIST OF TABLES

TABLE 1

Possible final forms of relationship between relatedness and entry

		Marginal Knowledge Effect		
		Decrease with Relatedness	Constant with Relatedness	Increase with Relatedness
Marginal Differentiation Effect	Decrease with Relatedness	Any	U-shape	U-shape
	Constant with Relatedness	IUS	Linear	U-shape
	Increase with Relatedness	IUS	IUS	Any

TABLE 2

Characteristics of each entrant type: An illustration

Entrant Type	General-Purpose Knowledge	Industry-Specific Knowledge	Total Knowledge Stock
Diversifier	[0, 100]	[0, 100]	100
Spinout	[0, 50]	[0, 50]	50
Spinoff	[0, 10]	[40, 50]	50

[x,y] denotes that the entrant types has a stock of the knowledge of the column in the range of x to y.

TABLE 3

Entrant types, original dataset

Entrant Type	Frequency	Percentage (%)
Diversifier	4539	59.33
Spinout	2178	28.47
Spinoff	933	12.20
Total	**7650**	**100**

TABLE 4

Descriptive statistics

	N	Min	Max	Mean	S.D.
Entry (3674)	4735	0	1	0.65	0.48
Entry (3571)	4735	0	1	0.25	0.43
Entry (7372)	4735	0	1	0.49	0.50
Entry (3663)	4735	0	1	0.62	0.48
Diversifier	4735	0	1	0.47	0.50
Spinoff	4735	0	1	0.17	0.37
Relatedness to 3674	4735	0.00	1.00	0.26	0.25
Relatedness to 3571	4735	0.00	1.00	0.22	0.24
Relatedness to 7372	4735	0.00	1.00	0.17	0.13
Relatedness to 3663	4735	0.00	1.00	0.22	0.19
Number of Founders	4735	1	5	1.17	0.49
Log (Number of Employees)	4735	0.00	12.39	4.52	1.61
Product Diversity	4735	0.00	0.36	0.06	0.05
Parent Age	4735	0	893[30]	29.56	46.34

[30] Note that about 94% of the originating organizations in the sample are younger than 100 years. The ones that are older than this tend to be universities, government agencies, etc., although some firms do exist. The maximum value, incidentally, belongs to Oxford University, whose founding year is listed as 1096.

TABLE 5

Results for 3674, semiconductor industry

	Model 1	Model 2	Model 3
Constant	0.7251	1.1873**	0.7130
	(0.5701)	(0.5885)	(0.5830)
Relatedness		-1.5567***	2.5255***
		(0.2449)	(0.6128)
Relatedness^2			-6.6739***
			(0.9114)
Diversifier		-0.4159***	0.0489
		(0.1087)	(0.1291)
Spinoff		-0.5818***	0.9366***
		(0.1612)	(0.3617)
Relatedness x Diversifier		1.6891***	-3.2395***
		(0.3078)	(0.7888)
Relatedness^2 x Diversifier			7.6659***
			(1.0449)
Relatedness x Spinoff		4.9371***	-8.8726***
		(0.4047)	(2.8699)
Relatedness^2 x Spinoff			22.7456***
			(5.4494)
Number of Founders	0.1370**	0.0339	0.0165
	(0.0690)	(0.0739)	(0.0750)
Log (Number of Employees)	0.0271	0.0359	0.0407*
	(0.0212)	(0.0221)	(0.0225)
Product Diversity	3.8965***	4.1924***	4.0362***
	(0.5749)	(0.6027)	(0.6045)
Parent Age	-0.0005	0.0002	-0.0003
	(0.0006)	(0.0007)	(0.0007)
Year Effects	Yes	Yes	Yes
State Effects	Yes	Yes	Yes
N	4735	4735	4735
Log Likelihood	-2956.53	-2852.28	-2813.01
	(df = 91)	(df = 96)	(df = 99)

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

TABLE 6

Results for 3571, computer industry

	Model 1	Model 2	Model 3
Constant	-2.6304***	-2.0842***	-2.7145***
	(0.6781)	(0.6943)	(0.7511)
Relatedness		-1.1608***	5.2019***
		(0.2294)	(1.3735)
Relatedness^2			-13.1660***
			(3.1359)
Diversifier		-0.7294***	-0.1816
		(0.1106)	(0.1399)
Spinoff		-0.7675***	0.5454**
		(0.1635)	(0.2260)
Relatedness x Diversifier		1.5895***	-6.8549***
		(0.3455)	(1.5036)
Relatedness^2 x Diversifier			15.5338***
			(3.2148)
Relatedness x Spinoff		4.8551***	-8.8291***
		(0.3876)	(1.8522)
Relatedness^2 x Spinoff			23.2790***
			(3.6393)
Number of Founders	0.3077***	0.1051	0.0984
	(0.0652)	(0.0710)	(0.0722)
Log (Number of Employees)	0.0804***	0.1118***	0.1206***
	(0.0247)	(0.0262)	(0.0265)
Product Diversity	-5.1213***	-5.4720***	-5.5504***
	(0.5566)	(0.5855)	(0.5949)
Parent Age	0.0006	0.0008	0.0005
	(0.0007)	(0.0008)	(0.0008)
Year Effects	Yes	Yes	Yes
State Effects	Yes	Yes	Yes
N	4735	4735	4735
Log Likelihood	-2556.91	-2425.60	-2379.20
	(df = 91)	(df = 96)	(df = 99)

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

TABLE 7

Results for 7372, computer software industry

	Model 1	Model 2	Model 3
Constant	-0.5089	-0.1817	-0.3293
	(0.5154)	(0.5224)	(0.5338)
Relatedness		-0.5204	1.6228**
		(0.3876)	(0.7342)
Relatedness^2			-4.5894***
			(1.1532)
Diversifier		-0.4202***	-0.2232*
		(0.1146)	(0.1322)
Spinoff		-0.3487*	1.2793***
		(0.1837)	(0.4092)
Relatedness x Diversifier		0.9421*	-1.7695*
		(0.5650)	(0.9845)
Relatedness^2 x Diversifier			5.7507***
			(1.5553)
Relatedness x Spinoff		4.8373***	-13.6016***
		(0.8196)	(4.0850)
Relatedness^2 x Spinoff			43.9227***
			(10.0433)
Number of Founders	0.3211***	0.1612**	0.1465**
	(0.0685)	(0.0712)	(0.0711)
Log (Number of Employees)	0.0696***	0.0873***	0.0874***
	(0.0209)	(0.0215)	(0.0216)
Product Diversity	-9.1294***	-9.2211***	9.3318***
	(0.6519)	(0.6592)	(0.6696)
Parent Age	0.0015**	0.0017**	0.0015*
	(0.0007)	(0.0008)	(0.0008)
Year Effects	Yes	Yes	Yes
State Effects	Yes	Yes	Yes
N	4735	4735	4735
Log Likelihood	-3110.09	-3046.44	-3031.98
	(df = 91)	(df = 96)	(df = 99)

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

TABLE 8

Results for 3663, communications equipment industry

	Model 1	Model 2	Model 3
Constant	2.5091***	2.4487***	2.2202***
	(0.6174)	(0.6262)	(0.6260)
Relatedness		-1.0554***	1.7275***
		(0.3458)	(0.6635)
Relatedness^2			-4.8487***
			(0.8293)
Diversifier		0.0532	0.2363*
		(0.1216)	(0.1423)
Spinoff		-0.3625**	0.8276***
		(0.1552)	(0.2792)
Relatedness x Diversifier		0.9473**	-1.2316
		(0.4342)	(0.8857)
Relatedness^2 x Diversifier			4.1445***
			(1.0468)
Relatedness x Spinoff		1.9430***	-6.8701***
		(0.4963)	(1.4637)
Relatedness^2 x Spinoff			12.1941***
			(1.5452)
Number of Founders	-0.1980***	-0.1299*	-0.1334*
	(0.0675)	(0.0713)	(0.0719)
Log (Number of Employees)	-0.0124	-0.0216	-0.0214
	(0.0236)	(0.0241)	(0.0243)
Product Diversity	-20.3645***	-20.1146***	-20.3492***
	(1.1096)	(1.1171)	(1.1341)
Parent Age	-0.0013*	-0.0008	-0.0013
	(0.0007)	(0.0007)	(0.0007)
Year Effects	Yes	Yes	Yes
State Effects	Yes	Yes	Yes
N	4735	4735	4735
Log Likelihood	-2620.02	-2609.02	-2588.19
	(df = 91)	(df = 96)	(df = 99)

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

TABLE 9

Results with alternative imputations

	Imputation by Annual Average	Removed NAs
Constant	0.6346	0.4361
	(0.6106)	(0.6668)
Relatedness	3.3882***	3.7705***
	(0.8769)	(0.9313)
Relatedness^2	-7.2346***	-7.8691***
	(1.0465)	(1.1416)
Diversifier	-0.0007	0.0459
	(0.2258)	(0.2349)
Spinoff	1.0229**	0.9826**
	(0.4046)	(0.4152)
Relatedness x Diversifier	-2.7115**	-3.6383***
	(1.1037)	(1.2124)
Relatedness^2 x Diversifier	7.0270***	8.1683***
	(1.2211)	(1.3460)
Relatedness x Spinoff	-8.9328***	-9.0369***
	(3.0015)	(3.2853)
Relatedness^2 x Spinoff	22.1626***	22.7196***
	(5.4864)	(6.2689)
Controls	Yes	Yes
Year Effects	Yes	Yes
State Effects	Yes	Yes
N	4735	3868
Log Likelihood	-2813.91	-2240.98
	(df = 99)	(df = 99)

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

TABLE 10

Results with alternative comparisons

	Spinoffs vs Others	Spinoffs + Same Industry Diversifiers vs Others
Constant	0.7920	0.4631
	(0.5752)	(0.5917)
Relatedness	0.0064	2.3345***
	(0.3876)	(0.4976)
Relatedness^2	-0.5254	-6.1288***
	(0.4032)	(0.6539)
Same Industry Diversifier		0.4835***
		(0.1266)
Spinoff	0.8784**	1.1171***
	(0.3876)	(0.3560)
Relatedness x Same Industry Diversifier		-4.9674***
		(0.9881)
Relatedness^2 x Same Industry Diversifier		14.3399***
		(1.7880)
Relatedness x Spinoff	-6.3314**	-8.7620***
	(2.8400)	(2.8458)
Relatedness^2 x Spinoff	16.602***	22.1876***
	(5.4033)	(5.4168)
Controls	Yes	Yes
Year Effects	Yes	Yes
State Effects	Yes	Yes
N	4735	4735
Log Likelihood	-2852.19	-2728.26
	(df = 96)	(df = 99)

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

TABLE 11

3674, results by decade

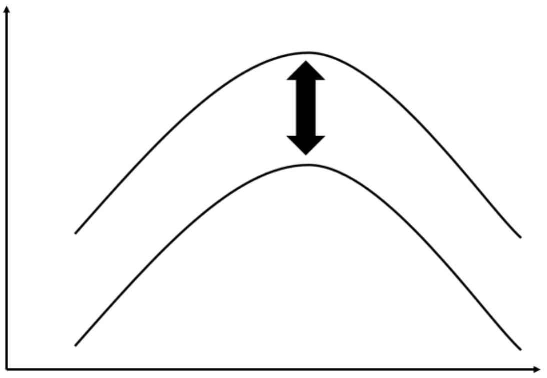
	1951-1960	1961-1970	1971-1980	1981-1990
Relatedness	4.4044	0.6071	3.8699***	1.7647*
	(2.9192)	(1.4099)	(1.3994)	(0.9431)
Relatedness^2	-6.7853**	-3.9689***	-9.0003***	-6.0617***
	(3.3956)	(1.4758)	(2.7815)	(1.3461)
Diversifier	0.4458	-0.3015	-0.0959	0.3408
	(0.4175)	(0.2905)	(0.2405)	(0.2390)
Spinoff	1.1031	0.7107	0.3242	2.3795**
	(0.8256)	(0.6175)	(0.6944)	(1.1196)
Relatedness x Diversifier	-4.8300	-0.2505	-4.5363***	-4.1970***
	(3.4864)	(1.6966)	(1.6340)	(1.4061)
Relatedness^2 x Diversifier	7.4740*	3.9108**	10.5816***	7.9632***
	(3.9909)	(1.7381)	(2.9263)	(1.7183)
Relatedness x Spinoff	-13.5510	-3.2542	-7.9858	-23.7856***
	(13.1861)	(3.5065)	(5.1001)	(8.7202)
Relatedness^2 x Spinoff	71.3883*	10.9426***	22.1391**	53.4528***
	(39.2851)	(3.5231)	(9.4758)	(15.6344)
Controls	Yes	Yes	Yes	Yes
N	632	1300	1440	1284

LIST OF FIGURES

FIGURE 1

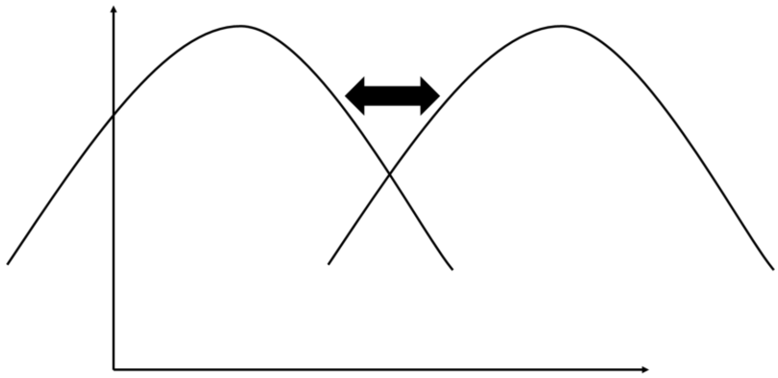
Visualization of theoretical description of changes of curves and their shapes

Upward/Downward Shift



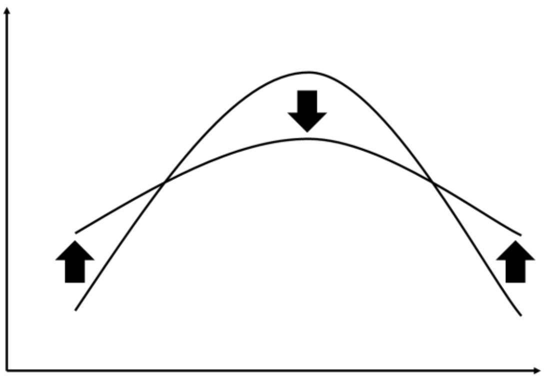
May be caused by changes in total knowledge/differentiation effect.

Left/Right Shift



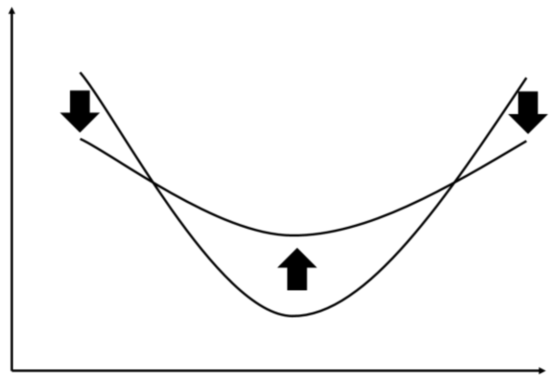
May be caused by changes in marginal knowledge/differentiation effect.

Flatten (Inverted U-shaped Curve)



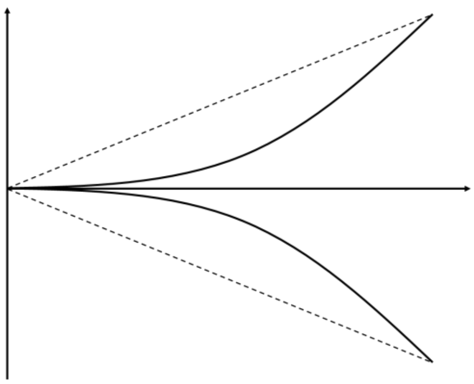
May be caused by decrease in concavity of knowledge and/or differentiation effect.

Flatten (U-shaped Curve)



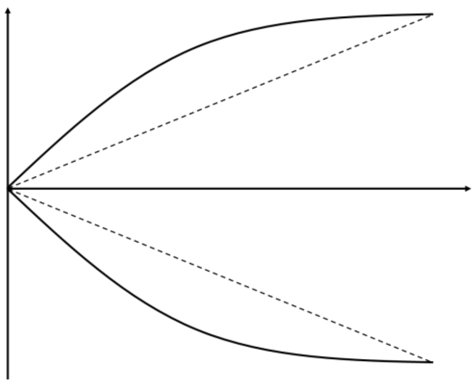
May be caused by decrease in convexity of knowledge and/or differentiation effect.

Convex (Knowledge and Differentiation Effects)



The dotted straight line represents a convexity of zero.

Concave (Knowledge and Differentiation Effects)



The dotted straight line represents a concavity of zero.

FIGURE 2

Number of entrants and composition over time

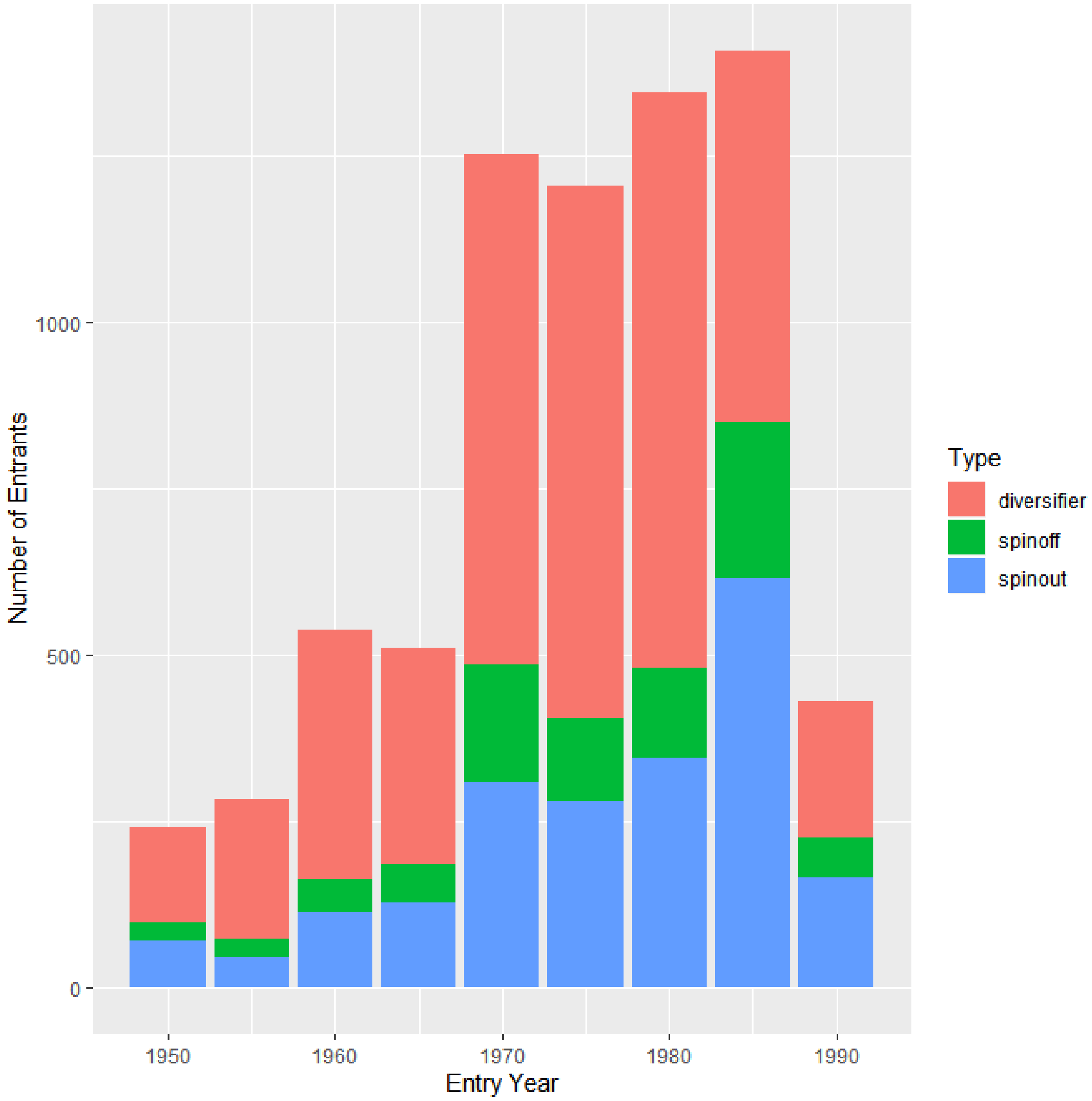


FIGURE 3

Proportion of entrant types by year

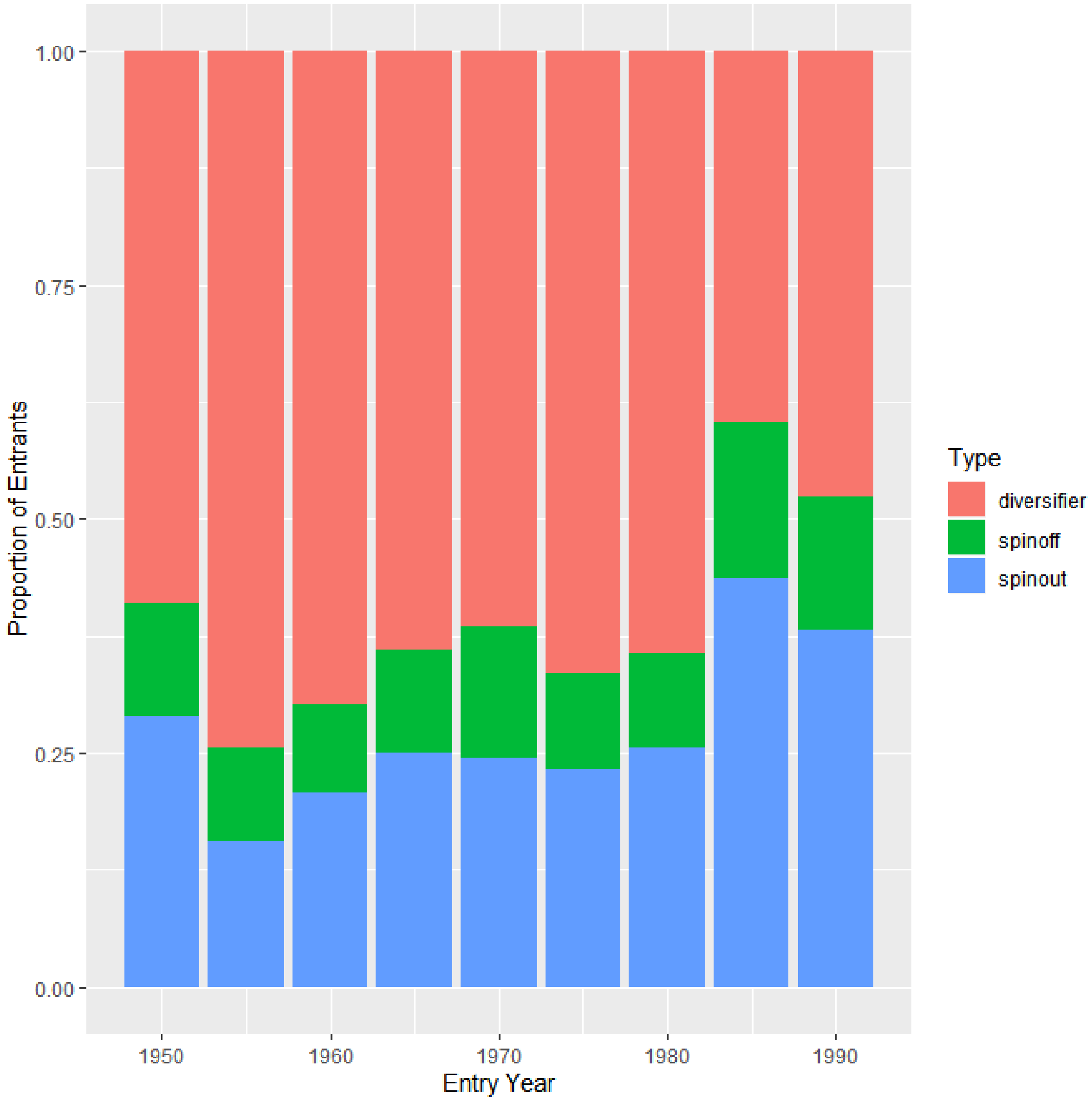
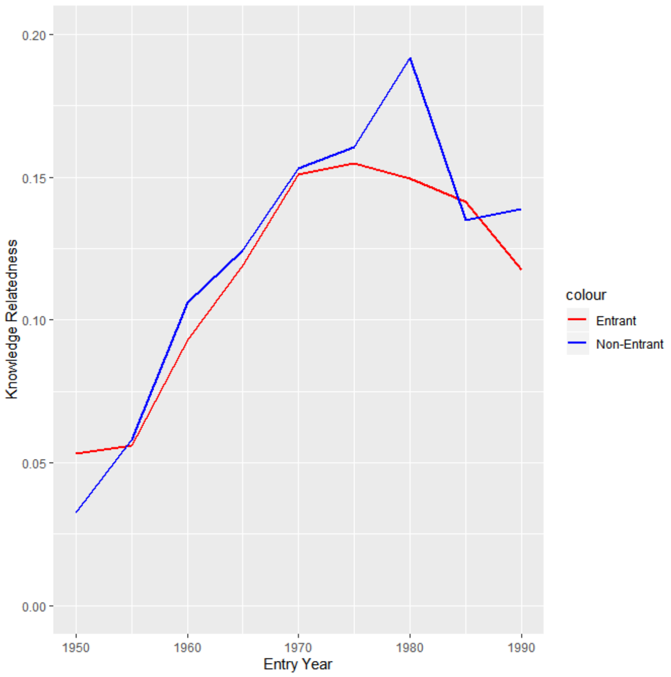


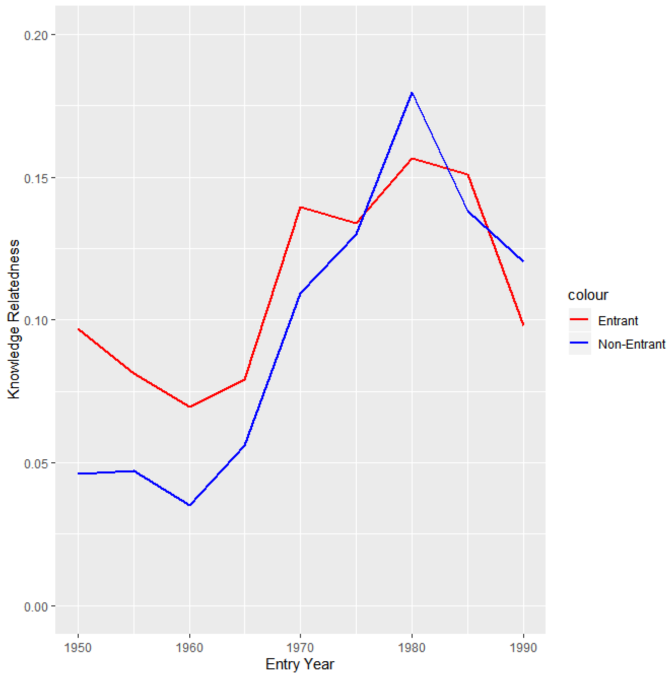
FIGURE 4

Average level of relatedness, entrant vs non-entrant

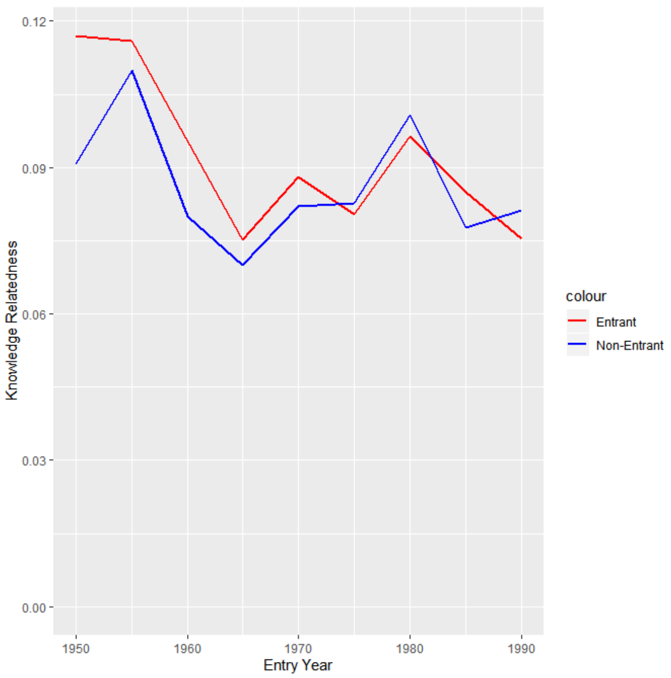
3674



3571



7372



3663

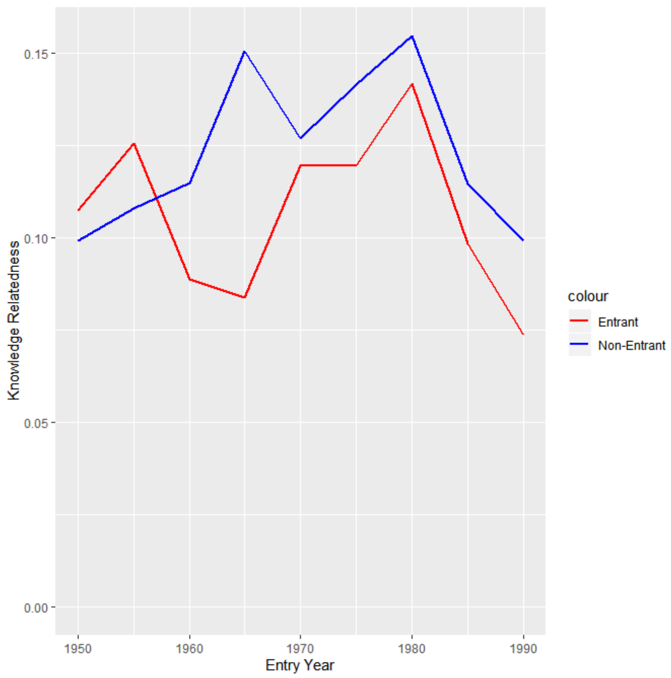
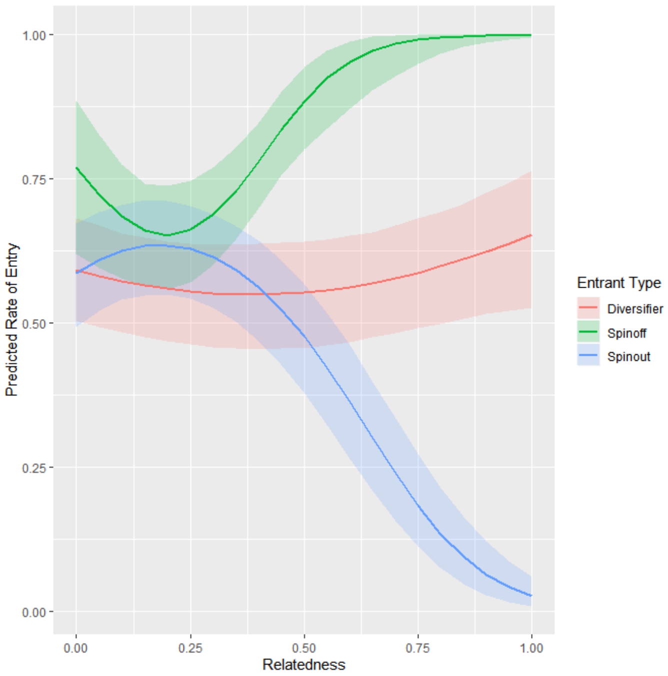


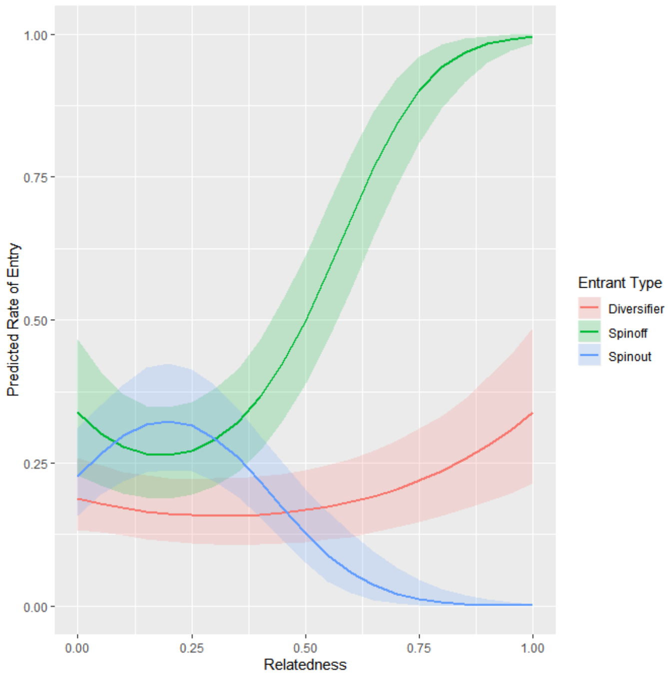
FIGURE 5

Visualization of results of predicted probability

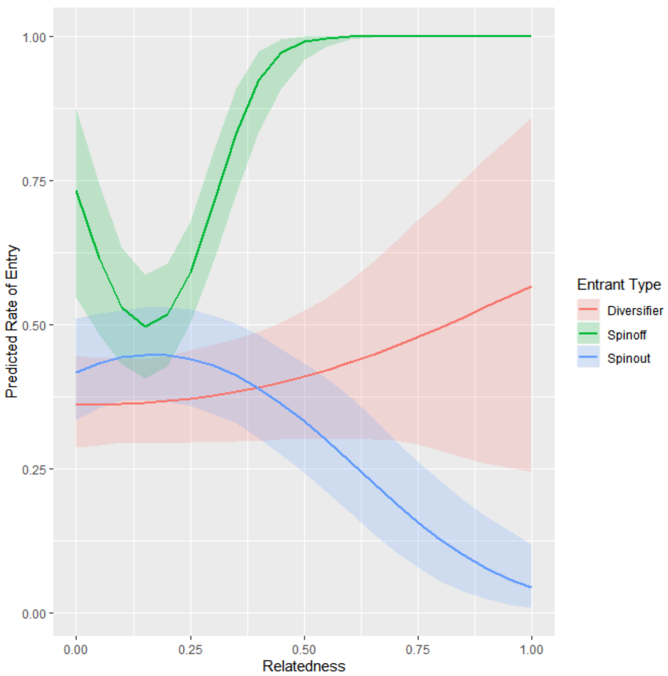
3674



3571



7372



3663

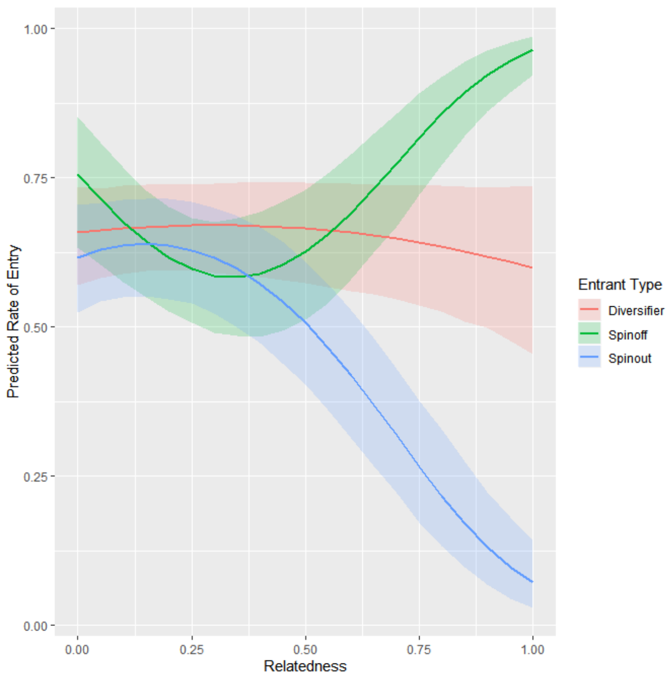
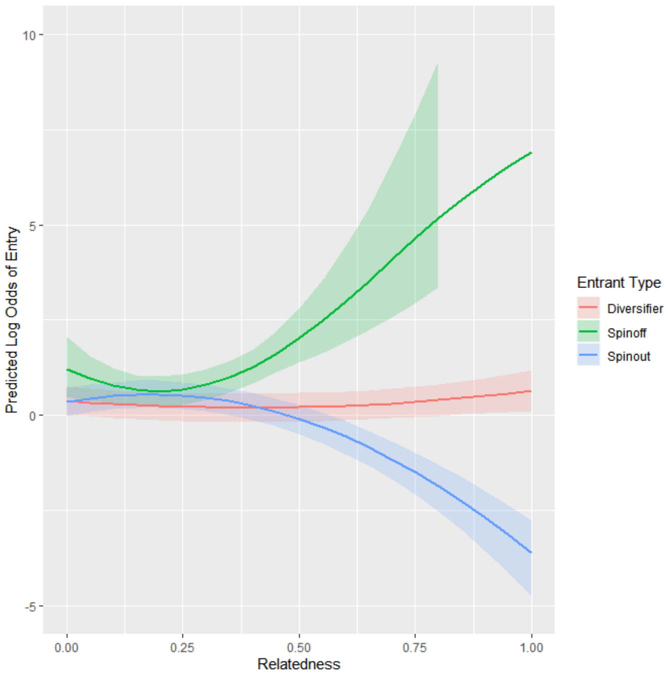


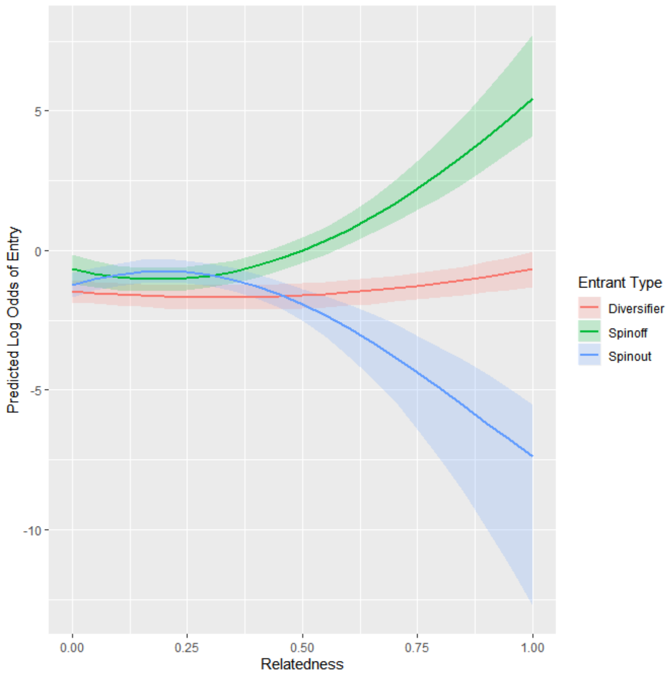
FIGURE 6

Visualization of results of predicted log odds

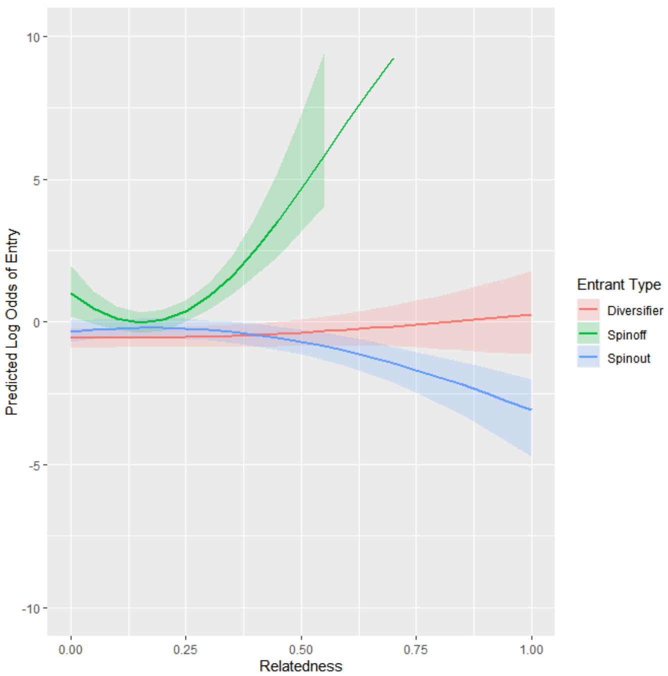
3674



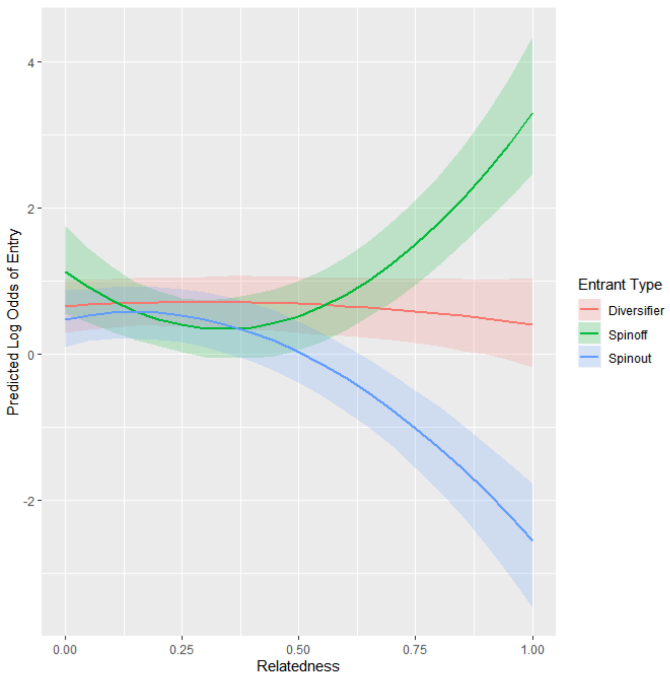
3571



7372



3663



APPENDIX

Explanation of Table 1: Knowledge relatedness and firm entry

The overall effect of knowledge relatedness on firm entry results from the combination of two effects. The first is the knowledge effect. According to this effect, higher knowledge relatedness implies possession of a higher stock of (target) industry-specific knowledge. This implies that the potential entrant has more knowledge that they can immediately deploy in the target industry, which would increase their expected profitability and survival, thus encouraging entry.

The second is the differentiation effect. According to this effect, higher knowledge relatedness implies that the potential entrant is very similar in knowledge composition to the incumbents in the target industry. This implies that the potential entrant will generally find it more difficult to differentiate itself from the incumbents, which would tend to hurt profitability as it faces direct competition from them. This would tend to discourage entry instead.

The two effect exert opposing forces, whose interplay will determine form the final relationship between relatedness and rate of entry will take. Generally, the final outcome will tend to depend on the curvature of each mechanism, or on how the marginal effect from each force changes with increasing relatedness.

What ultimately needs to be analyzed are two aspects.

(i) What is the curvature of the resulting relationship?

(ii) Where does the peak or trough occur in the relationship?

Curvature problem

The first question can be addressed by looking at the resulting sum of the curvature of the knowledge effect and of the differentiation effect. Let $k(r)$ and $d(r)$ be the functions describing the knowledge effect and the differentiation effect respectively, where r represents the level of knowledge relatedness. If we assume that the knowledge effect is always positive and monotonically increasing with relatedness, while the differentiation effect is always negative and monotonically decreasing with relatedness, we may come up with three possible scenarios.

(a) $k''(r) + d''(r) < 0, \ \forall r \in [0, 1]$

(b) $k''(r) + d''(r) = 0, \ \forall r \in [0, 1]$

(c) $k''(r) + d''(r) > 0, \ \forall r \in [0, 1]$

Scenario (a) implies that the relationship takes an IUS form, Scenario (b) leads to a linear relationship, and Scenario (c) implies a U-shaped relationship.

To discuss how each scenario might arise, it is necessary to analyze under what conditions the knowledge effect and the differentiation effect might display certain curvatures.

Knowledge effect. One factor that affects the curvature of the knowledge effect is the degree of complementarity between the individual "units" of industry-specific knowledge. If there is no or very low complementarity, it is reasonable to assume that the marginal contribution of the knowledge effect to the rate of entry will tend to decrease, as is often the case with marginal benefits. On the other hand, if there is high complementarity, such as in cases where knowledge tends to be integrated, it may be possible for an additional unit of knowledge to have a larger impact as it amplifies the value of the other knowledge as well.

Differentiation effect. The differentiation effect arises from the fact that lower industry specific knowledge also implies a relatively higher stock of general-purpose knowledge. If we assume that the benefits from higher general-purpose knowledge in differentiation decreases with an increase of its stock, then we can claim that the marginal impact of relatedness on the magnitude of the differentiation effect increases with relatedness.

Given the above argument, the conditions for the occurrence of each scenario, especially (a) and (c), can be analyzed.

IUS form vs U-shaped form. Scenario (a) will always arise if the magnitude of the marginal impact of the knowledge effect decreases with relatedness while the opposite is true with the differentiation effect. One of these effects may also show constant marginal impact for the condition to hold.

If this is not the case, then whether the scenario may still arise depends on which curvature is more dominant. If, for example, the marginal knowledge effect is decreasing at a rate faster than the rate at which the magnitude of the marginal differentiation effect is decreasing, we may still end up with the IUS form.

The conditions for which Scenario (c) will occur will be the opposite.

Existence of peak or trough

The second question, on the other hand, is more of an empirical question. For scenario (a) or (c) to occur, it will have to be the case that

$$\exists\ r \in (0, 1)\ such\ that\ k'(r) + d'(r) = 0$$

For this to occur, the conditions for each scenario will be as follows:

IUS form. For the peak to appear in the IUS relationship it needs to be the case that, up to a certain level of relatedness, the sum of the marginal impact from each effect is positive, assuming the conditions for curvature are met. In other words, the marginal impact from the differentiation effect must be smaller in magnitude than that from the knowledge effect. This may occur if the importance of having the industry-specific knowledge is high, which may be the case, for example, in high-tech industries.

However, it must also be true that the magnitude of the marginal impact of the differentiation effect must increase at a faster rate than that of the knowledge effect, since at some stage the differentiation effect must overwhelm the knowledge effect. Thus, in terms of Table 1, an IUS form is most likely to be observed if the marginal cost of the differentiation effect increases with relatedness while the marginal benefit from the knowledge effect decreases with relatedness.

U-shaped form. Essentially the opposite from the IUS case, we need the sum of the marginal impact to be negative at lower levels of relatedness. This would imply that the marginal impact from the differentiation effect is larger in magnitude than that of the knowledge effect. In other words, the marginal benefits from having an additional unit of industry-specific knowledge will have to be overshadowed by the marginal costs from the added difficulty in differentiation. This might occur especially in industries where competition is more heavily based on advertising or on variables other than R&D investment.

Of course, above a certain level of relatedness, the marginal benefit of the knowledge effect must overtake the marginal cost of the differentiation effect. Thus, this is most likely to occur if the marginal cost of the differentiation effect decreases with relatedness while the marginal benefit from the knowledge effect increases with relatedness.

Who enters and when? Knowledge relatedness, entrant heterogeneity, and industry characteristics

ABSTRACT

The knowledge relatedness of potential entrants to the target industry is often cited as one of the main drivers of entry. However, empirical evidence shows that the relationship between relatedness and the probability to enter often varies across industry and time, suggesting that industry characteristics may act as moderators in this relationship. This study provides a theoretical framework to identify these moderators, focusing on modularity, rate of change, technological discontinuities, and concentration across different types of entrants. It then empirically tests this framework using a dataset of firms active in the US electronics industry between 1960 and 1990. The results of this paper show that industry characteristics matter in moderating the relationship between knowledge relatedness and entry, but some are more relevant than others.

Keywords:

Firm Entry; Knowledge Relatedness; Entrant Heterogeneity; Modularity; Industry Evolution

Who enters and when? Knowledge relatedness, entrant heterogeneity, and industry characteristics

INTRODUCTION

The relatedness between the knowledge of a potential entrant and that of the target industry has often been cited as one of the major drivers of firms' entry decisions (Helfat & Lieberman, 2002; Klepper & Sleeper, 2005; Bayus & Agarwal, 2007; Agarwal & Shah, 2014; Adams, Fontana, & Malerba, 2015). However, it has also been found that the pattern of entry into an industry varies across industries and over time (Agarwal & Gort, 1996; Klepper, 1996; Agarwal & Gort, 2002). This suggests that industry characteristics and their change over time may affect the relationship between knowledge relatedness and firm entry. This study proposes a framework able to identify the mechanisms through which industry level characteristics may moderate the relationship between knowledge relatedness and entry, and how the effect of the moderators varies across heterogeneous entrants. The framework is then tested with a large dataset that tracks the US electronics industry from 1960 to 1990.

The consensus on the relationship between knowledge relatedness and rate of entry is that in general higher knowledge relatedness tends to lead to a higher rate of entry. However, the literature on industry evolution suggests that the relationship may not necessarily stay constant throughout the industry life cycle. The relationship may not even be linear to begin with, according more recent evidence, especially the one regarding spinouts and spinoffs (Sapienza, Parhankangas, & Autio, 2004; Clarysse, Wright, & Van de Velde, 2011; Adams et al., 2015; Sakakibara & Balasubramanian, 2020). This indicates that the link between knowledge relatedness and firm entry may not be as straightforward as originally thought.

One way in which the evolution of the industry may affect the relationship may be by altering the relative importance of different types of knowledge. This change in the relevance of different types of knowledge in an industry over time can be associated with the broader view of industrial change established by Schumpeter (Schumpeter, 1934; Schumpeter, 1942). While this idea is often used to explain why the number of entrants change over time, it may also contribute to understanding how it changes the role of knowledge relatedness in firms' entry decisions.

In this paper, a theoretical framework with which to analyze how industry characteristics affect the relationship between knowledge relatedness and the entry of different types of firms is proposed. Four moderators (i.e., modularity, rate of change, technological discontinuities, and concentration), and three types of entrants (i.e., diversifiers, spinoffs, and spinouts) are considered. This framework is then tested by using a dataset derived from the Western Electronics Manufacturers Association (WEMA) directory, which provides information on firms active in the US electronics industry from 1960 to 1990.

This paper explores new grounds in the relationship between knowledge relatedness and entry. Given its broad scope, it should be considered work in progress. The results from this paper, however, highlight some relevant avenues for further research. These will be discussed at the end of the paper.

LITERATURE REVIEW

Relatedness and entry

The role of knowledge, and knowledge relatedness, in firms' entry decisions are considered to be highly significant in the literature (Klepper & Simons, 2000; Helfat & Lieberman, 2002; Bayus & Agarwal, 2007). This was found to be particularly useful to explain the heterogeneity of entrants, and specifically to explain why entrants from some industries tend to enter an industry more than others (Klepper & Simons, 2000; Helfat & Lieberman, 2002; Krafft, 2004).

While the precise definition of knowledge relatedness between a potential entrant and an industry is debated (Makri, Hitt, & Lane, 2010), it is also accepted that relatedness be understood primarily in terms of the usefulness of a potential entrant's knowledge in the focal industry. Consequently, higher knowledge relatedness is often thought to lead to higher rates of entry, especially in earlier studies such as Klepper & Simons (2000) or Helfat & Lieberman (2002). Exactly how this relationship arises, however, is not well understood.

Indeed, some of the more recent empirical studies tend to find that overly high levels relatedness may have a deterring effect on entry (Clarysse, Wright, & Van de Velde, 2011; Sakakibara & Balasubramanian, 2020). The studies that show the entry encouraging effect of high knowledge relatedness tend to focus on the potential entrants having access to knowledge that can be (re)deployed (Agarwal, Echambadi, Franco, & Sarkar, 2004; Uzunca, 2011). The studies that emphasize the entry deterring effects, on the other hand, show that high relatedness can also make the potential entrants more vulnerable to competitive pressure from the incumbents. The fact that such contradicting results are found suggests that a more thorough analysis is needed to identify the mechanism behind the relationship between knowledge relatedness and firm entry.

One particular aspect that can be emphasized from the previous discussion is that the consequences of high knowledge relatedness seem to vary across the different types of entrants. While heterogeneity among entrants in terms of knowledge resources has already been established as a stylized fact in the literature (Helfat & Lieberman, 2002; Agarwal et al., 2004; Ganco & Agarwal, 2009; Agarwal & Shah, 2014; Adams et al., 2015), specific studies on the relationship between knowledge relatedness and entry show the relevance of categorizing entrants into different types.

Clarysse et al. (2011) or Sakakibara & Balasubramanian (2020), for example, focus on inter-industry spinouts and intra-industry spinoffs.[31] Spinouts and spinoffs are frequently discussed in the entry literature, especially in more recent studies, due to the fact that they have a very distinct property that sets them apart from the other types of entrants. Specifically, they are thought to possess knowledge, likely related to the industry they enter, that gives them an edge in performance and survival compared to most entrants (Helfat & Lieberman, 2002).

The differences in terms of knowledge resources across types of entrants have been documented in studies such as Agarwal et al. (2004), Agarwal & Shah (2014), and Uzunca (2011). This literature identifies at least three types of entrants: diversifiers, spinouts, and spinoffs. Diversifiers are essentially incumbents of another industry who decide to expand into the target industry (Helfat & Lieberman, 2002). Spinouts are entrants whose founders previously worked in a firm in another industry before establishing their own firm. While still technically start-ups, their defining characteristic is the fact that they have some stock of knowledge inherited from the parent via their founders (Adams, Fontana, & Malerba, 2015). Spinoffs are similar, except that the founder comes from a company that is in the same industry.

[31] Also referred to as out-of-industry spinouts and within-industry spinouts, respectively, in some studies (e.g., Sakakibara & Balasubramanian, 2020). The former are start-ups founded by previous employees of companies established in a different industry, while the latter are similar except that the founder comes from a company in the same industry.

Industry heterogeneity: Identifying some key industry characteristics

Another major stylized fact in the literature on entry is that the pattern of entry tends to differ across industries. This not only refers to the overall rate of entry, but also to the characteristics of the entrants. For example, Dunne, Roberts, & Samuelson (1988) showed that there tends to be a large variation across industries in terms of the size of the entrants. In addition, the industry life cycle literature points to the fact that a stage of an industry greatly affects the types of entrants and the patterns of entry (Agarwal & Gort, 1996; Klepper, 1996). Other studies suggest that certain industrial characteristics or technological contingencies may have an impact on firms' decisions to enter.

The first is the *rate of technological change* in the industry. The speed of technological change affects the usefulness of the knowledge of the potential entrant as well as the rate of learning. In addition, this will also affect what types of potential entrants are more likely to enter, given that not all firms are sufficiently adept in keeping up with changes in knowledge (Zott, 2003).

The second characteristic is the presence and/or frequency of *technological discontinuities*. Technological discontinuities are events that level the playing field, especially if their source is exogenous to the industry. The actual effect of the discontinuities on incumbents can go both ways, in that they can be competence-enhancing or competence-destroying, but both lead to periods of technological ferment (Tushman & Anderson, 1986; Anderson & Tushman, 1990). The consequence for knowledge relatedness is that a discontinuity can cause the actual level of knowledge relatedness to become uncertain, at least for some of the potential entrants. The direct impact of discontinuity on entry may vary depending on its type, but how it affects the relationship between knowledge relatedness and entry may be more consistent.

The third characteristic is the *modularity of knowledge*. While modularity can be defined at many levels[32], its definition at the product level is easiest to understand, which is "the degree to which functionality corresponds to a component and the interfaces between the components within the product architecture are de-coupled" (Ulrich, 1995; Brusoni & Prencipe, 2001). Essentially, modularity refers to the extent to which individual components can be freely modified without affecting the others, where "components" can mean elements of knowledge when defined at the knowledge level (Yayavaram & Ahuja, 2008). Higher modularity in the industry tends to be directly associated with more specialized firms in the industry, suggesting more entry overall. This also implies that the importance of architectural knowledge decreases, suggesting that potential entrants can survive with a smaller share of the knowledge in the industry.

The final industry characteristic is the *concentration of innovative activity*. This is also a common indicator employed to characterize whether an industry is Schumpeter Mark I or Mark II (Malerba & Orsenigo, 1997; Breschi, Malerba, & Orsenigo, 2000; Fontana, Martinelli, & Nuvolari, 2021). Innovative concentration is a usually cited as a source of barriers to entry and it is directly associated with some of the knowledge characteristics of the industry. A high concentration of innovative activity can, for example, impact the research costs involved in innovating in the industry (Klepper, 1996; Malerba & Orsenigo, 2002). In addition, a high concentration implies that incumbents would have larger stock of knowledge overall (Malerba & Orsenigo, 1997; Breschi et al., 2000), similar to how high market concentration is associated with the presence of larger firms in general (Demsetz, 1973).

[32] Some of the most common levels at which the concept of modularity is used are products, organizations, and knowledge (Ulrich, 1995; Brusoni & Prencipe, 2001).

THEORETICAL FRAMEWORK

Knowledge relatedness and entry

Various definitions of knowledge relatedness exist (Sapienza et al., 2004; Tanriverdi & Venkatraman, 2005; Benner & Tripsas, 2012; Chang, Eggers, & Keum, 2021), but they share the commonality that they refer to how much useful knowledge the firm has in relation to the focal industry. In this study, we define knowledge relatedness by applying the categorization of knowledge proposed by Pisano (2017), which allows us to conceptualize a firm's stock of knowledge as consisting of industry-specific and general-purpose knowledge.[33]

Specifically, we define knowledge relatedness as the average similarity of knowledge between the potential entrant i (which may be a diversifier, spinout, or spinoff) and the target industry j that they may enter, weighted by the degree of specificity to the target industry (see equation 1 below). The idea behind this notion is that, given a similar level of overlap, the knowledge that is more specific to the target industry is a better indication of the knowledge relatedness between the potential entrant and the industry.

$$Knowledge\ Relatedness_{ij} = \frac{1}{K}\sum_{k} S_j(Knowledge_i) \cdot \left(Industry\ Specificity_j\right) \qquad (1)$$

Here, $S_j(Knowledge_i)$ represents the similarity of the unit of knowledge held by potential entrant i to the knowledge of target industry j, $Industry\ Specificity_j$ refers to the degree at which the knowledge in question is specific to the target industry j, and $k = 1, \dots, K$ is the stock of knowledge the potential entrant has. In summary, the equation defines knowledge relatedness as the average similarity of the knowledge stock of the potential entrant relative to

[33] Industry-specific knowledge refers to knowledge that is applicable in a specific industry. General-purpose knowledge, on the other hand, is applicable in a wider range of activities (Pisano, 2017).

the target industry it may enter, but with the knowledge that is specific to that target industry contributing more significantly.

This conceptualization helps us to identify two main mechanisms through which knowledge relatedness may affects firms' entry decisions. On the one hand, we have what we will call the *knowledge effect*. Essentially, higher relatedness implies that the potential entrant has more knowledge that can be deployed in the industry it is considering as a target, thus leading to higher expectations of survival and profitability, and therefore encouraging entry (Helfat & Lieberman, 2002; Bayus & Agarwal, 2007). On the other hand, higher relatedness also implies that potential entrants are more likely to be similar to incumbents in the target industry in terms of the knowledge they can access. This can harm their ability to innovate and differentiate themselves from other firms, exposing them to more competitive pressure (Clarysse, Wright, & Van de Velde, 2011; Katila, 2002; Sakakibara & Balasubramanian, 2020; Sapienza, Parhankangas, & Autio, 2004). This will be referred to as the *differentiation effect*.

The two effects, knowledge effect and differentiation effect, work in opposite directions so the actual relationship between knowledge relatedness and the rate of entry could go in either direction or even be flat. It may even be the case that the relationship ends up being nonlinear, as one may suppose that at least one of the two effects will not necessarily be linear. For example, the marginal impact of the knowledge effect on the rate of entry may decrease with relatedness. Under certain conditions, it may be the case that the relationship between relatedness and entry is an inverted-U shape. This is assuming that either the marginal knowledge effect becomes smaller with relatedness or the marginal differentiation effect becomes larger with relatedness[34].

[34] If neither of these conditions are satisfied, then we would simply observe a linear effect or possibly even a U-

Before going into the discussion of the changes in the relationship between knowledge relatedness and the rate of entry, along with discussions on the two effects, Figure 1 provides a visualization of some of the descriptions of these changes. This figure should be useful in understanding the discussion.

[INSERT FIGURE 1 ABOUT HERE]

Relatedness and entrant types. In this paper, we consider three types of entrants: diversifiers, spinoffs, and spinouts. One of the principal differences between diversifiers and the other types of entrants would be that diversifiers tend to have access to a larger stock of knowledge (Bayus & Agarwal, 2007; Helfat & Lieberman, 2002). Thus, given the same level of relatedness, a diversifier would be endowed with a larger stock of both industry-specific and general-purpose knowledge.

Table 1 below summarizes some of the major differences, based on what the literature suggests, between the three types in an illustrative manner.

[INSERT TABLE 1 ABOUT HERE]

This can have two consequences in terms of the effects examined above. First, diversifiers would tend to have a larger stock of industry-specific knowledge that can be directly deployed in the target industry when compared to a spinout with the same level of relatedness (Helfat and Lieberman, 2002). This would help them achieve better profitability compared to other types of entrants such as spinouts, leading to a higher rate of entry by diversifiers. This implies a larger knowledge effect overall, resulting in an upward shift of the curve of the relationship between relatedness and the rate of entry.

---

shaped relationship instead.

Second, the larger stock of general-purpose knowledge can help diversifiers differentiate themselves better from incumbents in the target industry. As noted in previous studies such as Pisano (2017), general-purpose knowledge can be a valuable complement to industry-specific knowledge and allows the firm to combine them in different ways. Given this access to a greater source of variety, diversifiers would find it easier to differentiate their products and shield themselves from competition. Thus, the differentiation effect will become flatter overall. In addition, since diversifiers would tend to be better at differentiating than spinouts, it is less likely that there will be intervals in which there is a dramatic change in the magnitude of the differentiation effect. In other words, the differentiation effect will also tend to become less concave or less convex overall, making it closer to a straight line in shape. This would tend to flatten the curve of the final relationship relative to that of the spinout, if it is originally an IUS (inverted U shape).

Spinoffs are those entrants that enter the same industry as their parent. This implies that, according to studies such as Sakakibara and Balasubramanian (2020), these firms have a much better knowledge of potential niches within their industry of origin, making it a more attractive target when compared to other industries. In addition, spinoffs may also expect to enjoy larger marginal knowledge effects in their industry of origin compared to other industries, which would make them even less attractive as targets. Consequently, the knowledge effect will become dominant only with a much higher level of relatedness. At lower levels of relatedness, the differentiation effect will dominate instead.

Summarizing this discussion, if the relationship between knowledge relatedness and rate of entry turns out to be an IUS for spinouts, then the relationship for the diversifiers would tend to be flatter for diversifiers and spinoffs. Depending on the degree to which the knowledge and differentiation effects change for these two types, the relationship may even mutate into a U-shaped relationship.

In what follows, we introduce the four industry characteristics discussed above: knowledge modularity, innovators concentration, relative rate of change, and discontinuity. We examine how they moderate the relationship between knowledge relatedness and the rate of entry across different types of entrants.

Knowledge Modularity

The first industry-level moderator considered is knowledge modularity. Modularity has been considered a major factor in determining the capabilities of firms and even affecting industry structure (Sanchez & Mahoney, 1996; Argyres & Bigelow, 2010; Campagnolo & Camuffo, 2010). Modularity can be defined at many levels, but in general this refers to the degree at which functionality corresponds to a component and whether the interfaces between the components within the product architecture are de-coupled (Ulrich, 1995; Brusoni & Prencipe, 2001). One example can be found with the semiconductor industry, where the emergence of fabless firms can be largely attributed to the loosening of the bonds between design and fabrication (Ernst, 2005). Another can be found in the computer industry, where the IBM's System/360, considered the first "truly modular" computer, allowed many peripheral devices to be added and substituted without too much difficulty (Baldwin & Henkel, 2015). In the context of knowledge, modularity essentially describes the extent to which different "modules" of knowledge are coupled with each other (Brusoni & Prencipe, 2001; Yayavaram & Ahuja, 2008).

Modularity and relatedness- Spinouts. One implication of higher modularity is that it facilitates learning for the firm, at both the component and the architectural level (Sanchez & Mahoney, 1996). In this study, this could be interpreted as the stock of knowledge required tending to be lower. More specifically, when modularity is high firms have less need of the knowledge of the system as a whole since there are fewer interdependencies between the components they need to worry about, and they can simply mix and match existing components without having to come up with a new design (Fleming & Sorenson, 2001; Ethiraj, Levinthal, & Roy, 2008; Argyres & Bigelow, 2010). At the same time, however, the actual changes gained from these new combinations will tend to not have large functional differences (Fleming & Sorenson, 2001; Fleming & Sorenson, 2003), effectively limiting the extent to which differentiation can be achieved. In other words, while achieving differentiation is easier, the degree of differentiation that is attainable would also be lower.

This would first imply that the knowledge effect will tend to be larger for the same level of knowledge relatedness when knowledge modularity is high. In addition, the marginal knowledge effect will tend to be smaller overall, especially at higher levels of knowledge relatedness, since there is no longer any synergy to be achieved in possessing different modules of knowledge. In other words, the curve that describes the knowledge effect will become more concave.

At the same time, the differentiation effect will also tend to become smaller overall. Another implication is that differentiation becomes easier to achieve, as entrants can simply change a part of the system without having to worry about what happens to the other parts (Ethiraj et al., 2008; Sanchez & Mahoney, 1996). In this case, the magnitude of the differentiation effect would tend to be smaller overall. In addition, since differentiation can simply target parts of the system, the marginal differentiation effect will not vary as much with changes in relatedness, which will tend to make the curve describing the differentiation effect closer to a straight line.

In summary, under higher knowledge modularity, the curve that describes the relationship between knowledge relatedness and the rate of entry will tend towards becoming an IUS. However, this is dependent on whether the change in the knowledge effect is more dramatic than that of the differentiation effect, especially if the curve describing the differentiation effect at zero modularity is concave. Since the range at which the change can occur is higher for the knowledge effect, we may suppose that the IUS outcome is likely. The curve would also shift upwards simply because higher modularity tends to make entry easier overall.

Modularity and relatedness- Diversifiers. The change in magnitude or the concavity of the curve describing the differentiation effect should be less dramatic for diversifiers compared to spinouts. This is because, for the same level of relatedness, diversifiers are less affected than spinouts by the differentiation effect given their larger stock of general-purpose knowledge they have compared to spinouts with the same level of relatedness. If the change in the knowledge effect remains the same for diversifiers as well, the curve for diversifiers should tend towards an IUS but to a lesser degree. In other words, the curve describing the relationship for the diversifiers would tend to become a steeper IUS with higher modularity, but still flatter than that for spinouts.

Modularity and relatedness- Spinoffs. For spinoffs, an interesting difference may come into play. First, while the curve describing the knowledge effect will likely become more concave as well, the change should be more gradual. This is because spinoffs will tend to be much better at exploiting the existing complementarity between the units of knowledge specific to their own industry compared to other spinouts. In other words, for spinoffs in industries with high relatedness to the target industry, the degree to which the curve describing the knowledge effect becomes concave should be lower. In addition, since spinoffs with high level of relatedness will tend to be more knowledgeable about the possible niches in the target industry, the curve describing the differentiation effect will tend to appear relatively convex. In summary, the tendency for the IUS to become steeper will be even weaker for spinoffs. In fact, the curve for spinoffs may even become U-shaped depending on whether the change in the differentiation effect is more dramatic than the change for the knowledge effect.

Innovators concentration

In an industry in which innovations are concentrated in the hands of a few innovating firms, incumbents would tend to have a relatively larger stock of knowledge. This implies that potential entrants end up having to compete with firms with high knowledge stock which will likely discourage entry, especially if the potential entrants come from smaller or less concentrated industries.

High concentration also implies that incumbents may have saturated most of the potential technological niches in the product space of the target industry. If this is the case, entry is more difficult, since there would be very little room for differentiation for the new entrants. This would add to the entry deterring effect of innovator concentration.

Innovators concentration and relatedness- Spinouts. The knowledge effect would imply that as concentration increases, the curve describing the relationship between relatedness and rate of entry would tend to shift downwards. At the same time, the curve that describes the knowledge effect may actually become more convex, since the potential entrants will require more deployable knowledge in order to compete effectively with the incumbents.

An interesting point to note is that the differentiation effect becomes less relevant since there are fewer niches in the product space for the new entrants to enter. In other words, the magnitude of the differentiation effect would tend to decrease overall. In addition, the marginal differentiation effect will also tend to be smaller, especially at higher levels of relatedness, since the degree to which additional general-purpose knowledge will lead to better differentiation will be limited. This would cause the curve describing the differentiation effect to become generally less concave, leading to a flattening of the curve that describes the relationship between relatedness and the probability of entry. The curve might even become U-shaped.

Innovators concentration and relatedness- Diversifiers. The entry deterring effects of innovator concentration will tend to be smaller for diversifiers, simply because they would tend to be larger as well. In particular, they would have access to the knowledge stock of their parents, which may be comparable in size to the stock of incumbents in the target industry. This would imply that the change in the knowledge effect will tend to be much less dramatic for diversifiers. This would imply, overall, that the relationship between relatedness and the probability to enter will change at a lesser degree than for spinouts.

Innovators concentration and relatedness- Spinoffs. For spinoffs, the curve describing the knowledge effect would behave similarly to the one for spinouts, with the exception that it would start off from being closer to a convex shape. The curve that describes the differentiation effect, on the other hand, will tend to be more convex compared to spinouts, since spinoffs from highly related industries will tend to be comparatively better at identifying niches than spinouts. There is unlikely to be a large difference here, however, since the availability of niches are low to begin with. This would imply that overall, the relationship will tend to become less U-shaped than spinouts.

Relative rate of change

Another factor that may affect entry is the rate of change in the target industry (Ganco & Agarwal, 2009; Gort & Klepper, 1982; Klepper, 1996). In our framework, the rate of change in the target industry is likely to affect entry through a change in the industry-specific knowledge. More precisely, confronted with changes in the industry, a firm must adjust its overall stock of knowledge to avoid making the level of industry-specific knowledge it holds less and less relevant over time.

For a firm to stay competitive in the industry, the firm would need to reconfigure itself so that it can at least maintain its stock of industry-specific knowledge. The firm would also have to reconfigure itself so that it actually keeps up with the changes in the industry in order to maintain its performance. This line of reasoning is partly inspired from Zott (2003), which shows that the timing of resource (re)deployment is an important factor explaining heterogeneity in firm performance. There, the mechanism through which timing affects firms is explained as a kind of race towards the optimal deployment of resources, which may not necessarily stay constant.[35] In essence, firms that reconfigure themselves more often are likely to perform better for at least some periods of time, which is thought to be especially true where the environment is more dynamic (D'Aveni, Dagnino, & Smith, 2010).

One possible implication of this argument is that, depending on the rate of change in their industry of origin, entrants may differ in their familiarity with a slow or fast timing of decisions. At the very least, it seems reasonable to suppose that entrants that come from an industry with a lower rate of change are slower in their decision making, while the opposite holds for firms from an industry of origin with a high rate of change.

Relative rate of change and relatedness- Spinouts. The relative rate of change may moderate the effect of relatedness on entry, especially if we visualize change as changes in the knowledge of the industry. In an industry that has a high rate of change, the level of relatedness of a given stock of knowledge to the industry would only be relevant for a short period of time. In other words, any initial benefit from high relatedness would tend to be

[35] The analysis in Zott (2003) is based on an industry that is stable over time in terms of technological change, etc., which leads to the eventual convergence in firm performance. However, the study does suggest that if this is not true, then firm heterogeneity would actually persist.

short-lived in fast-changing industries. This would be especially true if the potential entrant comes from an industry that has a slower rate of change, since they would also tend to be slower in reconfiguring themselves. In other words, the magnitude of the knowledge effect would tend to decrease overall as the relative rate of change increases. It is also likely that the curve describing the knowledge effect will tend to become more concave as well, as the benefits of additional industry-specific knowledge are likely to hit the limit more quickly.

This would imply that the relationship between relatedness and entry changes so that the curve describing it would shift downward for potential entrants that come from an industry where the rate of change is slower. In addition, the relationship would tend towards becoming more IUS the higher the relative rate of change of the target industry.

Relative rate of change and relatedness- Diversifiers. To compete effectively in the target industry, potential entrants should maintain their stock of industry-specific knowledge so that they keep up with the changes in the industry. Much like in Pisano (2017), this means that the firms need to somehow (re)create, adjust, and reconfigure their industry-specific knowledge to maintain its relevance to the changes underway in the target industry. This process would require the use of general-purpose knowledge to obtain new industry-specific knowledge or deepen/broaden their general-purpose knowledge in order to obtain the necessary knowledge. Of these, the former strategy would tend to be less costly than the latter.

In this context, it would be reasonable to suppose that since diversifiers would tend to have a larger stock of general-purpose knowledge than other entrants, they will find it much easier to enter the industry compared to spinouts, since they will find it easier to reconfigure themselves to keep their stock of industry-specific knowledge relevant. In other words, any change in the knowledge effect will tend to be less dramatic for diversifiers, suggesting that while the relative rate of change may cause the curve describing the relationship for diversifiers to become more IUS, this tendency will tend to be smaller.

Relative rate of change and relatedness- Spinoffs. For spinoffs, the curve describing the knowledge effect already tends to start out relatively more convex than in the other cases. As the relative rate of change increases, the knowledge effect at lower levels of relatedness will be relatively unaffected since the effects are already small. The change will also tend to be smaller at higher levels of relatedness, but for different reasons. Essentially, at higher levels of relatedness, spinoffs will tend to better able to keep up with the rate of change in the target industry than spinouts because any stock of general-purpose knowledge they have is likely to be more complementary with the industry-specific knowledge they already possess. Thus, the effect of this moderator on the relationship between relatedness and the probability to enter for spinoffs will tend to be less pronounced than that for spinouts.

Discontinuity

The arrival of a technological and/or market discontinuity can be conceptualized as an event that levels the playing field for all players, especially if the source of the discontinuity is exogenous. More specifically, we can think of this as a situation where the stock of industry-specific knowledge is reset to zero for all firms. This also includes incumbents, which means that, as far as their stock of industry-specific knowledge is concerned, they may become indistinguishable from potential entrants.

Discontinuity and relatedness- Spinouts. One of the major consequences of a discontinuity is that the previous stock industry-specific knowledge the firm held is no longer relevant, leading to a reset of the level of relatedness of the firm to the target industry. However, a discontinuity does not necessarily render all previous knowledge useless. In fact, depending on the type of discontinuity, firms might even end up with an increased level of relatedness.

The point, in any case, is that in the presence of discontinuity, potential entrants would tend not to know what their new level of relatedness is. In other words, the curves describing the knowledge effect and differentiation effect would both effectively become flat. In this case, the relationship between relatedness and rate of entry would cease to exist, since for all intents and purposes relatedness is not defined for the potential entrant. Thus, the overall relationship between relatedness and the probability to enter would become flatter, or even downright disappear.

Discontinuity and relatedness- Diversifiers. While a discontinuity affects the relevance of the industry-specific knowledge, it does not wipe out the differences among potential entrants in terms of general-purpose knowledge. In fact, the general-purpose knowledge stocks are still preserved for all entrants and incumbents, since this is not immediately affected by discontinuities in technology or the market. In addition, the general-purpose knowledge may make it easier for the potential entrant to identify the new industry-specific knowledge more quickly, which may mitigate the resetting effect of the discontinuity.

For diversifiers, this would mean that while the curve describing the knowledge effect may become flatter, this is likely to happen in a less dramatic way. The curve describing the differentiation effect would also behave similarly. Thus, we may suppose that while the curve describing the overall relationship for diversifiers would also tend to become flatter, it would do so at a lesser degree than spinouts.

Discontinuity and relatedness- Spinoffs. For spinoffs, the loss of relevance of the industry specific knowledge would tend to be higher for those that had originally had high relatedness with the target industry. On the other hand, for spinoffs in industries of low relatedness, the target industry is now no less attractive than its own industry of origin. As such, the curve describing the knowledge effect would tend to become more concave for the spinoffs, causing the overall relationship to become more IUS.

Table 2 summarizes the expected coefficients of the four moderators for each of the three types of entrants. While omitted from the table, note please that in general the coefficients for the interaction terms with the relatedness will tend to be opposite to that with squared relatedness. This is because otherwise it is likely that the resulting relationship will not be meaningfully different from a linear relationship.

[INSERT TABLE 2 ABOUT HERE]

EMPIRICAL ANALYSIS

The creation and the processing of the database

The empirical analysis is based on a large original dataset on the American electronics industry over a period of thirty years, from 1960 to 1990. The dataset is based on the collection and elaboration of information from annual directories published by the Western Electronics Manufacturers Association (WEMA), an organization that collects the membership of companies that operated in the electronics industry since its inception.[36] These directories are only available in paper form, so the directory had to be scanned page by page, then processed via an OCR (Optical Character Recognition) process and to extract the text into an Excel format, which was then cleaned for potential errors.

Based on this dataset, we obtained information on the company name, location, year of establishment, number of employees, companies' products, and information on executive and

[36] We thank Leslie Berlin, the Project Historian for the Silicon Valley Archives at Stanford University, for pointing to our attention the existence of this invaluable source.

board members. It also allowed us to determine the entrant type, since the directory explicitly states whether a company is a subsidiary or division of an existing company.

We also obtained information on the companies' founders and the associated parent companies, which involved extensive web searching. Various sources such as newspapers.com were used, where the founders were mentioned in association with the company in news articles, obituaries, etc. This also provided information on where the founders originally worked, identifying the parent company. Sources such as ORBIS were also used in order to determine the SIC of the companies and the parent companies[37].

The identification of unique SIC codes for companies was not always available, especially when they were divisions of an existing company. To address this issue, the product data which the directory is very rich in, was used extensively. The products were classified into a smaller number of categories via a semi-supervised learning approach based on a kNN classifier relying on the Levenshtein distance between product names and the rate of cooccurrence in patents, and the classifications were then associated with suitable SIC codes. The classification itself is a variation of the framework proposed by Saviotti & Metcalfe (1984), assigning to each product a set of three categories: technical, service, and market.

Method

One of the main issues in empirically analyzing firms' entry decisions is that, under normal circumstances, we can only observe the firms that entered ex-post, as opposed to those that could have entered ex-ante but did not. This issue often makes regressions that take entry as

[37] Parent companies are companies from which the founder originated. For example, if a founder used to work in Company X before founding Company Y, then Company X is the parent company of Company Y.

the dependent variable uninformative, since by definition all firms in the sample have entered.

To circumvent this problem, this study treats all firms established in each year as the potential entrant to an industry. Essentially, the empirical analysis is framed as determining whether firms decide to enter the chosen target industry over others, given their level of knowledge relatedness to the target industry based on their industry of origin, using the logit model. Thus, the actual dependent variable becomes a binary variable that equals 1 if the firm is established in the chosen target industry and 0 if it is established in a different industry. Thus, the dependent and explanatory variable, especially relatedness, varies depending on which target industry the analysis focuses on, which in this study is the semiconductor industry.[38]

Dependent variable

The main dependent variable is the entry into the semiconductor industry. The value of this variable is equal to 1 if the company has commercialized any product that is associated with the semiconductor industry. As was explained previously, whether a product is associated with the industry is defined based on the classification results by a semi-supervised learning algorithm, which assigned technical, service, and market categories to each product.

Explanatory variables

Knowledge relatedness. Our main explanatory variable is the measure of knowledge relatedness of the potential entrant to the target industry. The measure is based on the

[38] This approach is inspired by studies such as Cockburn & MacGarvie (2011), for example, although in this case the framing instead was that all firms in the sample were potential entrants to a segment over the sample period, effectively changing the question to *when* the firms enter.

indicator proposed by Chang, Eggers, & Keum (2021) for technological relatedness, which in turn is based on the measure pioneered by Breschi, Lissoni, & Malerba (2003). In our study, however, this measure is computed at the industry-dyad level, since we determine the potential entrants' pre-entry knowledge by their industry of origin as determined by their parent company.[39] This choice also mitigates potential problem with the original measure, which is the fact that it is based on the use of patents, which not all companies possess. While there is also the problem that patents are not always the most representative of the technological knowledge in the industry, the fact that this study tends to focus on the electronics industry should alleviate this problem as this industry is known to rely substantially on patents.

The measure in this study, like the original one in Chang et al. (2021), is calculated in two steps[40]. First, we determine the relatedness between dyads of CPCs, based on the similarity of citation patterns of the patents within each CPC (Cooperative Patent Classification) class, as in Breschi et al. (2003). This is calculated annually, using patents granted between 1 to 3 years prior.

Then, after determining the CPCs of patents in each SIC industry, the knowledge relatedness between a pair of SICs is determined by aggregating these similarities. This is where our measure diverges from the original one in Chang et al. (2021), in that instead of taking the maximum similarity among the possible combinations of CPCs across the two industries, we compute the relatedness between industries i and j as follows.

[39] In our study, the industry of origin refers to the industry of the parent company. For diversifiers, subsidiaries, or divisions, this is the industry in which the firm were originally active. In the case of spinoffs or spinouts, it is the industry where the founder was previously employed.

[40] See Chang, Eggers, & Keum (2021), especially the online appendix, for a more thorough description of how the original measure is calculated.

$$R_{ijt} = \begin{cases} \frac{1}{n(C)} \sum_{c \in C} \max_{\mathrm{d} \in \mathrm{D}} (S_{cdt}) \cdot (1 - Generality_{ct}), & i \neq j \\ 1, & i = j \end{cases} \qquad (2)$$

Here, S_{cdt} refers to the similarity of citation patterns between CPC $c \in D$ and $d \in D$, where C and D refer to the set of CPCs associated with industry i and j respectively, evaluated in year t. Note that relatedness simply equals 1 if it is calculated for the same industries.

Generality is calculated as suggested by Hall, Jaffe, & Trajtenberg (2001), but extended so that the generality of the CPC is calculated instead of the individual patent. The exact calculation is based on the following formula.

$$Generality_{ct} = 1 - \sum_{i} s_{cdt}^{2} \qquad (3)$$

Here, s_{cit} refers to all forward citations made to any patent in CPC c up until 5 years after year t by all patents belonging to CPC d. Thus, generality is essentially the reverse of the Herfindahl index of CPCs of citing patents, such that the higher this value, the higher the generality of the CPC. The similarity of the CPC with higher generality is given a lower weight in the calculation of relatedness.

The reasons behind this approach are twofold. First, we ultimately want the indicator to match our theoretical framework as closely as possible. One of the key features in our definition of relatedness is that it depends heavily on there being general-purpose and industry-specific knowledge. The measure treats this distinction as a continuum.

Second, using the original measure (which relies on maximum CPC-level similarity) often has the tendency to label many pairs of SICs to be highly related to each other, especially since sharing even one CPC class automatically causes the value to be equal to 1. This would tend to artificially reduce the variance in the measure, making it difficult to interpret the results in the context of our theoretical framework.

There were also cases where the relatedness measure simply could not be defined. There were two situations in which this might occur. The first occurred when patents simply did not exist in the industries in the specific time period. The second occurred when patents existed, but the generality measure could not be defined because the patents were not cited within the 5-year time frame. In the main analysis, the value for the variable was treated as zero for both cases, as both represent situations where no useful knowledge can be brought to the target industry from the industry of origin.

Knowledge Modularity. The measure for the modularity of knowledge in the industry is based on the methodology proposed by Yayavaram & Ahuja (2008).[41] The measure captures knowledge decomposability, which is considered a good proxy for modularity by the authors. The original method calculates the value at the firm level but in our paper, the value is measured at the industry level instead, since we are more interested in the level of knowledge modularity in the target industry (in this case, the semiconductor industry).

The measure is based first on the identification of all patents assigned to the focal SIC industry, 3674 in this case, within the previous 3 years. Then, the degree of coupling between a pair of CPCs, or the strength of ties between the two CPCs is calculated for each pair. This is calculated as the proportion of patents that are assigned both CPCs relative to the number of patents that are assigned either of the CPCs. Equation (4) below describes this for CPCs i and j at time t.

$$Coupling_{i,j,t} = \frac{N[(\text{Patents in CPC } i \text{ at time } t) \cap (\text{Patents in CPC } j \text{ at time } t)]}{N[(\text{Patents in CPC } i \text{ at time } t) \cup (\text{Patents in CPC } j \text{ at time } t)]} \quad (4)$$

[41] Other examples of measures of modularity based on patents exist as well, such as the one proposed by Fleming & Sorenson (2001), which is calculated at the patent level and looks at the average number of sub-classes that appear with the sub-class assigned to each patent. This study chooses to use the measure proposed by Yayavaram & Ahuja (2008) because arguably reflects better the extent to which patent classes are interdependent to each other.

Then for each CPC i, the level of integration is calculated. This is defined as the proportion of strong ties, or ties with a level of coupling higher than the median relative to total possible number of ties among CPC i and all CPCs that have any level of coupling with i. This is expressed by Equation (5) below for CPC i at time t, supposing that there are $m-1$ CPCs with which CPC i has any level of coupling.

$$Integration_{i,t} = \frac{N(Coupling_{i,j,t} | Coupling_{i,j,t} \geq median(Coupling_{i,j,t}))}{\binom{m}{2}} \quad (5)$$

This is then summed over all CPCs, weighted by the percentage of patents assigned to each class to give the modularity of the semiconductor industry at time t, as expressed in Equation (6) below.

$$Modularity_t = 1 - \sum_i \frac{N(Patents\ assigned\ CPC\ i)}{\sum_k N(Patents\ assigned\ CPC\ k)} \cdot Integration_{i,t} \quad (6)$$

Relative rate of change. The rate of change in the industry is defined as the rate at which the number of patents increases over time. This is based on the notion that each patent is meant to represent novel knowledge, thus capturing the change in the knowledge of the industry. The rate of change in industry i calculated for year t is as follows.

$$Rate\ of\ Change_{i,t} = \frac{Number\ of\ Patents_{i,t}}{\sum_{t-10}^{t-1} Number\ of\ Patents_{i,\tau}} \quad (7)$$

The number of patents granted in a year is compared to the number of patents that were granted in the last 10 years. The relative rate of change, the variable used in the analysis, for a firm originating from industry i is defined as follows.

$$Relative\ Rate\ of\ Change_{3674,i,t} = Rate\ of\ Change_{3674,t} - Rate\ of\ Change_{i,t} \quad (8)$$

Innovators concentration. The innovators concentration is calculated as the Herfindahl index (HHI) of the industry, based on the share of patents held by each firm in the industry

(Breschi et al., 2000). HHI is calculated on an annual basis, like the other moderators in this analysis.

Discontinuity. The discontinuity is simply represented by a dummy that equals 1 if the entry occurred after 1971. This year is based on the characterization of a major technological discontinuity in the semiconductor industry, as in studies such as Anderson & Tushman (1990) and Malerba et al. (2008). In 1971, microprocessors were introduced in the industry.

Entrant type. The other main independent variables are dummies that define whether the potential entrant is a diversifier or a spinoff. The diversifier dummy equals 1 if the potential entrant is a subsidiary or division of another company, regardless of which industry that company is in. The spinoff dummy is equal to 1 if the company is not a diversifier or a plain start-up, and if the SIC of the parent company is associated with any of the products produced by the company, regardless of what that SIC is. They may still eventually decide to enter the same industry as their parent. For example, for a company whose parent company was in the SIC industry 3571 (computers) and entered in that same industry, the value of the spinoff dummy is equal to 1, regardless of whether they also produce products associated with SIC 3674 (semiconductors).

Note that, in contrast, relatedness is calculated between the parent company's SIC and the target SIC, unlike the spinoff dummy. So, for the same company described previously, the value of the relatedness to the target industry, SIC 3674 for example, has nothing to do with the fact that the company is a spinoff in SIC 3999. The same goes for the diversifier case, where the relatedness between the parent and the actual industry it entered is not a factor.

The econometric analysis also contains some control variables: location, the log of the number of employees of the company, the diversity of the SICs associated with the products produced by the company, and the age of the parent company.

Descriptive statistics

The entire dataset consists of 14,774 unique firms. Of these, 7,668 firms are non-startups. As Table 3 shows, about 4,550 are divisions or subsidiaries, while 3,118 are spinouts or spinoffs.

[INSERT TABLE 3 ABOUT HERE]

Table 4 shows the descriptive statistics of the variables used in the analysis.

[INSERT TABLE 4 ABOUT HERE]

The share of each type of entrant varies by year, as can be seen from Figures 2 and 3. In both figures, each bar represents 5-year periods such as 1961-1965 or 1986-1990 due to presentation issues. Figure 2 shows that the total number of entrants tend to show an increase over time, while Figure 3 shows that the proportions of each type of entrants fluctuate.

[INSERT FIGURES 2 AND 3 ABOUT HERE]

Figure 4 shows the average level of relatedness between the actual entrants and non-entrants to the semiconductor industry. The entrants are defined as companies that produce any product that corresponds to the industry, which includes products such as microcircuits and microprocessors. One may note that, perhaps surprisingly, that there tends to be little difference in the relatedness between the two groups, suggesting that if there is a relationship, it is not likely to be linear.

[INSERT FIGURE 4 ABOUT HERE]

Results

Table 5 shows the results of the analysis on the semiconductor industry. While the first specification does not include moderators, every other specification in Table 5 involves one of the moderators, including their interaction with relatedness.

[INSERT TABLE 5 ABOUT HERE]

The results are different depending on the moderator involved. Model 2 suggests that higher modularity tends to cause the relationship between relatedness and rate of entry to change from a flat line to an upward line for spinouts. Model 4 implies that higher innovator concentration tends to cause the curve to become steeper. Model 3 and 5, on the other hand, suggest that relative rate of change and discontinuity respectively have no real effect on the relationship.

One of the reasons why these results are observed can be related to the different effect of moderators depending on the type of entrants, as suggested by our theoretical discussion. Table 6 shows the results of re-running the analysis, taking into account the distinction between each potential entrant.

[INSERT TABLE 6 ABOUT HERE]

Interestingly, most of the results that are significant are in line with the predictions summarized in Table 2, except for discontinuity, which turns out be positive instead.

It has to be noted, however, that since the relative rate of change indicator has a fairly large range of values while the opposite is true for modularity, the coefficients are perhaps not well-suited to describe the effects of the moderators. Table 7 presents the results from the same specification as Table 6, except that modularity is rescaled so that its range is now from 0 to 1, and the relative rate of change is now defined as a dummy indicating whether the value is above zero or not.

[INSERT TABLE 7 ABOUT HERE]

The results are generally similar to Table 6. They also show that the relative rate of change now has a significant effect on the relationship for spinoffs, causing the relationship to become more U-shaped.

Figures 5 and 6 present a rough visualization of the results for modularity and innovators concentration of Table 7, as a plot of relatedness with predicted rate of entry, with the 90% confidence interval represented as well. [42]

[INSERT FIGURES 5 AND 6 ABOUT HERE]

Note that while the relations would appear as IUS, U-shaped curve, etc. when the log odds are plotted, the probability curves tend to be presented as logistic curves instead. Still, the shape around the peak or trough of the curve looks like the more familiar IUS or U-shaped curves. Therefore, focusing on those areas will make it simpler to understand the relationship.

Discussion

While Model 1 in Table 5 shows that the direct relationship is of an IUS type for spinouts and U-shaped for the other types of entrants, Models 2 to 5 in Table 1, Table 6, and Table 7 show that much of the relationship is driven by the moderators. Thus, the results as a whole seem to show some support for the theoretical framework.

Modularity seems to be the main driver of the IUS relationship between knowledge relatedness and rate of entry for spinouts, and of the U-shaped relationships for the other types of entrants. These results are mostly in agreement with the theoretical framework,

[42] The visualization for the other moderators is available upon request. They were omitted because they tend to be not significant in the regression to begin with.

although for spinoffs the effect is more extreme than expected. This may be an indication that either the change in the knowledge effect is much smaller than expected, or that the differentiation effect changes more dramatically than expected. For example, the differentiation effect may change from being concave to being convex for diversifiers and spinoffs, which would then cause the relationship between relatedness and entry to become more U-shaped for these two types of entrants. This may occur if diversifiers and spinoffs tend to be better at benefitting from modularity so that, beyond a certain level of relatedness, the differentiation effect does not increase substantially.

Also, the results for the effects of innovators concentration are in line with expectations. The result for spinoffs is slightly different from what is predicted in Table 2, but may not necessarily contradict the theoretical analysis. Specifically, the results still agree with the theoretical analysis in that under higher innovators concentration, the relationship between relatedness and rate of entry tends to become more U-shaped for spinoffs. The difference may be due to the fact that the extent to which this occurs is not so different from spinouts (at least according to Table 7). While the theoretical prediction was that spinoffs will be similarly affected by innovators concentration but to a smaller degree than spinouts, the results suggest that differences are insignificant. In other words, the changes due to the knowledge and differentiation effects in terms of their concavity appear to be very small, if there are any changes at all. On the other hand, results show that the curve for spinoffs is shifted to the left relative to that of spinouts, which suggests two not mutually exclusive possibilities. First, the marginal knowledge effect for spinoffs may be larger than that for spinouts. This may be the case if spinoffs tend to be more effective in deploying the industry-specific knowledge compared to spinouts. Second, the marginal differentiation effect may be smaller for spinoffs than spinouts. This would correspond to situations where spinoffs tend to have less trouble in identifying niches to differentiate into compared to spinouts.

On the other hand, the results for both the relative rate of change and the discontinuity seem to have a significant effect only on spinoffs, at least according to Table 7. More importantly, both show the opposite sign than expected. Given that the corresponding coefficients are not significant for the other types, however, this may be an indication that the expected changes in the knowledge and differentiation effects apply only to spinoffs and not the others. One possible explanation may be that spinoffs preserve their level of relatedness to the target industry once they enter it, while the other types are less concerned. Therefore, the knowledge effect only changes for the spinoffs, since for the other types of entrants the possibility that their current stock of industry-specific knowledge loses relevance is not of concern.

Similar logic can be applied in the explanation of the estimated coefficients for discontinuity. In fact, it may be the case that the other types of entrants are not even aware that a discontinuity has occurred, at least not enough to actually consider it in their entry decisions.

One interesting observation in the results is that changing the measure for the relative rate of change to a binary variable affects not only the results for relative rate of change, but also innovators concentration. A particular point to note is that while Table 6 suggests that the relationship between relatedness and rate of entry changes into an IUS for spinoffs under higher innovators concentration, Table 7 suggests that there is only some left-shifting of the U-shaped curve for spinoffs. This may be an indication that the effect derived from from the relative rate of change is somehow related to the effect from innovators concentration. One possibility in this regard is that for both factors, the main source of change might be the change in concavity of the differentiation effect.

One of the main mechanisms through which innovators concentration was considered to affect the relationship between relatedness and the rate of entry is that the knowledge effect tends to become more convex due to the higher requirements of industry-specific knowledge in order to compete effectively. While this is not expected to happen in general when the relative rate of change is considered, spinoffs may be in a different situation. In fact, one of the expected consequences of a higher relative rate of change is that the knowledge effect curve will tend to be convex compared to spinouts, the logic being that any general-purpose knowledge held by spinoffs from highly related industries tends to be more complementary to their industry-specific knowledge. If something similar is occurring when innovators concentration changes, this reasoning may explain why we observe the change in the results.

We may then suggest that, as in the case of Table 6, the effect of the complementarity between industry-specific and general-purpose knowledge is captured mostly by innovators concentration, because the relative rate of change has a comparatively higher variance and therefore is less visible. Once the relative rate of change is turned into a binary variable, however, the variance is no longer so large, allowing the relative rate of change to capture most of the effect.

CONCLUSION

This paper presented an exploratory attempt to shed light on the role of industry characteristics, such as modularity, innovators concentration, rate of change, and discontinuity, in affecting the relationship between relatedness and entry for diversifiers, spinouts, and spinoffs. Two significant results from our analysis may be pointed out. The first is that industry characteristics matter in affecting the overall relationship between knowledge relatedness and entry. Ignoring some key industry features of the industry may lead to an incomplete understanding of the role that context plays in affecting the way relatedness affects entry decisions.

The second is that the effects of the industry moderators differ across types of entrants. Overlooking the heterogeneity of entrants may lead to a lack of awareness that the roster of entrants may vary in terms of diversifiers, spinoffs, and spinouts. This, in turn, may cause us to ignore the role this composition plays in understanding the effect of knowledge relatedness on entry in a context with certain industry characteristics.

Some results in the econometric analysis of this paper show support for the premise that certain industry-level moderators have a significant impact on the nature of the relationship between knowledge relatedness and firm entry, while others do not. This may depend on the type of industry selected for the analysis (semiconductors), as well as on the construction of the indicators to proxy for some of the chosen moderators. This calls for more work on the conceptual framework as well as more in-depth and finer-grained analysis of the data. It is also possible that perhaps not all industry-level moderators are necessarily significant, but confirming this will require more in-depth analysis of the data.

Given that this paper is a work in progress in rather unexplored terrain, the theory needs refinement in terms of consistency of logic, while the empirics would benefit from the identification of alternative measures or better specifications. In particular, one may argue that certain variables used in the current analysis, such as relative rate of change or discontinuity, could benefit from more robust measures.

Despite these limitations, this paper represents a novel and interesting line of research, because solid results for the industry-level moderators may help significantly clarify the reasons behind why patterns of entry tend to vary in composition and across industries and time. This in turn would have important implications for potential entrepreneurs and existing firms, since a better understanding of how the various mechanisms work would help them identify the competitive environment they can expect in an industry. Also, the policy implications might be significant, as the mechanisms examined in this paper identify the factors that would encourage or discourage entry.

REFERENCES

Adams, P., Fontana, R., & Malerba, F. 2015. User-industry spinouts: Downstream industry knowledge as a source of new firm entry and survival. ***Organization Science***. 27 (1): 18-35.

Agarwal, R., & Gort, M. 1996. The evolution of markets and entry, exit and survival of firms. ***The Review of Economics and Statistics***. 78 (3): 489-498.

Agarwal, R., Echambadi, R., Franco, A., & Sarkar, M. 2004. Knowledge transfer through inheritance: Spin-out generation, development, and survival. ***Academy of Management Journal***. 47 (4): 501-522.

Agarwal, R., & Gort, M. 2002. Firm and product life cycles and firm survival. ***American Economic Review***. 92 (2): 184-190.

Agarwal, R., & Shah, S. 2014. Knowledge sources of entrepreneurship: Firm formation by academic, user and employee innovators. ***Research Policy***. 43: 1109-1133.

Anderson, P, & Tushman, M. 1990. Technological discontinuities and dominant designs: A cyclical model of technological change. ***Administrative Science Quarterly***. 35 (4): 604-633.

Argyres, N., & Bigelow, L. 2010. Innovation, modularity, and vertical deintegration: Evidence from the early U.S. auto industry. ***Organization Science***. 21 (4): 842-853.

Baldwin, C., & Henkel, J. 2015. Modularity and intellectual property protection. ***Strategic Management Journal***. 36: 1637-1655.

Bayus, B., & Agarwal, R. 2007. The role of pre-entry experience, entry timing, and product technology strategies in explaining firm survival. ***Management Science***. 53 (12): 1887-1902.

Benner, M., & Tripsas, M. 2012. The influence of prior industry affiliation on framing in nascent industries: The evolution of digital cameras. ***Strategic Management Journal***. 33: 277-302.

Breschi, S., Lissoni, F., & Malerba, F. 2003. Knowledge-relatedness in firm technological diversification. ***Research Policy***. 32: 69-87.

Breschi, S., Malerba, F., & Orsenigo, L. 2000. Technological regimes and Schumpeterian patterns of innovation. ***The Economic Journal***. 110: 388-410.

Brusoni, S., & Prencipe, A. 2001. Unpacking the black box of modularity: Technologies, products, and organizations. ***Industrial and Corporate Change***. 10 (1): 179-205.

Campagnolo, D., & Camuffo, A. 2010. The concept of modularity in management studies: A literature review. ***International Journal of Management Reviews***. 12 (3): 259-283.

Chang, S., Eggers, J.P., Keum, D. 2021. Bottleneck resources, market relatedness, and the dynamics of organizational growth. ***Organization Science***. Articles in Advance: 1-19.

Clarysse, B., Wright, M., & Van de Velde, E. 2011. Entrepreneurial origin, technological knowledge, and the growth of spin-off companies. ***Journal of Management Studies.*** 48: 1420-1442.

D'Aveni, R., Dagnino, G., & Smith, K. 2010. The age of temporary advantage. ***Strategic Management Journal***. 31: 1371-1385.

Demsetz, H. 1973. Industry structure, market rivalry, and public policy. ***Journal of Law and Economics***. 16 (1): 1-9.

Dunne, T., Roberts, M., & Samuelson, L. 1988. Patterns of firm entry and exit in U.S. manufacturing industries. ***The RAND Journal of Economics***. 19 (4): 495-515.

Ernst, D. 2005. Limits to modularity: Reflections on recent developments in chip design. ***Industry and Innovation***. 12 (3): 303-335.

Ethiraj, S., Levinthal, D., & Roy, R. 2008. The dual role of modularity: Innovation and imitation. ***Management Science***. 54 (5): 939-955.

Fleming, L., & Sorenson, O. 2001. Technology as a complex adaptive system: evidence from patent data. ***Research Policy***. 30: 1019-1039.

Fleming, L., & Sorenson, O. 2003. Navigating the technology landscape of innovation. ***MIT Sloan Management Review***. 44 (2): 15-23.

Fontana, R., Martinelli, A., & Nuvolari, A. 2021. Regimes reloaded! A reappraisal of Schumpeterian patterns of innovation, 1977-2011. ***Journal of Evolutionary Economics***. 31: 1495-1519.

Ganco, M., & Agarwal, R. 2009. Performance differentials between diversifying entrants and entrepreneurial start-ups: A complexity approach. ***The Academy of Management Review***. 34 (2): 228-252.

Gort, M., & Klepper, S. 1982. Time paths in the diffusion of product innovations. ***The Economic Journal***. 92 (367): 630-653.

Helfat, C., & Lieberman, M. 2002. The birth of capabilities: market entry and the importance of pre-history. ***Industrial and Corporate Change***. 11 (4): 725-760.

Helfat, C., & Raubitschek, R. 2000. Product sequencing: Co-evolution of knowledge, capabilities, and products. ***Strategic Management Journal***. 21: 961-979.

Katila, R. 2002. New product search over time: Past ideas in their prime?. ***The Academy of Management Journal***. 45 (5): 995-1010.

Klepper, S. 1996. Entry, exit, growth, and innovation over the product life cycle. ***The American Economic Review***. 86 (3): 562-583.

Klepper, S., & Sleeper, S. 2005. Entry by spinoffs. ***Management Science***. 51 (8): 1291-1306.

Krafft, J. 2004. Entry, exit and knowledge: Evidence from a cluster in the info-communications industry. ***Research Policy***. 33: 1687-1706.

Makri, M., Hitt, M., & Lane, P. 2010. Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions. ***Strategic Management Journal***. 31: 602-628.

Malerba, F., Nelson, R., Orsenigo, L., & Winter, S. 2008. Vertical integration and disintegration of computer firms: A history-friendly model of the coevolution of the computer and semiconductor industries. ***Industrial and Corporate Change***. 17 (2): 197-231.

Malerba, F., & Orsenigo, L. 1997, Technological regimes and sectoral patterns of innovative activities. ***Industrial and Corporate Change***. 6 (1): 83-117.

Pisano, G. 2017. Toward the prescriptive theory of dynamic capabilities: connecting strategic choice, learning, and competition. ***Industrial and Corporate Change***. 26 (5): 747-762.

Sakakibara, M., & Balasubramanian, N. 2020. Human capital, parent size, and the destination industry of spinouts. ***Strategic Management Journal***. 41: 815-840.

Sapienza, H., Parhankangas, A., & Autio, E. 2004. Knowledge relatedness and post-spin-off growth. ***Journal of Business Venturing***. 19: 809-829.

Sanchez, R., & Mahoney, J. 1996. Modularity, flexibility, and knowledge management in product and organization design. ***Strategic Management Journal***. 17 (Winter Special Issue): 63-76.

Schumpeter, J. 1934. ***The Theory of Economic Development***, Cambridge, MA: Harvard Economic Studies.

Schumpeter, J. 1942. ***Capitalism, Socialism, and Democracy***, New York: Harper and Brothers.

Tanriverdi, H., & Venkatraman, N. 2005. Knowledge relatedness and the performance of multibusiness firms. ***Strategic Management Journal.*** 26: 97-119.

Tushman, M., & Anderson, P. 1986. Technological discontinuities and organizational environments. ***Administrative Science Quarterly***. 31 (3): 439-465.

Ulrich, K. 1995. The role of product architecture in the manufacturing firm. ***Research Policy***. 24: 419-440.

Uzunca, B. 2011. Comparative advantages of spinoff firms: An evolutionary perspective. ***Journal of Technology Management and Innovation***. 6 (4): 80-92.

Yayavaram, S., & Ahuja, G. 2008. Decomposability in knowledge structures and its impact on the usefulness of inventions and knowledge-base malleability. ***Administrative Science Quarterly***, 53: 333-362.

Zott, C. 2003. Dynamic capabilities and the emergence of intraindustry differential firm performance: Insights from a simulation study. ***Strategic Management Journal***. 24: 97-125.

LIST OF TABLES

TABLE 1

Characteristics of each entrant type: An illustration

Entrant Type	General-Purpose Knowledge	Industry-Specific Knowledge	Total Knowledge Stock
Diversifier	[0, 100]	[0, 100]	100
Spinout	[0, 50]	[0, 50]	50
Spinoff	[0, 10]	[40, 50]	50

[x,y] denotes that the entrant types has a stock of the knowledge of the column in the range of x to y.

TABLE 2

Expected signs of major results

	Spinout	Change in Effect For	
		Diversifier	Spinoff
Relatedness	+	-	-
Relatedness^2	-	+	+
Relatedness^2 x Modularity	-	+	+
Relatedness^2 x Relative Rate of Change	+	-	-
Relatedness^2 x Innovators Concentration	+	-	-
Relatedness^2 x Discontinuity	+	+	-

TABLE 3

Entrant types, original dataset

Entrant Type	Frequency	Percentage (%)
Diversifier	4539	59.33
Spinout	2178	28.47
Spinoff	933	12.20
Total	**7650**	**100**

TABLE 4

Descriptive statistics

	N	Min	Max	Mean	S.D.
Entry	4726	0	1	0.65	0.48
Diversifier	4726	0	1	0.48	0.50
Spinoff	4726	0	1	0.17	0.38
Knowledge Relatedness	4726	0.00	1.00	0.26	0.25
Modularity	4726	0.15	0.30	0.17	0.02
Relative Rate of Change	4726	-5.45	0.28	0.02	0.16
Innovators Concentration	4726	0.02	0.69	0.23	0.14
Number of Founders	4726	1	5	1.17	0.49
Log (Number of Employees)	4726	0.00	12.39	4.52	1.61
Product Diversity	4726	0.00	0.36	0.06	0.05
Parent Age	4726	0	893[43]	29.53	46.25

[43] Note that about 94% of the parent companies in the sample are younger than 100 years. The ones that are older than this tend to be universities, government agencies, etc., although some commercial companies do exist. The maximum value, incidentally, belongs to Oxford University, whose founding year is listed as 1096.

TABLE 5

Results with moderators

	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	0.0726	-0.8840	0.0557	-0.1484	0.3326
	(0.3806)	(0.5745)	(0.3814)	(0.3830)	(0.3858)
Relatedness	2.4126***	-5.6046	2.4652***	1.5825*	2.8570***
	(0.6433)	(4.4658)	(0.6525)	(0.8203)	(0.8590)
Relatedness^2	-7.0187***	0.2581	-7.0715***	-5.8590***	-7.1747***
	(1.0837)	(5.1497)	(1.1002)	(1.0849)	(1.1245)
Diversifier	0.1806	0.1041	0.1852	0.1242	0.1093
	(0.1260)	(0.1290)	(0.1262)	(0.1274)	(0.1281)
Spinoff	1.0958***	0.9906***	1.1161***	1.0404***	1.0035***
	(0.3393)	(0.3474)	(0.3401)	(0.3481)	(0.3514)
Modularity		6.1140**			
		(2.7127)			
Relative Rate of Change			0.1964		
			(0.2948)		
Innovators Concentration				1.0454**	
				(0.4417)	
Discontinuity					-0.3822***
					(0.1263)
Relatedness x Diversifier	-3.7776***	-3.4732***	-3.8004***	-3.6084***	-3.5581***
	(0.8063)	(0.7996)	(0.8103)	(0.7868)	(0.7870)
Relatedness^2 x Diversifier	8.4430***	7.9547***	8.5136***	7.9461***	7.9295***
	(1.1925)	(1.1028)	(1.2095)	(1.0555)	(1.0439)
Relatedness x Spinoff	-9.4446***	-9.0947***	-9.5325***	-9.6081***	-9.4137***
	(2.6663)	(2.7565)	(2.6671)	(2.8239)	(2.8244)
Relatedness^2 x Spinoff	22.8359***	22.8292***	22.915***	23.8061***	23.4917***
	(5.0365)	(5.2397)	(5.0351)	(5.4568)	(5.3990)
Relatedness x Modularity		50.2489*			
		(27.6923)			
Relatedness^2 x Modularity		-43.8353			
		(31.5413)			
Relatedness x			-0.3769		
Relative Rate of Change			(3.0017)		
Relatedness^2 x			5.1371		
Relative Rate of Change			(7.0465)		
Relatedness x				5.1441*	
Innovators Concentration				(3.0830)	
Relatedness^2 x				-4.6628	
Innovators Concentration				(3.4294)	
Relatedness x Discontinuity					-0.6120
					(0.7788)
Relatedness^2 x Discontinuity					0.7905
					(0.8621)
Controls	Yes	Yes	Yes	Yes	Yes
N	4726	4726	4726	4726	4726
Log Likelihood	-2858.96	-2836.42	-2858.08	-2836.77	-2838.31
	(df = 58)	(df = 61)	(df = 61)	(df = 61)	(df = 61)

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

TABLE 6

Results with moderators, by entrant type

	Spinout	Change in Effect For	
		Diversifier	Spinoff
Constant	1.8411	-3.4623**	-0.4935
Modularity	-9.3158	19.1752**	13.4013
Relative Rate of Change	0.2911	0.2004	-0.0269
Innovators Concentration	0.4895	-0.0568	-3.2542
Discontinuity	-0.4735	0.5159	0.1198
Relatedness	-27.0149*	25.1313	25.5250
Relatedness^2	57.5134*	-56.3259	-147.7383
Relatedness x Modularity	190.7577*	-186.9544*	-270.3087
Relatedness^2 x Modularity	-415.4684*	415.5357*	1078.4160**
Relatedness x Relative Rate of Change	-5.4126	10.4828	-22.0024
Relatedness^2 x Relative Rate of Change	11.3329	-24.2289	50.6138
Relatedness x Innovators Concentration	-14.1486	18.5652	53.9648**
Relatedness^2 x Innovators Concentration	30.4851*	-33.9421*	-71.8536*
Relatedness x Discontinuity	-2.0264	-2.4127	-5.9957
Relatedness^2 x Discontinuity	-4.6484	5.3842	25.7941

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

TABLE 7

Results with moderators (alternative measures) by entrant type

	Spinout	Change in Effect For	
		Diversifier	Spinoff
Constant	0.8449	-1.3522*	0.5805
Modularity[44]	-1.6721*	3.3423***	3.0335
Positive Relative Rate of Change[45]	-0.3903	0.7647*	1.1443
Innovator Concentration	0.5929	-0.1523	-3.8843
Discontinuity	-0.4637	0.4837	0.2684
Relatedness	-1.5066	2.0864	-1.3678
Relatedness^2	1.8658	-4.5478	-23.5049
Relatedness x Modularity	30.0295**	-30.7063**	-54.9377*
Relatedness^2 x Modularity	-66.4239**	67.4931**	212.2536**
Relatedness x Positive Relative Rate of Change	1.9553	-3.9540	-12.4592*
Relatedness^2 x Positive Relative Rate of Change	-5.3437	9.1313	31.8940**
Relatedness x Innovators Concentration	-13.9753	18.8577	53.9334*
Relatedness^2 x Innovators Concentration	32.7819*	-36.4283*	-76.0335
Relatedness x Discontinuity	1.7383	-1.8596	-9.5536
Relatedness^2 x Discontinuity	-4.1666	4.6579	36.8657*

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

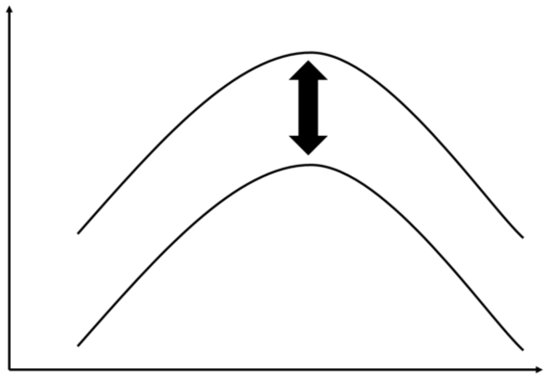
[44] In this table, modularity is rescaled so that it now takes a value between 0 and 1.
[45] This is a binary variable that takes the value of 1 if the relative rate of change is larger than or equal to 0. In other words, this simply denotes whether the rate of change in the target industry is at least as high as the industry of origin.

LIST OF FIGURES

FIGURE 1

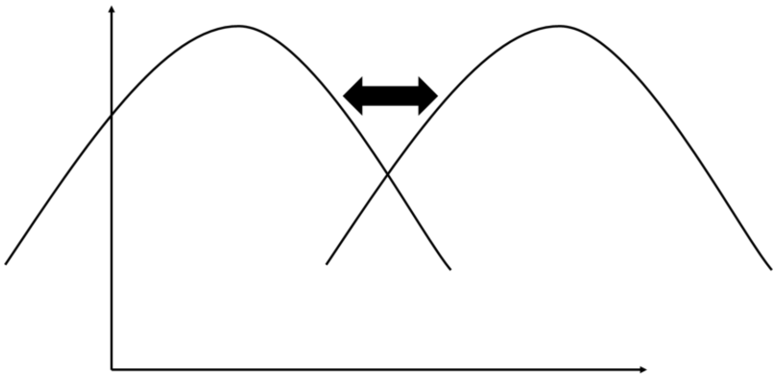
Visualization of theoretical description of changes of curves and their shapes

Upward/Downward Shift



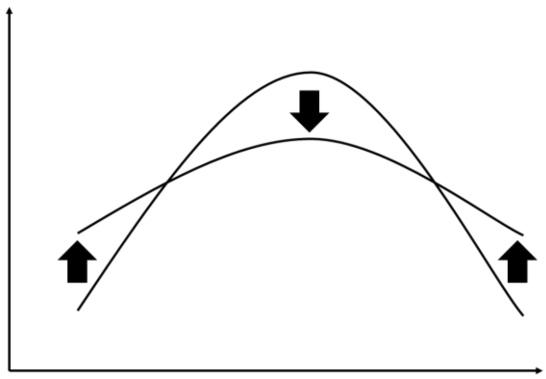
May be caused by changes in total knowledge/differentiation effect.

Left/Right Shift



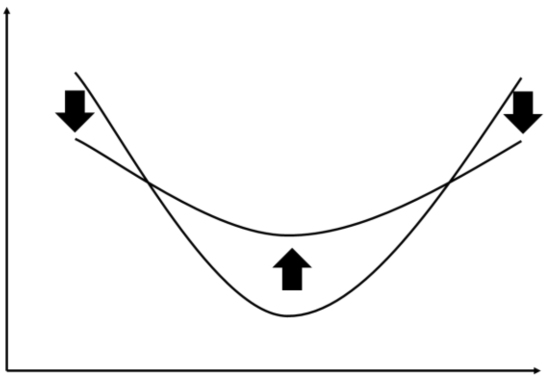
May be caused by changes in marginal knowledge/differentiation effect.

Flatten (Inverted U-shaped Curve)



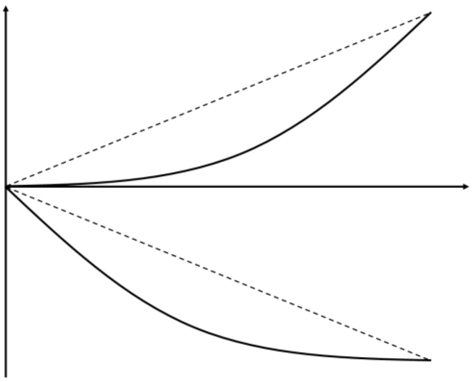
May be caused by decrease in concavity of knowledge and/or differentiation effect.

Flatten (U-shaped Curve)



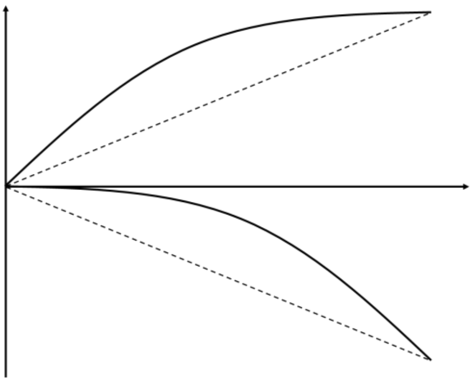
May be caused by decrease in convexity of knowledge and/or differentiation effect.

Convex (Knowledge and Differentiation Effects)



The dotted straight line represents a convexity of zero.

Concave (Knowledge and Differentiation Effects)



The dotted straight line represents a concavity of zero.

FIGURE 2

Number of entrants and composition over time

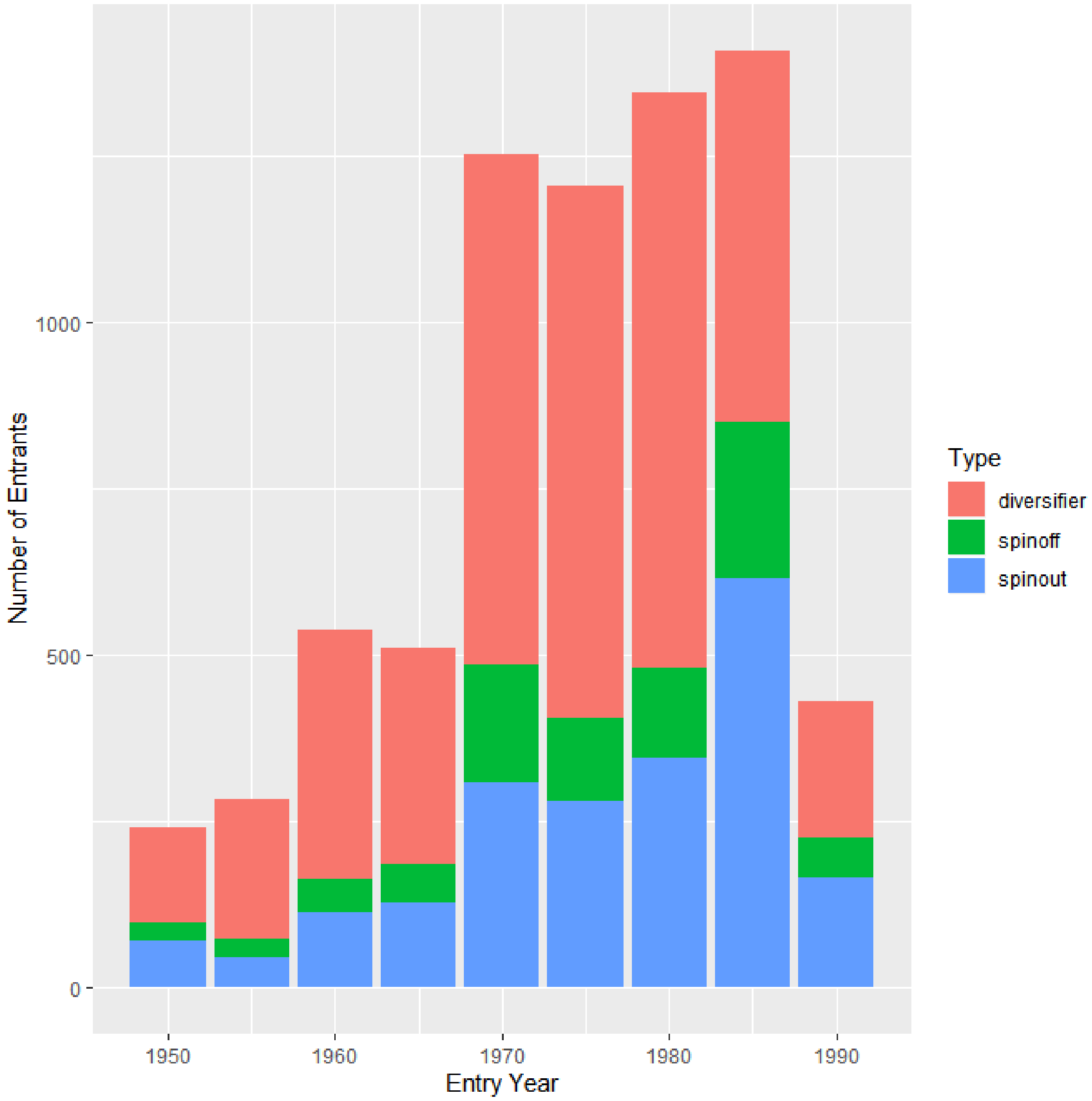


FIGURE 3

Proportion of entrant types by year

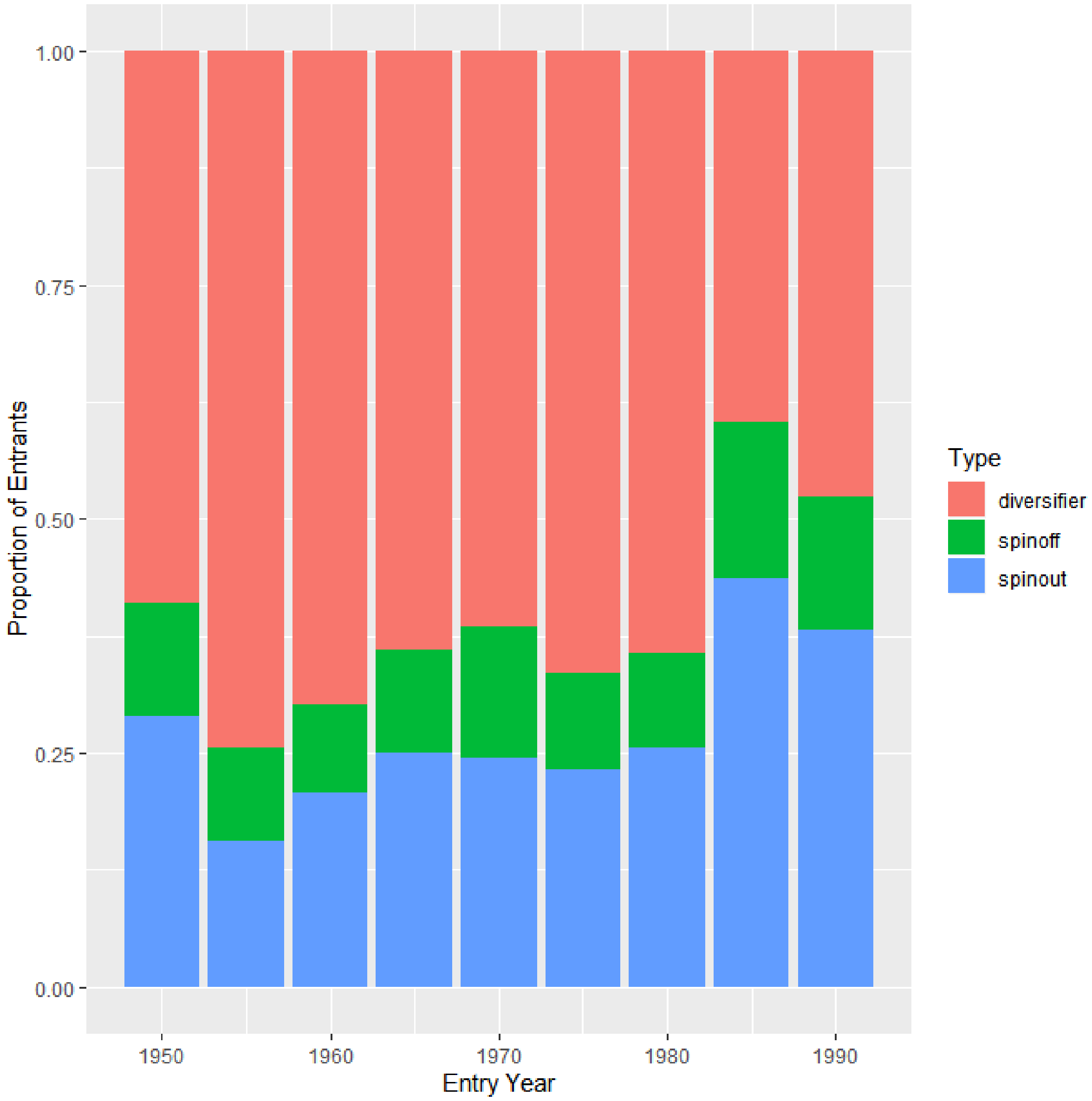


FIGURE 4

Average level of relatedness, entrant vs non-entrant

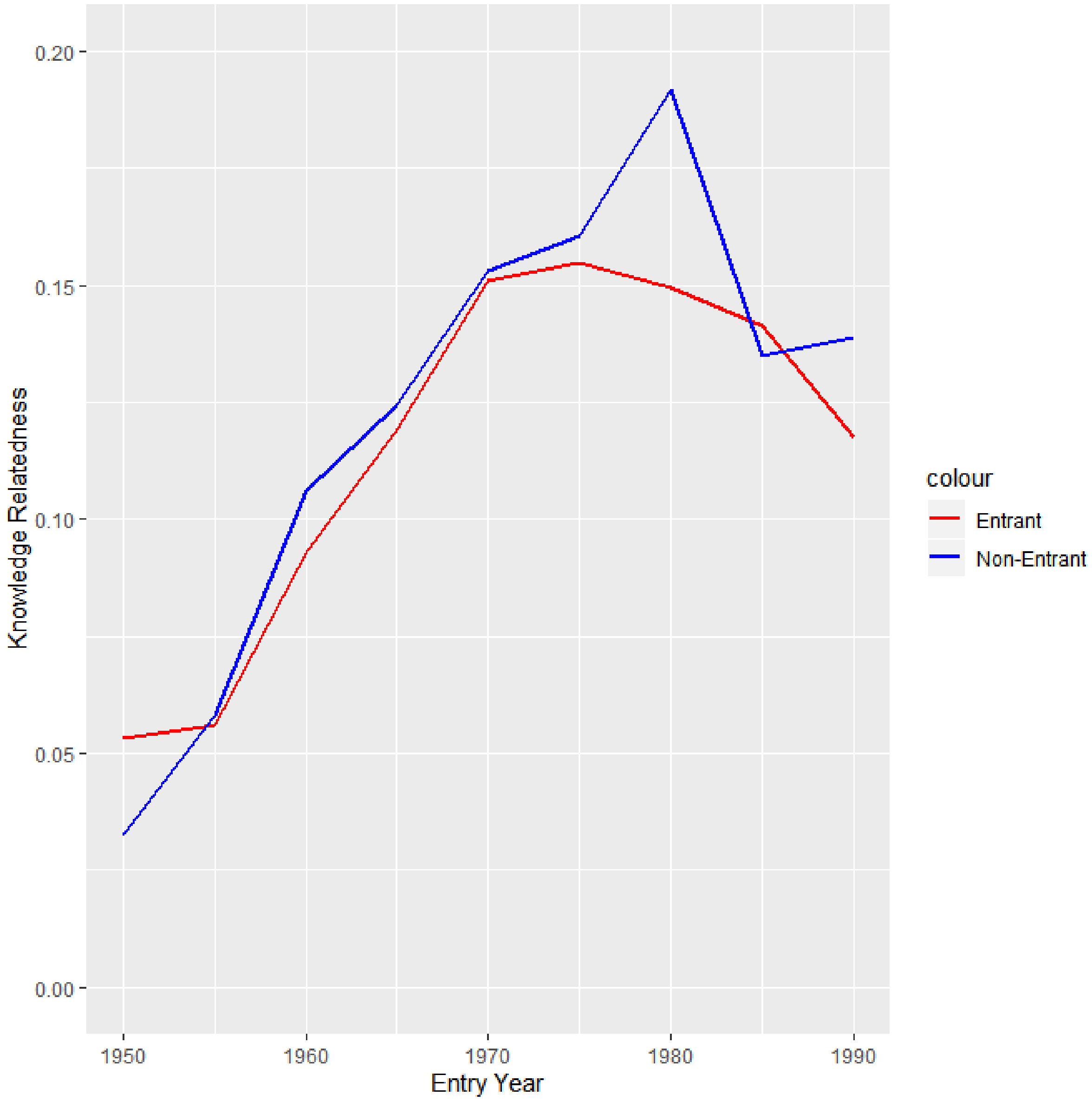
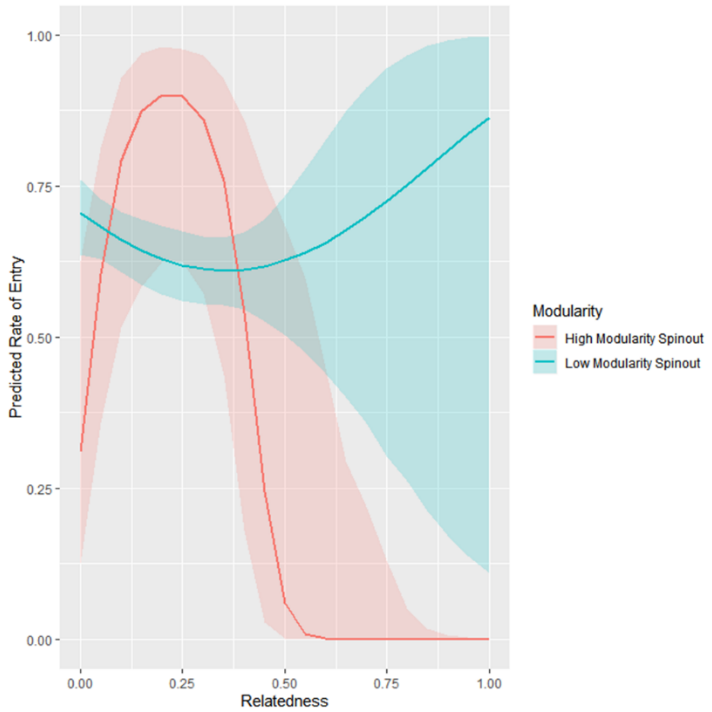


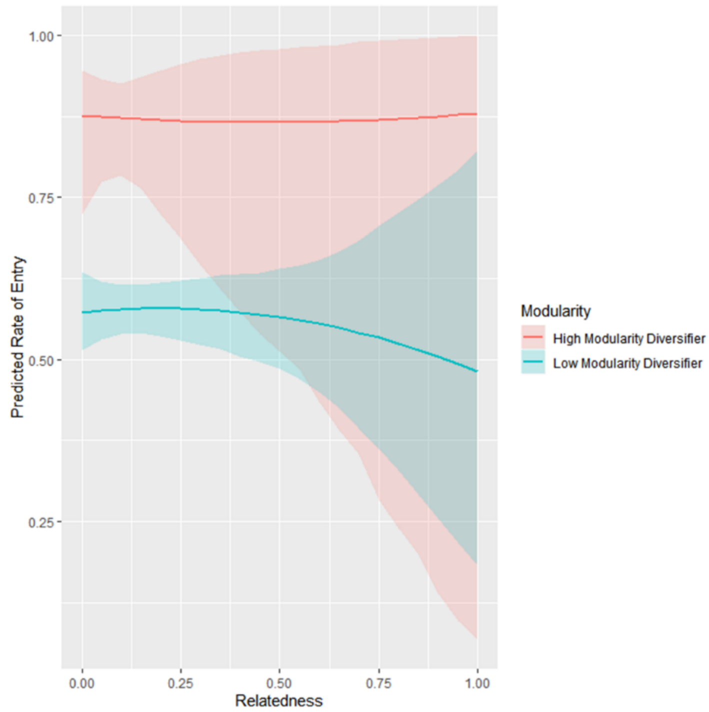
FIGURE 5

Visualization of results across modularity

Spinouts



Diversifiers



Spinoffs

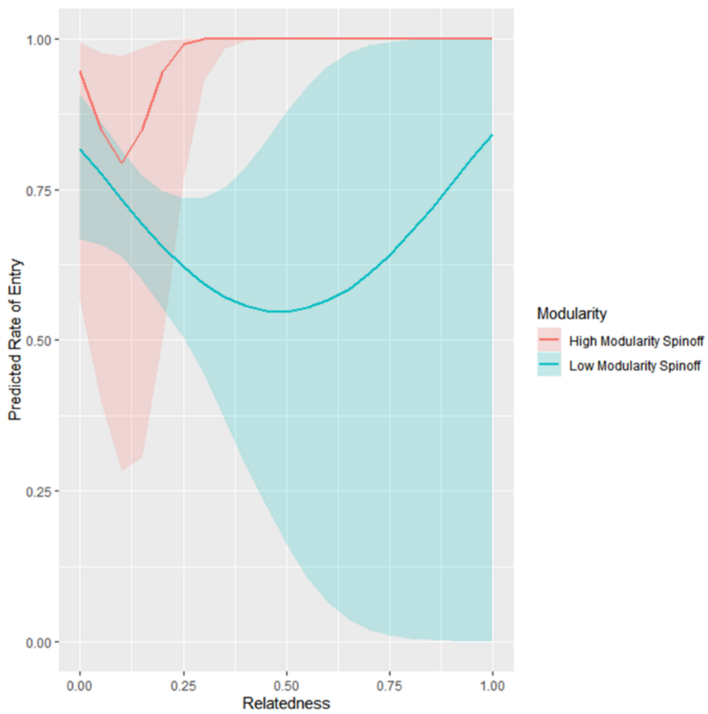
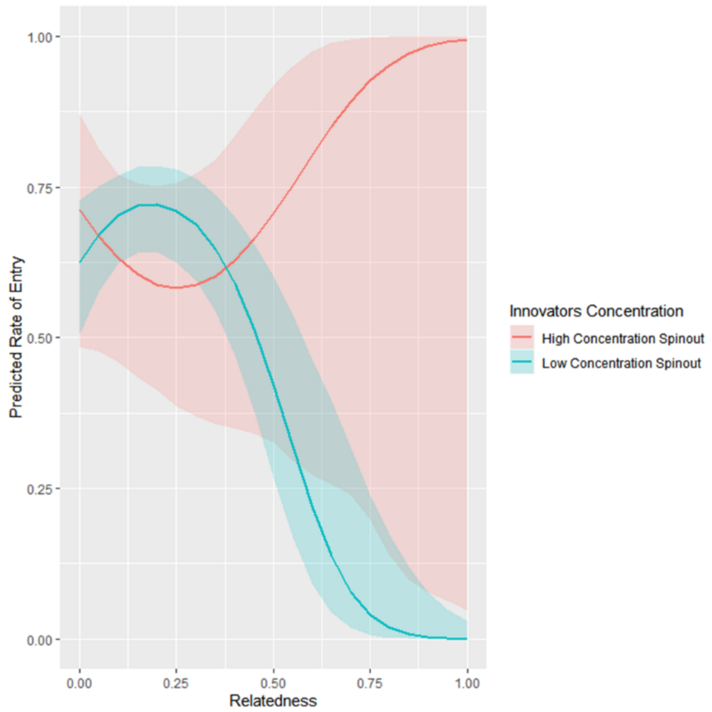


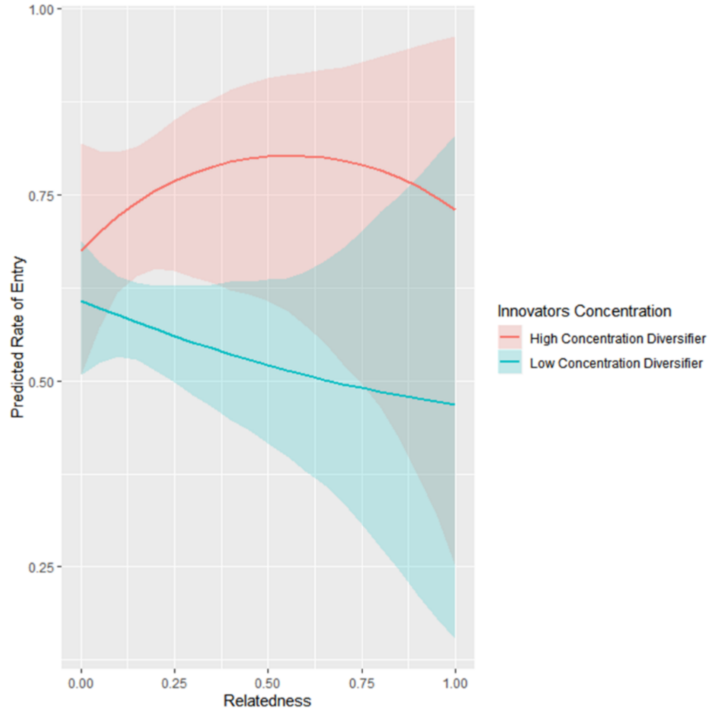
FIGURE 6

Visualization of results across innovators concentration

Spinouts



Diversifiers



Spinoffs