GOTHENBURG STUDIES IN INNOVATION AND ENTREPRENEURSHIP 2

Exploring Knowledge Intensity in Entrepreneurship

A quantitative study of knowledge, innovation and performance in entrepreneurial firms

Ethan Gifford

GÖTEBORGS UNIVERSITET
HANDELSHÖGSKOLAN

Institute for Innovation and Entrepreneurship
Department for Economy and Society
University of Gothenburg
Gothenburg, Sweden

ISBN: 978-91-981507-3-5

Available online: http://hdl.handel.net/2077/51773

Printed by Campusservice Lorensberg

Gothenburg, Sweden. 2017

Abstract

This Ph.D. dissertation investigates the statistical and theoretical relationships between different dimensions of knowledge intensive entrepreneurship (KIE) in Europe and how knowledge intensity and performance in entrepreneurial firms can be related. KIE is modeled as an application of resource-based theory, connecting pre-entry inputs like education and experience to external search activities to innovativeness and firm financial performance, growth, and survival. The data used was collected during a wide-scale EU financed framework project (FP7 - AEGIS), combined with additional panel-based firm level data gathered by the author, in order to investigate knowledge intensity, innovation, and performance in entrepreneurial firms: Moreover, the thesis explores how these concepts might be defined and modeled. Confirmed results indicate: Positive associations between depth of external search with innovative performance and a partial inverse curvilinear association between breadth of external search and innovative performance; Positive yet inversely curvilinear associations between the beneficial aspects of functional heterogeneity (or, knowledge scope) of the founding team with that of financial performance and survival, and negative linear associations between detrimental aspects of functional heterogeneity (or, knowledge disparity) of the founding team with the same response variables; Finally, positive associations were identified between the radicalness of innovations produced both with that of financial performance over time, and with the likelihood of firm survival. Conclusions use these results to reflect on broader relationships between knowledge intensity, innovation and performance in entrepreneurial firms. Recommendations for future research include more advanced modeling of complex latent factors constituting different forms of internal and external knowledge intensity, innovativeness, and performance on the part of entrepreneurial firms. Furthermore, drawing more extensively on existing tools such as resource-based theory may prove more enlightening than constructing new concepts and typologies to explain knowledge intensive entrepreneurship in new light, and policy wishing to promote knowledge intensive entrepreneurship may find it beneficial to focus on the educational and experiential underpinnings of creating such firms in diverse industries including low- and medium-technology industries as well as different types of services.

Key Words: Innovation, Entrepreneurship, Firm Performance, Knowledge Intensity, Search, Human Capital

Sammanfattning på svenska

I EU- och OECD-länder har stora delar av policyskapande fokuserat på hur ett land eller region kan öka både sin produktivitet och stimulera startandet av nya entreprenöriella företag som inte bara skapar jobbmöjligheter utan även bygger på ny teknik, nytt design-tänkande, och andra typer av innovationer. Denna doktorsavhandling handlar om entreprenörskap och innovation i EU-området, och hur små- och mikroföretag i olika industrier använder sig av olika former av kunskapsintensitet för att öka sin prestationsförmåga och konkurrenskraft. Kvantitativa metoder används för att bland annat kartlägga och mäta samband mellan den interna och externa kunskapsintensiteten i entreprenöriella företag och hur dessa förhåller sig till företagens innovationsförmåga. Samtidigt mäts sambandet mellan innovationsförmåga och affärsmässig prestationsförmåga. Författaren jobbar med enkätbaserad data som samlades in under EU:s 7:e ramprogram, vilket syftade till att utvidga förståelsen för kunskapsintensitet på företagsnivå och hur den bidrar till ekonomisk tillväxt och välfärd i stort.

Vad det gäller företagets agerande i relation till öppen innovation, så finns det positiva samband mellan hur brett och hur djupt de söker efter ny extern kunskap - det vill säga hur många olika typer av källor de använder sig av och till vilken grad de används - och deras innovationsförmåga. Det finns även starka samband mellan ett entreprenöriellt företags interna kunskapsintensitet - dvs. utbildningsnivå hos grundarna och de anställda-, grundarnas tidigare arbetslivserfarenhet samt förmåga att utnyttja nya möjligheter som uppstår ur ny teknik och institutionella förändringar, och företagets förmåga att använda sig av vissa typer av extern kunskapsintensitet, så som hur högt de värderar kunskap från universitets- och forskningsvärlden kontra kunskap från aktorer i företagens värdekedjor. Det finns också starka samband mellan intern kunskapsintensitet och företagets affärsmässiga prestation, i form av antalet anställda och rörelseintäkter, och i vissa fall dess överlevnad. Innovationsförmåga på företagsnivån har också ett positivt samband med sådan affärsmässig prestation.

Det nystartade företagets interna sammansättning är viktigt i fler avseenden än vem man är och vilka erfarenheter man har som entreprenör, vem man väljer att anställa och ta med sig på sin resa spelar också en stor roll. Grundarna är givetvis viktiga, men mer i termer av att de kompletterar varandra väl när det kommer till deras arbetsmässiga bakgrund och erfarenheter än exempelvis deras utbildningsnivå, vilket istället spelar roll när det kommer till de anställda. Dessutom verkar det mer fördelaktigt i relation till innovation för nystartade företag i de industrier som studerats att satsa på att söka djupt snarare än brett när det kommer till extern kunskap. Ju mer innovativt ett företag är ju bättre lyckas det i längden jämfört med mindre innovativa företag i dessa sektorer, men det finns risker kopplade till innovation vilket gör läget osäkert. Dessa resultat förmedlar insikter till innovationspolicy gällande kunskapsintensitet i olika sektorer.

Acknowledgments

This body of work could not have been completed without the following people's contributions, advice, or support:

My supervisors Maureen McKelvey and Olof Zaring. Your help and guidance during this process has been invaluable to me, and the final results speak I think for themselves. Thank you for allowing me to pursue this area of study, and for providing me with level-headed, practical, and theoretically sophisticated advice and feedback throughout the process. I could not have imagined a better complement to my own limitations as a researcher during this process than your careful and competent approach to supervision.

My senior colleagues at the Institute for Innovation and Entrepreneurship in Gothenburg: Daniel Ljungberg, for always making yourself available for whatever issue might be plaguing me, be it theory, statistics, or anything else. Thanks for throwing a paper at me and saying "why don't you start with this?", which ended up shaping in large parts the base of this entire piece of work! Evangelos Bourelos, thanks for the invaluable advice during the work on statistical models, and for your aid during the finalization of this thesis with formalities and bureaucracy of actually obtaining a Ph.D. Having people with your caliber of knowledge and experience with empirical research has been a true privilege. Johan Brink, thanks for the discussions about the AEGIS survey itself; data, techniques, and the theory behind everything. It has been instrumental in forming my final product (It's everything except cars and rubber!). Thanks to Rick Middel, and to Harald Dolles, for encouraging me to do a Ph.D. in the first place! You both helped to give me an idea of what a research career entails, and provided me with good role models early in the process.

To my current Ph.D student colleagues at the IIE; Snöfrid Börjesson Herou, Erik Gustafsson, Karin Berg, and Linus Brunnström: Our discussions and Ph.D. meetings have been a huge help to me in completing this work. Also, I would like to thank those guest scholars, post-doctoral scholars, and visiting professors (in no particular order) associated with the institute at one time or another during my Ph.D. process, who have been in some way instrumental: Marcus Holgersson, Ida Hermansson, Rani Dang, Elena Mas Tur, Magnus Holmén, Astrid Heidemann Lassen, Jun Jin, Guido Buenstorf, Lina Bakker, Ali Mohammadi, Rögnvaldur Sæmundsson, Sharmista Bagchi-Sen, Staffan Albinsson, Solmaz Sajadirad, Berna Beyhan, and any others who at the moment escape me.

I'd like to give special thanks to all those at the Department of Innovation and Organizational Economics, with whom I spent some time gathering data mid-way through the process: but special thanks Peter Lotz, Gitte Hornstrup, Anders Ørding Olsen, Toke Reichstein, Keld Laursen, Mirjam van Praag, Mark Lorenzen, Adrian Gutierrez, and Theodor Vladsel. Also, thank you to all those whom I interacted with both at the DRUID Academy conference in January 2015, as well as during DRUID15 in Rome. Special thanks to Lars Fredriksen, Bram Timmermans, Dominik Hepp, Katia Pina, Anders Ørding Olsen and Ron

Boschma for opposition, comments and inspiration. Thank you to all those at the Inter-university Consortium for Political and Social Research (ICPSR) who I interacted with during the Summer of 2015 in Ann Arbor, Michigan. I am especially grateful to Bill Jacoby, John Fox, Shawna Smith, Kelly Gleason, and Trent Mize for their inspiring instruction in statistics and for all correspondence thereafter. Thank you as well to the following statisticians; James Honaker, Sanford Weisberg, and Kenneth Bollen, for their correspondence and help.

I would also like to thank my all friends, who at this point are widely spread across the globe, for all the support and kind words I have received during this process. I would also like to thank my family, especially to: My wife, Kristin Gifford, for always being such an amazing listener, opponent, and discussant above and beyond the call, and for her endless encouragement and support in all aspects of life; thanks to Teresa Gifford, my mother, for her constant encouragement and thanks to Ted Gifford, my father, for the same support, and for advising and helping me to learn the R Statistical Computing Environment, which has become a tremendous asset to me in professional life, and for being a constant source of advice and inspiration in both statistical as well as theoretical aspects of both the R program and the Ph.D. process itself.

Ethan Gifford
March, 2017

Table of Contents

List of Figures

List of Tables

List of Key Variables

`AgeMax` - The age of the oldest founder in the founding team.

`Breadth` - Construct measuring total number of sources used for exploring knowledge external to the firm.

`Depth` - Construct measuring total number of sources the firm deems deeply important for exploring external knowledge.

`EmpEdu` - Proportion of employees (including founders) with at least a tertiary degree

`EmpHiEdu` - Proportion of employees (including founders) with at least a Masters or PhD degree.

`EXPC1` - Importance of knowledge stemming from specialized knowledge providers including state, national, or regional research-based or academic entities.

`EXPC2` - Importance of knowledge stemming from clients, customers, and suppliers, or intra-industry knowledge.

`EXPC3` - Importance of knowledge stemming from academic publications and trade conferences.

`FF1` - The importance of opportunity based factors for founding the firm

`FF2` - The importance of experiential and network based factors for founding the firm

`FF3` - The importance of design and technical knowledge for founding the firm

`FoundEdu` - Highest level of education attained by the founder (1-5)

`FoundEnt` - Binary variable denoting whether the founding team possesses any prior entrepreneurial experience.

`FoundInd` - Variable denoting the number of years of industry experience of the founder(s).

`FoundUni` - Binary variable denoting whether the founding team possesses any prior university experience.

`InnoGoods` - Proportion of new or significantly improved goods to total sales.

`InnoServ` - Proportion of new or significantly improved services to total sales.

`KDisp` - Knowledge Disparity: A combined measure of the dissimilarity and non-redundancy of the functional backgrounds of the founding team.

`KScope` - Knowledge Scope: A combined measure of the variety and diversity of functional backgrounds of the founding team.

`logEmp` - Natural logarithm of the number of employees a firm had in a given year.

`logOpRev` - Natural logarithm of the amount of operating revenue a firm has earned in a given year.

`RadInnOS` - Rescaled ordinal variable denoting the highest degree of radical innovation of a firm.

`Spinoff` - Binary variable denoting whether the firm came from a prior organization or not.

Chapter 1

Introduction - New ventures and how they impact societal growth and development

For many years now, both scholars and policy makers around the world have argued that newly established business ventures should be seen as important engines for growth in modern economies. This is not a statement that provokes much argument in the present day, but this was not always the case. One need only compare our current understanding of economic growth with that which was prominent in the early to mid- 20th century; a time when policy makers in capitalist economies were largely content with putting their trust in neoclassical growth theory, something which largely failed to account for the entrepreneur and entrepreneurial activities such as the starting of a new business venture in its tenets. Though much of the research community in modern economics still ascribes to largely non-dynamic models which exclude the entrepreneur, the view of entrepreneurship as a potential driver of economic growth has changed a great deal. This change in perception is in part thanks to the work done in the fields of evolutionary economics, entrepreneurship studies, and economic geography. Moreover, it is important to remember that how much an economy benefits from new ventures also depends on how these ventures contribute to, and what they provide to society. That is to say, entrepreneurship in and of itself may not be the answer to stimulating economic growth; however, some particular types of entrepreneurship may be better poised to do so.

Both the static (today) and dynamic (over time) well-being of society are not so easily measured or aggregated. It has however been argued that there are certain new ventures which, through innovative use and application of new knowledge, are able to definitively impact society and strongly contribute to economic growth more so than others. And, despite many claims to the contrary, this contribution is not isolated, or even concentrated, to specific industries or sectors. Actually, these growth-enhancing firms may span multiple sectors and industries, and while the industrial context remains a deciding factor in the nature of their contribution, this contribution occurs in part through the application and/or usage of novel types of knowledge in core and peripheral industries to the firm, resulting in new opportunities, methods, devices, processes or products, which may progressively impact economic growth and well being. This type of firm may be referred to as a *knowledge intensive entrepreneurial firm* (Malerba and McKelvey, 2015).

Over the years, scholars have explored what it is about new ventures that actually stimulates the economy, and this often leads to a discussion about the nature of entrepreneurship. Entrepreneurship is *praxiological*, that is, in many aspects, the art of humans acting on opportunities (Mises, 1949). Entrepreneurs often start companies based on identified markets in order to make their living and to provide for their families; in order to exploit their talents for profit; or in some cases, in order to carry out some mission or vision that has a direct benefit to society. Alternatively a founder to-be may leave an employer in frustration, after being unable to fulfill his or her personal mission or vision for certain intellectual property created within the confines of the employer firm, or to exploit an opportunity the parent firm has passed on, or failed to perceive in full. Often, a new venture will in one way or another arise as a manifestation of an opportunity to exploit or profit from a change in technology or productive knowledge; alternatively it may find and exploit holes in the market, where needs are not being filled sufficiently; a so-called 'market failure' effect.[1] Despite some variation in purpose and scope, research has shown that quite many firms are constantly using and applying new scientific, technological, organizational or design-based knowledge, and harnessing and using their resources and capabilities, in new ways; some having widespread implications for the economy as a whole, some merely affecting smaller niche markets. One great challenge lies in constructing a coherent and useful typology of these firms, and what drives them, from which more understanding might come regarding how and to what extent they actually drive the growth of economies

This Ph.D. thesis is, in a broad sense, chiefly concerned with this type of firm; which we may call the *knowledge-intensive entrepreneurial (KIE) firm.* It is posited to *actively use or apply novel forms of scientific, technological, organizational or design-based knowledge in its competitive and remunerative activities* (Malerba and McKelvey, 2015). The examples of this particular type of firm are numerous and quite varied, and work has been done, mainly in connection to research programmes funded by intergovernmental organizations, to build up an understanding of how this type of firm develops in relation to its surrounding innovation system (ibid., Malerba et al., 2015; Malerba, 2010). Varying classifications of

[1] Market failure that is exploited by entrepreneurs often deals with information asymmetries and resource allocation (Barbaroux, 2014) of supply and demand, often stemming from some variation of Knightian (1921) uncertainty about the future, or other forms of market disequilibrium.

KIE firms have been used in the recent literature (Malerba, 2010): Corporate spin-off firms radically redeveloping or re-envisioning their parents' technology for new aims; Academic spin-off firms making their first step into a market with new technology based on new scientific developments; A firm providing business and technologies services, like enterprise resource planning-based application and consultancy, improving and developing new systems and techniques to enhance their clients business and resource management; or a new firm involved in food production harnessing new technological breakthroughs in feed and feeding procedures, could all be seen as KIE firms.

In terms of what makes these firms such a driving force in societal development, it may be simply argued that it is how they harness this 'knowledge intensity', and how this activity drives their growth, survival, and performance, creating both implicit and explicit benefits for society. In this regard, KIE firms that perform well are likewise assumed to contribute relatively more to society and to growth. While much has been done on mapping the KIE firm and how it is embedded in its environment, more in-depth work is needed concerning analyzing and exploring the empirical associations between knowledge, innovation, and performance in these firms.

1.1 The role of knowledge and the new firm in economic growth: Perspectives and directions

The link between knowledge and economic growth is one that is now well established. Much of the literature in the fields of innovation and entrepreneurship has revolved around the impact of individual- and firm-level knowledge on economic growth and technical change in society (cf. Solow, 1957; Nelson and Winter, 1982; Romer, 1990; Metcalfe, 2002; Bloch and Metcalfe, 2011). Indeed, Carl Menger, widely acknowledged as the father of the Austrian school of economics, expressed already in 1871's (1976, p. 74) *Principles of Economics* that the "degree of economic progress of mankind will still, in future epochs, be commensurate with the degree of progress of human knowledge."[2] Not long after, Schumpeter (1934; 1939; 1942) and other growth theorists (cf. Young, 1928; Burns, 1934; Kuznets, 1954) established a connection between what knowledge

[2]Menger postulated that Adam Smith had in effect only scratched the surface with his division of labor hypothesis of how firms can grow and improve productivity and drive the economic machine forward. He pointed out that it is not merely the division of labor that is of note but the immersion into a specific activity by a human being that allows them to develop specific knowledge and capabilities, which over time will improve efficiency and sow the ground for future innovation (Menger, 1871/1976).

resides in innovative firms, the actions of entrepreneurs, and the growth of the modern capitalist economy.

Schumpeterian innovation scholars see entrepreneurs and entrepreneurial ventures as a crucial, dynamic, driving force of economic activity. Despite extensive research over the past decades, how innovative, entrepreneurial firms contribute to economic growth remains a phenomenon that is not completely understood (Block et al., 2016), and a significant area of research and policy agendas still thrives around these and other related issues. In light of this, it is not surprising that in recent years, young, dynamic, high-growth firms have been identified by inter- and non-governmental organizations (and many scholars) as particularly beneficial for economies, accounting for a significant amount of job creation and employment opportunities, not least in developed countries (OECD, 2003; 2013; 2015)[3]. While many have taken this observed phenomenon and tried to classify and understand the types of firms most responsible for growth generation, select research groups have attempted to pin down this typology of a new firm driving economic growth and technical change as being *knowledge intensive* as well as *entrepreneurial* (cf. Delmar and Wennberg, 2010; Malerba and McKelvey, 2010; McKelvey and Lassen, 2013; Malerba et al., 2015).

This research has helped many realize that knowledge intensity can be a multifaceted construct, and that it is not something that only exists in the traditional high-technology industries (Smith, 2002; Hirsch-Kreinsen et al., 2008; Malerba and McKelvey, 2015). This notion has traditionally been a common conception within the literature surrounding innovation and entrepreneurship, and as represented in many studies of innovation and change in pharmaceutical, engineering, and other high-tech fields. Potentially manifesting in diverse sectors and activities, knowledge intensive entrepreneurial firms are said to be distinguished by their application of new knowledge or innovation (McKelvey and Lassen, 2013). The way in which one defines this type of firm is crucial for its identification and use in theory, practice, and policy making. *The knowledge intensive entrepreneurial venture has been defined by some as a new firm that strategically uses new scientific, technological, or organizational knowledge to reap economic rewards and harness innovative opportunities* (Holmén et al., 2007; Malerba and McKelvey, 2015).

[3]OECD 2013 Science technology scoreboard, p, 13: "Young, dynamic firms contribute more to job creation than previously recognized ... Young firms with fewer than 50 employees represent only around 11% of employment, but they generally account for more than 33% of total job creation in the business sector; their share of job destruction is around 17%."

There has been a recent upswing of the importance of knowledge intensive entrepreneurship, ‘high-potential entrepreneurship’ (Delmar and Wennberg, 2010; Autio and Acs, 2007), or innovative or Schumpeterian entrepreneurship (Block et al., 2016). However, despite much theoretical and exploratory work being done to address these related concepts (Malerba, 2010; Malerba et al., 2015), there are still relatively few empirical in-depth studies in regards to the relationship between knowledge, innovation and performance of the firm. Somewhat difficult to capture in an aggregated perspective, a knowledge intensive (entrepreneurial) firm can take on many different forms[4] in definition and conceptualization.

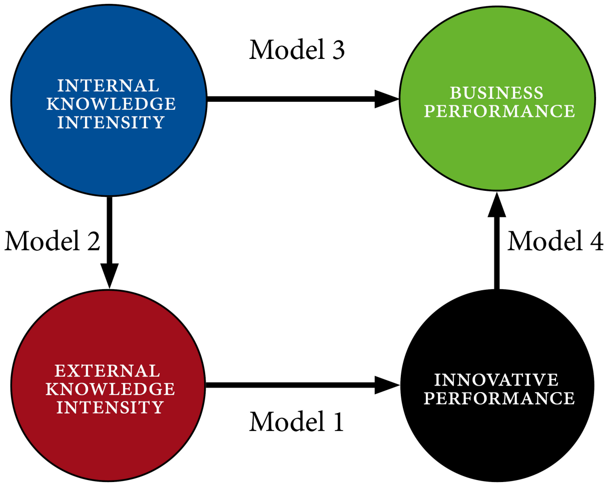
1.2 Research to be carried out

This thesis will attempt to clarify the interaction between the different properties possessed by such a firm, and how these properties interact with and influence one another. Simply put, it deals with *to what extent an entrepreneurial firm’s knowledge intensity affects how the venture performs, and how different types of knowledge intensity and performance affect one another*, according to different metrics. What is meant by this is that there is not only more work needed in exploring relationships between knowledge intensity and how it affects, or associates with, performance, but that there is also a need to look deeper into the theory and the conceptualization of knowledge intensity, as well as performance itself (in which I include innovation performance) in order to explore the inner workings of the KIE phenomenon, and see what relationships exist both between and within.

To achieve this, I will draw on an array of quantitative empirical methods, and later draw conclusions and implications for practice, policy, and future research. The overarching conceptual framework used draws on the resource-based view of the firm, entrepreneurship theory, innovation systems approaches, and economic theory in linking inputs of

[4]Some argue that it does not make sense to talk about aggregated individual knowledge on an organizational level, since knowledge per definition implies some kind of verified truth residing in the human mind, and that new knowledge must be a product of mental processes therein (Hayek, 1945; Loasby, 2000). However, it can be argued that firms draw on the subjective, individual knowledge of its members in order to build a sort of collective or shared understanding, which takes the form of information, and functions as an approximate aggregate, drawing on bundled routines (Nelson and Winter, 1982; Kogut and Zander, 1992) that, in tandem, embody and at the same time develop its purpose and functionality (Metcalfe and Ramlogan, 2005). Grant (2002) has rightly pointed to the fact that once an organization is viewed as a ‘knowing entity’, difficulties arise in terms of how one can differentiate the mechanisms through which individual knowledge is combined or linked together, and the ‘knowledge base’ of the firm which results of this collected knowledge. This is a distinction that is not touched upon here, but see Tell (2004) for a more in-depth take on organizational knowledge

Figure 1.1: Summary of the models



resources and capabilities, especially those related to knowledge intensity, to outputs relating to performance and creating a sustained competitive advantage. Furthermore, this Ph.D. dissertation carries out an in-depth analysis of knowledge intensity as a concept and as a construct.

The main theoretical point of departure is as follows: There is both an internal and an external component to a firm's knowledge intensity. Thus, two working definitions or conceptualizations may be introduced: Internal knowledge intensity: or, *the knowledge intensity that is largely inherent in a firm when it comes into being, rooted in different types of human capital investments and outcomes, as well as other knowledge-based factors have driven the firm to formation*; and external knowledge intensity: or, *the way and extent to which a firm searches out, relies on, and valuates external knowledge post-formation.* I will go on to explore how these concepts are inter-related empirically, as well as how they affect different outcomes of performance in entrepreneurial firms., which are also inter-related. At the end of the thesis, these two knowledge intensity dimensions will be re-assessed according to the empirical results and discussions, and I will link them to both innovative and economic/business performance.

Figure 1.1 below summarizes the full conceptual relational map of the thesis. Each model (and thereby research objective) is represented in the figure by numbers 1-4, and shown here is the scheme of associations used in the analyses, that is, how the different input and output concepts affect one another.

The research objectives to be addressed are as follows:

- **Research Objective 1**: *Explore the association between external knowledge intensity and innovative performance in the entrepreneurial firm*

Innovative performance, and how it relates to external search practices for knowledge, has seen much attention in the field of innovation studies. Mainly this has been in the context of large manufacturing firms. Using the unique dataset of European entrepreneurial ventures across numerous sectors, I test these relationships in different forms and compare them with what we know already about external knowledge usage by firms, but integrated within the knowledge intensity framework.

- **Research Objective 2**: *Explore the association between internal knowledge intensity and external knowledge intensity in the entrepreneurial firm.*

Some research exists mapping how the firm's pre-founding history affect performance, but not as many exist regarding how the main factors of this pre-history, which I attribute to internal knowledge intensity, deal with the association to external knowledge intensity as I characterize it.

- **Research Objective 3**: *Explore the association between internal knowledge intensity and business performance of the entrepreneurial firm.*

Much research exists regarding inputs of entrepreneurial firms as affecting outputs like performance, but few which systematically analyze across multiple sectors and regions containing knowledge intensive activities on different scales.

- **Research Objective 4**: *Explore the association between innovative performance and business performance of entrepreneurial firms.*

One of the key propositions about knowledge intensive entrepreneurship is that innovativeness of firms should positively associate with economic growth and performance. In lieu of being able to compare the effects on whole economies, I carry out an investigation looking at if innovativeness of entrepreneurial firms actually does positively associate with firm-level growth, volume and performance.

Using these different objectives as a tool to structure the different intra- and inter-relationships between on the one hand, knowledge intensity and on the other, performance (including both innovation and business performance), I hope to find some confirmation that these relationships are both interlinked and also influential. This should, in my view, aid in tackling such broad and complex realities as innovativeness, entrepreneurial firms, and the elusive relationship to economic growth and well being by use of finely honed research tools.

To begin this line of inquiry, it will first be necessary to assess the scientific literature on the knowledge intensive entrepreneurial firm, and to find the most suitable method of analyzing it. In the aim of further illuminating the KIE firm, and to learn more about the way in which it interacts with, learns from, and evolves with its external environment, this dissertation will go deeper towards understanding the links between knowledge intensity and performance in the entrepreneurial venture, and thereby contributing to the ongoing discourse between scholarly and political groups attempting to disentangle how to move forward with this type of conceptualization.This will be the subject of the following two chapters.

The objective of Chapter 2 will be to *identify the conceptual and literary origins of the knowledge intensive entrepreneurial firm.* In order to understand and analyze it, one needs to understand the conceptual roots of the KIE firm and concept in a broad sense, including how it relates to the world industrial economy, what role policy plays and has played, and how the research of KIE firms has developed over time. This type of background material will be compiled, and presented, in Chapter 2. Once it becomes clearer how the KIE firm came into focus in research and practice, it is necessary to delve into relevant literature in the realms of strategic management, organizational science, business administration, and other relevant fields of study, in order to derive an appropriate theoretical framework to assess the research problem. Additionally, more specific research objectives need to be derived and placed into context to give the project adequate depth and scope. So, the objective of Chapter 3 is to *build a conceptual framework through which the knowledge intensive character of entrepreneurial firms can be analyzed and understood in terms of the detailed relationship between, and involving, types of knowledge intensity and performance.* This chapter shows how resource-based theories of the firm can be applied to Knowledge Intensive Entrepreneurship, and how concepts based in human capital, organization origins, and external search activities may be used to represent the different latent aspects of Knowledge Intensity. Following these chapters this report will go further and operationalize the concept of knowledge intensity in a new venture, as well as appropriate measures of firm performance, into variables that can be empirically measured and analyzed across various industries. This is in order to align the definitions employed with the constructs developed and used in subsequent empirical chapters. This analysis will take place in Chapters 4 and 5, each containing two related models. Put quite succinctly, in Chapter 3, using organizaitional and entrepreneurship theory, I separate the idea of knowledge intensity into an external and an internal dimension, the former being associated with sources of knowledge external to the firm and how the firm gauges their importance for its business, and the latter being based on various human capital inputs and outputs like education, experience, and other conditioning factors leading to the formation of the firm itself. I then try to analyze how both of these concepts (and their constructs) affect firm performance.

In terms of data used in the subsequent analyses, Chapter 4 will cover the first two research objectives using two model sets, which are based on quantitative data from the AEGIS survey conducted during the EU-funded AEGIS project investigating knowledge intensive entrepreneurship in Europe, with data gathered in 2010/2011[5]. Capturing data from over 4000 firms in 10 different European Union member states, the AEGIS survey is predominantly a blend of Likert scale ordinal/interval variables and numeric (ratio) variables. This chapter looks at two components; the effect the intensity of use of knowledge external to the firm on innovative performance, and the effect of firm pre-history on the intensity of use of this external knowledge. I use different regression and modeling techniques depending on the nature of the variables of interest and what fits the data best.

In Chapter 5 I will address the third and fourth research objectives. The resulting empirical models will expand on the AEGIS survey material and complement it with another source of data, namely financial data about the same firms. The chapter outlines how firm pre-history, by way of human capital constructs at the firm level, affect performance. The results are obtained by looking at the same firms' financial data from 2010 to 2015, as opposed to only 2010/2011. Finally, Chapter 6 will discuss and analyze the results in relation to the theory derived in Chapters 2 and 3, and Chapter 7 will conclude, reflecting on the future of the KIE concept for researchers, practitioners, and policy makers, along with the major implications of the text. This includes a re-orientation of what constitutes knowledge intensity and thus knowledge intensive entrepreneurship from my own perspective.

[5]This constitutes in and of itself a kind of empirical contribution to the field of innovation and entrepreneurial studies, as many have pointed out the underrepresentation of young firms, and especially firms of this particular character, in large scale databases, and survey data analyses (Head and Kirchoff, 2009; Coad et al., 2013; Coad et al., 2016).

The reasons for conducting this type of analysis is multifaceted. Practitioners in these types of entrepreneurial firms need to expand their own understanding of how knowledge intensity and knowledge-intensive processes take form or may be accounted for; human action that may be on many levels unaccounted or unconscious, in order to better harness their own resources and capabilities to the best effect, to improve their innovative potential, or to better understand what may be driving their firm's competitive advantage. Also, factors contributing to survival, growth, and different types of performance should be of interest to entrepreneurs in diverse sectors covered in this study. Policy makers need to refine and develop tools for encouraging and effectivizing the research on, and support of, knowledge intensive entrepreneurial firms, especially concerning their potential categorization and how they are targeted for support. It has been established that simplistic and unidirectional funding of sectors of economic activity where some baseline indicator of knowledge intensity is in use has had mixed success in stimulating growth. A more nuanced picture of which new firms require support and why, in which sectors they might be acting in, and which technological and scientific resources they use and apply, is sorely needed.

Today's sovereign state governmental bodies, inter-governmental organizations like the UN and its contemporaries, and relevant non-governmental organizations need to become better at organically promoting and encouraging firms and industries having 'high potential', and also those that commonly escape this qualification yet are vital for growth. They also need to improve in thinking more systemically about the supply and demand functions that drive knowledge intensive entrepreneurship in different contexts. This can hopefully be aided by establishing methods for identifying and explaining activities of entrepreneurial firms in sectors with high likelihood of knowledge intensive entrepreneurship occurring. Lastly, for scholars as well as policy actors, there is a great need for working towards clearer definitions of what constitutes knowledge intensity in a firm, how theory can be built around this concept, how it might be measured or operationalized as a construct[6], as well as how knowledge intensity shapes actual firm performance. Achieving these objectives will help in enhancing our thinking on what knowledge intensive entrepreneurship might be, and how this manifestation of knowledge intensity can contribute to a theory of economic growth driven by certain types of entrepreneurship.

[6] Avoiding tautology in definitions of knowledge is a real problem in social science research; see Grant's example (1996): 'that which is known'.

Chapter 2

Knowledge Intensive Entrepreneurship and the Firm: A Conceptual Background and Account of Previous Research

2.1 Knowledge, society, and the firm: Research and policy

This section reviews the role of knowledge at the economy level as proposed by management and economics research. It also explains the growth in the usage of the so-called knowledge economy as a term or concept involved in steering national and/or economic development by policy makers: It investigates the relationships between knowledge and economic growth in macro-economies as a whole, and how policy has been used to address this.

As touched upon in Chapter 1, the creation and application of knowledge has recently taken root in discussions about economic growth on a macro-level, and the world economic climate is widely argued to be becoming more and more of a knowledge economy,a knowledge society (Drucker, 1994; Stehr, 1994; Wilke, 1998; Granstrand, 2000; David and Foray, 2003; Hirsch-Kreinsen et al., 2008) or a learning economy (Lundvall and Borrás, 1997). And, based in part on the increasing degree of globalization in business, knowledge-based economic activity is argued to be gaining a distinct comparative advantage in the international political arena (Audretsch and Thurik, 2001). This is exemplified through an increased focus by firms, states, and their policy makers on the importance of knowledge as a productive factor or asset (Hirsch-Kreinsen et al., 2008). One of the earliest to document and analyze this was Bell (1973), who wrote about this increased focus in the early 1970s in The Coming of Post-Industrial Society. He argued therein that this societal direction being pursued by many of the world's economies during this time was due to an increased focus on knowledge in both the formal and abstract sense. The development has continued since Bell's early observations, and in past few decades, advances in science and technology, combined with vast improvements in instrumentation have further increased the importance of applied knowledge in new and varied contexts across the international economy (Dunning and Lundan, 2008). Many have argued throughout the 20th and 21st centuries that, with the increasing role of knowledge and knowledge spillovers as a critical source of economic growth (cf. Kuznets, 1954; Romer, 1986; Acs et al., 2009), entrepreneurship has taken on a new status within this growing knowledge economy because “it serves as a key mechanism by which knowledge created in one organization becomes commercialized in a new enterprise” (Audretsch and Keilbach, 2006: 10).

Though they play a role, the changes in business environments across the globe, as detailed above, are not the sole or even primary cause of the strengthened interest by the scientific community in firm-specific knowledge (Grant, 2002). Even more important has been the relatively recent rediscovery of knowledge as a productive resource in the first place. This has been exemplified by a resurgence of knowledge-related intellectual debate and activity: More specifically, a revitalized contemporary scientific interest in works by the likes of Polanyi (1962), Arrow (1962), March and Simon (1958) and Hayek (1945) has resulted in a tremendous upswing of firms and researchers thinking about knowledge and its characteristics.

The knowledge economy ‘literature’ gained a lot of influence in the early 21st century. For many, a ‘knowledge’ or ‘learning’ based economy became associated with an ideal state, and in some sense became less and less an object of rigorous analysis. In many policy circles, the term knowledge intensity has become equated with a proposed target or goal for nations to achieve (cf. OECD, 2003; 2008; 2013; European Commission, 2013). This type of formulation is still quite popular in national and regional policy-making. For example, Europe is currently extremely focused on tapping into its knowledge resources to achieve growth. The European Commission recently restated the importance of making the European Union “a more knowledge-based, competitive economy . . . investing in knowledge and making structural changes towards more knowledge intensive activities” in order to reach the Europe 2020 goal of being “a smart, sustainable, and inclusive economy” (European Commission, 2013, p. 1, 3-4). It has also emphasized the need to “increase [its] capacity to channel knowledge, creativity and technology into innovative, internationally competitive products and services that respond to societal needs” (ibid., p. 8).

Much of the recent interest in the knowledge economy on a policy level coincided and the actual active implementation of governmental organizations which (attempt to) guide the innovative trajectory of nations. Many countries have begun to, at least in principle, apply an innovation systems approach by creating their own innovation and development offices. Sweden and Finland, who established their own national innovation office in 2001 and 1983 respectively, were two of the frontrunners in this regard. Some see this as an attempt to more effectively harness the interaction of actors within innovation systems to channel the "creation and use of knowledge for economic purposes" (Sharif, 2006, p. 745). Other nations and intergovernmental organizations are similarly becoming pre-occupied with competitiveness through harnessing knowledge-based economic activity. For instance, the EU as a governing body has place great emphasis on research and development funding through a series of quadrennial framework programs, starting with FP1 in 1984 all the way up to the present day Horizon 2020 (FP8) program.

This new focus on the knowledge economy has not been without criticism. While many involved in policy and research communities have viewed the knowledge economy, and knowledge based activities, as new phenomena, Keith Smith (2002) observes that in fact they are not. Indeed, he cites Andrew Ure, Charles Babbage, and Karl Marx as progenitors of the knowledge-based approach in the 19th century. Grant (2002) echoes the sentiment in his own phrasing; that recent approaches concerning knowledge-bases do not necessarily constitute paradigm shifting thought, but that they more represent a recognition of a way of thinking about firms, industry, and management that are and have remained valid regardless of era. However, Grant (2002: 135) propagates the knowledge based view of firm activity as something that, while not revolutionary, could certainly "[provide] a perspective that can augment and extend, possibly even transform, existing theory and management techniques." So, while there are lessons to be learned through viewing the economy as being knowledge-based, there exist many caveats which scholars worry that policy makers may fail to appreciate.

However one interprets the policy statements, or indeed the theoretical underpinnings of these statements, it is clear that the role of knowledge in the economy has been pushed to the forefront of the international economic agenda, and that the general focus remains on 'knowledge intensive' industries and sectors. (Oakey, 1991; OECD, 1996; Hoffman et al., 1998; Smith, 2002; OECD, 2008). However, what sectors this refers to, and what knowledge intensity has traditionally meant in different industrial contexts, needs to be explicated further. The next section addresses this by providing an account of how knowledge intensity has been classified in different industries and groupings of industries.

2.2 An economic view of knowledge intensity in different industries

This section reviews and compares how knowledge intensity manifests in different industrial contexts. It explains how knowledge intensity is often characterized in high tech, low/medium tech, and service sectors, especially knowledge intensive business services (KIBS).

Generally, knowledge intensity has traditionally been a difficult concept to define, though there are a few examples of some well-established measures that can be applied across sector and industries. Many studies use the term based on its most established definition popularized by the Organization for Economic Cooperation and Development (OECD, 2002, 2013; Hatzichronoglou, 1997); as a measure of R&D Intensity at the industrial or sectoral level. More recently, the idea of Knowledge Intensive Activities (KIA) has been used by Eurostat and EU level studies: This measure aggregates the level of tertiary education in a sector, such that those sectors in which employees who have obtained an equivalent of on International Standard Classification of Education (ISCED) level of 5 and 6 or above represent 33% or more of the workforce are deemed as KIAs (Eurostat, 2014 [Eurostat indicators of high-tech industry and knowledge-intensive services, annex]). Though aggregated, this measure focuses on the importance of manifested internal knowledge to the industry, and thus indirectly, the firm. These types of general indicators remain widely applied by IGOs and research units investigating national and industrial performance.

How small, entrepreneurial firms fit into these types of categorizations has become quite topical. Although the importance of new and established small firms has been known for several decades as being a vital component of economic growth and industrial change (Birch, 1979; Rothwell and Zegveld, 1982; Rothwell, 1989), this statement is nonetheless a broad one, which includes small firms that have been established for many years as well as entrepreneurial ventures. And, as I alluded to earlier on, recent streams of research focus upon the particular importance of a subset of these start-up firms, and the key conceptualization which underlies several of these streams of literature is that innovative entrepreneurial firms seem to contribute more to economic growth than other types of small firms.

On a more sectoral level, many scholars have worked specifically linking firms in certain sectors with different technological classes, leading to some divergence between what constitutes knowledge intensity depending on which technological class and sector one analyzes. The following sub-sections constitute an overview of firm classifications as well as proposed methods of measurement of knowledge intensity within different technological classes.

2.2.1 High-tech knowledge intensity

Originally, much of the research on knowledge intensity occurred in the context of high tech industries. Within this context, Von Tunzelmann and Acha (2005) attribute much of the importance of knowledge intensity to that of the development of new technologies. They argue that technology is explicitly central to commercial success. High technology firms and industries have traditionally attracted attention as being particularly relevant for national and economic growth (Oakey, 1991; Hoffman et al., 1998; Klepper, 2016), and booming industries such as information technologies, automobiles and auto-components, and pharmaceuticals were characterized early on as being highly knowledge intensive (Shane and Venkatraman, 2000; O'Regan and Sims, 2008). As these ideas began to gain prominence in the literature, scholars became more critical of what meaning the applied terminology actually carried, since accounts of what the term *high-technology* means have not always been uniform. Commonly, the categorization has its basis in how much expenditure a firm (or industry) spends on average in research and development activities (Leonard, 1971; Mansfield, 1972; Nelson and Winter, 1977; OECD, 1994; 2002; O'Regan and Sims, 2008), largely mirroring the macro-level measurement methods explained above. However, certain scholars have pointed out that this usage over-emphasizes the role of R&D in high tech firms' knowledge intensity and innovative activities (Mowery and Rosenberg, 1989; McKelvey and Lassen, 2013; Malerba et al., 2015)[1]. Nonetheless, in past studies, R&D Intensity has often been equated with knowledge intensity in high-tech industries and firms, at least for operational ends (OECD, 2002).

Additionally, human capital in the form of education of founding teams and employees has been richly used to assess the knowledge intensity or knowledge components of high-tech firms (Story and Tether, 1998; Colombo et al., 2004; Colombo and Grilli, 2005; Gimmon and Levie, 2010). This is because knowledge about or pertaining to the new developments in science and technology often resides in the education and experience of the founding team, and to some extent, of the employees of

[1]This R&D based measure also has received criticism especially regarding its application to samples containing many small and medium sized firms (Spender and Grant 1996; Grueber and Studt, 2011; Haahti et al., 2005), due in part to the difficulty in assessing R&D costs and indeed the lack of R&D departments in many SMEs (Autio et al. 2000).

high-tech firms. Many different identifiers for firm-level experience are based on founders' and employees' previous occupations, constituting different forms of spinoffs in the literature (more on this later).

2.2.1.1 Existing typologies of high-tech entrepreneurial firms

Gradually, research has become more interested in categorizing and defining different types of high tech, high growth firms. This has led to many taxonomies being devised in order to address certain peculiarities that are found in the origins and growth patterns of high-technology firms. Two classifications that stand out in the high-technology firm literature are 'gazelle' firms, and 'new technology based' firms. Birch (1989) refers to gazelles as a high growth, high-tech focused group of firms that are responsible for a large proportion of employment growth in national economies. Typically, gazelles are defined as being young (less than 5 years old commonly) and experiencing employee growth of upwards of 20% *per annum* (OECD, 2006), however, this is not universally accepted. While they have been established as ranging from good to great in terms of job creation, they have also been found to grow intensely because they are often so young (Henrekson and Johansson, 2010). New technology based firms (NTBF) have also been extensively studied without clearly defined industrial categorization and firm characteristics (Rickne and Jacobsson, 1999). Generally, a NTBF can be viewed as a relatively young firm whose competitive advantage is derived from new technologies, and are often founded as a vehicle for bringing this new technology to the market. There is a distinct overlap with KIE firms in that not only NTBFS, but sometimes also of gazelles, being categorized as firms that are strategically using new scientific and technological knowledge for competitive gains (Autio, 1997; Rickne and Jacobsson, 1999; Yli-Renko et al., 2001). Other typologies of firms such as Science-based entrepreneurial firms (SBEFs), or, firms that directly use academically derived scientific or technological novelty in their strategic orientation, have also arisen in recent years, which, while not necessarily gazelles or high growth firms, may be important for "the dynamic efficiency of the economic environment in which they are embedded" (Colombo et al., 2010, p. 2). Also, Technology-based New Firms (TBNFs) are proposed as benefiting from external and internal social capital as well as knowledge resources in order to achieve international growth (Autio et al., 2000; Yli-Renko et al., 2002). Another categorization relating to the KIE concept has been that of the young, highly innovative company, or YIC (EC-DG ENTR, 2009; Schneider and Veugelers, 2010). There is some overlap between this conceptualization and that of knowledge intensive entrepreneurship: Scheider and Veugelers (2010, p. 972) define it as a company that has "the potential to develop important innovations with significant potential commercial applications and social value"[2]. Additionally, the YIC is assumed to have high R&D Intensity (ibid; Czarnitzki and Delanote, 2013).

Additionally, Delmar and Wennberg (p. 28) have also fashioned another conceptualization of Knowledge Intensive Entrepreneurship, albeit with a slightly different framing. They view knowledge intensive entrepreneurship as a type of "high potential entrepreneurship" contributing to economic growth, and argue that *"entrepreneurship in new independent firms [as being] the "missing link" between generally new sources of knowledge and economically relevant knowledge"* in the Schumpeterian sense of creative new innovations destroying the rents of older innovations. By doing empirical analyses of Statistics Sweden data as well as individual level data, they attempt to answer a lot of the same types of big questions as EU projects like AEGIS and KEINS, that is, the effect of KIE on economic growth and well-being, etc. They define knowledge intensive entrepreneurship by employing the OECD definition of knowledge intensive sectors, but focus only on high tech and medium high-tech manufacturing along with knowledge intensive business services sectors. Delmar and Wennberg argue, in line with Acs and colleagues (2009) before them, that knowledge spillover theory can account for the majority of the benefit of innovation in knowledge intensive industries. Spillovers can occur by way of exchange of educated and specialized personnel (Eliasson, 1996), by technical production by the developing firm (despite protective measures), and knowledge stocks related to R&D, the creation of entirely new firms based on new knowledge, and employee interaction spurring new knowledge combinations. Delmar and Wennberg have followed Eurostat and OECD's classification of KI sectors, based on R&D expenditure to GDP or R&D Intensity (based largely on Götzfried (2004)).

These categories of firms have some degree of representation in the classification of KIE firms to be used in this thesis, though the firm and environmental characteristics that they focus on are neither necessary nor sufficient indicators of a KIE firm. KIE represents the potential for firms involved in other areas of industry than those directly focused on high technology output and high growth. Knowledge intensity, especially in the context of KIE processes, need not be limited to only the transformation of science and technology into commercial gains (Caloghirou et al., 2015).

[2]This is highly consistent with KIE as long as the firm is the innovator, however, in KIE firms, the application of the innovation is often what distinguishes the firm.

Some firms that are gazelles, NTBFs, or YICs maybe be KIE firms, and some KIE firms may fall into some of the latter categories,so in a sense there is some degree of synthesis between KIE and such categorizations (Malerba and McKelvey, 2015). However, it is important to remember that KIE however, may be found even outside high-tech industries. The next section will account for different past representations of knowledge intensity that can be found in low- and medium-tech industries.

2.2.2 Low- and medium-tech knowledge intensity

Many industrial and organizational research on knowledge intensity have traditionally been limited *only* to high-growth high-tech firms, or to knowledge intensive business service firms. Keith Smith (2002) has called for increased focus on low-tech industries as also being potentially knowledge intensive. He claimed that a real problem lay in that we do not really know that much about the ways in which embodied and disembodied knowledge flow between different societal agents, organizations, and institutions. He further argued that "knowledge creation is a sectorally distributed, economy wide process, not dependent on R&D" (ibid., p. 5), and attacks commonly tossed around terminology related to the *knowledge based economy* (predominantly as used by the OECD (1996) as seeming to "cover everything and nothing: all economies are in some way based on knowledge, but it is hard to think that any are directly based on knowledge, if that means the production and distribution of knowledge and information products" (ibid., p. 6). Finally he encourages us to think of knowledge in more epistemological or cognitive terms, and that this will assist in clarifying, or in many instances establishing, definitions of constructs describing knowledge phenomena in a modern societal context.

Recently, added work in clarifying the differing nature of knowledge intensity in low tech industries was done by Hirsch-Kreinsen and colleagues (2008). In line with Smith, they argue that generally, low-tech, non-research intensive industries are misunderstood in terms of what specifically about them makes them potentially innovative, as well as the role they play in the current technological trajectory of the economy and how it will grow and change in the years to come (Hirsch-Kreinsen, Hahn and Jacobson, 2008). The narrowness of Hatzichronoglou's (1997) OECD indicators often miss-specifies how technology may materialize, be used, and is passed on throughout whole economies: Indeed, 'low-tech' sectors have been found to be quite dynamic technologically despite low average sectoral R&D investment (Hirsch-Kreinsen, Hahn and Jacobson, 2008;

Smith, 2009). Von Tunzelmann and Acha (2005) similarly argue that conventional classifications of the OECD fashion are "becoming less and less useful for academic analysis, though their sway still prevails in government policy making" (p. 409). This is because innovative activities in low and medium-tech industries fall outside the Frascati definition of R&D (OECD, 1994), and Von Tunzelmann and Acha emphasize ideas like knowledge search and identification processes as being of high importance to these sectors, not just traditional research and development. As pointed out by Smith (2002) a low R&D Intensive industry may make considerable use of, and development of, knowledge generated in other industries.

Thus, there are useful ways of assessing knowledge intensity for low tech industries that do not rely on R&D indexes, though they can be challenging to capture on the construct level. One important characteristic of knowledge intensity in low tech and medium tech industries is their *absorptive capacity* (Cohen and Levinthal, 1989; 1990), or the "ability to value, assimilate, and apply new knowledge" (Zahra and George, 2002, p. 188) since transforming outside knowledge into new usable knowledge and routines is required to make "productive use of these upstream developments" (Von Tunzelmann and Acha, 2005, p. 4). One caveat however is that absorptive capacity as proposed by Cohen and Levinthal assumes that the firm has sufficient research and development activities internally to understand and apply external knowledge.

One other prominent indicator of knowledge intensity in 'low tech' contexts is that of *organizational innovation*. This is commonly associated with (by the Oslo manual, among other sources) a new or significantly improved system of knowledge management in order to streamline and improve the exchange of information, general skills and knowledge inherent in the organization itself, major change to the organizational infrastructure of the work environment, or new changes related to outside actors like other firms or public institutions in the form of formal or informal collaboration methods (Gallego, Rubalcaba and Hipp, 2013).

Practical or operational knowledge is also argued to be an indicator of knowledge intensity in low and medium tech sectors. This concept refers to an amalgamation of explicit elements such as design and specifications for new products, as well as less tangible constructs such as deeply ingrained experience and routines for problem solving. Commonly, these types of knowledge are associated with learning by doing or using (Hirsch-Kreinsen, Hahn, and Jacobson, 2008; Nonaka and Takeuchi, 1995).

2.2.3 Knowledge intensity in business services

The European Commission (2012) recently stated that the main challenge in studying innovation is now attaining a better understanding of the measure of innovation activities in services and service firms. This, the report claims, is all the more challenging due to the blurred distinction between what constitutes a manufacturing firm *contra* a service firm, since manufacturing firms often offer services, and vice versa (Baines et al., 2009; Strambach, 2009). Given this, it is important to outline what particular type of service firms we are interested in. Services are often manifestations of supplier-user interaction and are co-terminal in the sense that they are consumed at a particular time and place (Miles, 2005, p. 435). The most rigorously studied branch of service activities with regards to innovation are knowledge intensive services, where service activities have a 'high knowledge component' (European Commission, 2012, p. 10), of which *knowledge intensive business services*, or service activities aimed at commercial gains sold to other businesses, are of chief interest. Strambach (2008) describes KIBS as an artefact of cumulative learning between user and supplier (which she derives from Muller and Zenker, 2001); and the consultancy role played by a firm in order to facilitate problem solving by drawing on expert knowledge to serve the client. KIBS firms invest considerably more in innovation than less knowledge intensive service firms (Tether and Hipp, 2002) Innovation in these activities may rely more on social and cultural factors, as well as supplier driven application of new technology, than on technological innovation in the service sector *per se* (ibid.). The term *knowledge intensive*, in the context of service activities, often refers to either *labor qualification*, or specifically to the nature of these *transactions between user and supplier* (Strambach, 2001; Hauknes, 1999), and a knowledge intensive firm is then a firm that undertakes complex operations where human capital is the crucial distinguishing factor (Alvesson, 1995; Muller and Doloreux, 2009).

The literature on KIBS innovation has noted this as a key distinction between innovative activities in KIBS-oriented firms and manufacturing firms (Sundbo and Gallouj, 2000; Camacho and Rodriguez, 2005; Tether, 2005; Freel, 2006; Tödtling et al., 2006; Muller and Doloreux, 2009), in that *training employees and actively pursuing innovation* are often somewhat more pronounced than in manufacturing firms, though collaboration in innovation is somewhat less likely. Thus, investments in human capital such as high quality and highly-qualified employees who undergo intensive collaboration with local users are a defining characteristic of KIBS firms (Muller and Doloreux, 2009).

Table 2.1: Comparing knowledge intensity indicators and sectors

KI - High Technology Sectors	KI - Low to Mid Technology Sectors	KI - Service Sectors
R&D Intensity and expenditure	R&D Intensity and expenditure	R&D Intensity and expenditure
Human capital (founders and employees) in the form of skills, training and education	Human capital (founders and employees) in the form of skills, training and education	Human capital (founders and employees) in the form of skills, training and education
Use and application of new knowledge	Use and application of new knowledge	Use and application of new knowledge
	Organizational Innovation	Organizational Innovation
	Practical or operational knowledge	Training and labor qualification
		Active innovation strategy
		Nature of transactions between user and supplier (knowledge exchange, co-terminal knowledge

As with low tech and medium tech industries, the literature has also emphasized *organizational innovations* as a distinct way for KIBS-oriented firms to innovate in their field, often a combination of technological and soft skills (Muller and Doloreux, 2009).

2.2.4 Knowledge intensity across sectors

By comparing these different measures and ideas of knowledge intensity in diverse sections, a few observations may be made (See Table 2.1). *Despite well documented differences between measures of knowledge intensity across different types of industries and activities, traditional measures of innovative output like R&D Intensity and expenditure remain relevant, though their importance depends greatly on the sub-level categorization of different firms in different sectors.* For instance, more technical KIBS firms like gazelles or new technology-based service firms tend to spend more on, and place higher value on, R&D than non-technical KIBS firms (Tether and Hipp, 2002). *Thus, the argument can be made that R&D can still be a relevant indicator across sectors, the same going for human capital in the form of employee and founder education, prior work experience, training and the imprinting that these actors make on an organization.* For these reasons, R&D intensity may be an effective baseline control for assessing other dimensions of knowledge intensity.

Though not overtly emphasized by this brief review (more will be said about it in the coming chapter), the extent to which a firm utilizes knowledge from external sources for innovation can also be argued to play a decisive role in how knowledge intensive an organization may be, not least in high-tech industries, but in low/medium-tech and services as well. Smith's (2002) distributed knowledge base theory and the knowledge spillover theories of Acs and colleagues (2009) attribute much of the knowledge intensity and competitiveness of ventures to their external knowledge networks. Indeed, low-tech and medium-tech firms are thought to rely heavily on this approach to absorbing external knowledge through search activities (Von Tunzelmann and Acha, 2005). Service firms are also externally reliant for innovation, not least due to the close reliance on customer and client relationships in their innovative processes (cf. Muller and Zenker, 2001)[3].

The idea of knowledge intensity in economic activity has become an extremely prominent topic in both research and policy in recent years. Despite being relatively new in terms of theoretical development, the amount of coverage the subject matter (KIE and KIE-related) receives from both IGOs like the OECD and EU as well as the scientific research listed so far is at least in part a signal that there is some merit to the idea, i.e. that there is something very striking and intuitive about defining economic activities in terms of knowledge intensity and societal benefit, especially given the state of the economic reality of today, with technology and science becoming more and more intertwined with our daily lives. Some research has strived for a more precise way of defining and empirically studying the idea, and out of this has come the literature focusing specifically on knowledge intensive entrepreneurship. The following section outlines some of the most recent work done in this area of entrepreneurship and innovation studies.[4]

2.3 Recent research on KIE: Development and conceptualization

This section reviews some of the recent research placing knowledge intensity specifically into an entrepreneurial context, and building of the KIE concept through a series of international research projects. It reviews the aims of the projects, as well as some key results.

During much of the early 21st century, research on the ‘knowledge economy’ was done largely on the national and regional level, usually

[3]While investigating the other aspects of knowledge intensity contained in Table 2.1, such as organizational innovation and practical or operational knowledge, which may more aptly uncover important relationships found mainly in low and mid-tech sectors, I have chosen not to do this. The main reason being that the project simply becomes too large in scope when these factors are added. Additionally, the empirical data used does not provide adequate coverage in its variables to prove worth using. I hope to cover these aspects more in detail in future projects.

[4]This literature to be reviewed next has served as a springboard of sorts to launch the objectives that will be the main focus of this work. Eventually, the nature of what may constitute measurable estimations of knowledge intensity for entrepreneurial firms will be proposed, and their applicability will be empirically and theoretically examined.

targeting entire industries. The next phase in the stream of interest came with the idea that *new* or *young* firms had an equally large role to play in knowledge-based growth. This is thanks in part to a resurgence of ideas of endogenous technological change in the economy that happened during the 20th century. While early literature on technical change and economic growth identified the prevalence of heterogeneous factors like education, institutions and national contexts driving the process across industries (Nelson and Phelps, 1966; Phelps, 1966; Nelson et al., 1967; Rosenberg, 1974; Nelson et al., 1976; Rosenberg, 1976; Rosenberg, 1982), the idea of endogenous technological change became more widespread in economics around the time that Romer (1986, p. 1003) proposed a model "in which long-run growth is driven primarily by the accumulation of knowledge by forward-looking, profit-maximizing agents". While the Schumpeterian entrepreneur sense has long been perceived as the driving force of the economy in certain circles (see the above references on technical change and growth from Nelson, Rosenberg and colleagues), the specific study of knowledge intensive entrepreneurship has only recently come into a more attuned focus for both scholarly communities and policy makers. This is occurring predominantly in Europe and its surrounding economic zones. While work is still unfolding in this specialized research domain, a few large-scale European Union/Commission projects have been paving the way for future research, and a large bulk of the current studies about the knowledge intensive entrepreneurial firm has been carried out as part of, or in conjunction with, these projects. Two of these are: *Knowledge-Based Entrepreneurship: Innovation, Networks and Systems*, or KEINS; and *Advancing Knowledge-Intensive Entrepreneurship and Innovation for Economic Growth and Social Well-being in Europe*, or AEGIS.

2.3.1 KEINS

The KEINS project attempted to shed light on fundamental questions and unknowns about the KIE phenomenon: What may constitute it, what factors affect it, how it impacts economic growth, and how it is relevant in Europe's advanced and transition economies and manifested in these different growth zones (Malerba and McKelvey, 2010). To ease collaboration between research units and their special skillsets, KEINS approached the concept as more of a flexible umbrella concept under which numerous classifications were possible. The 3 building blocks for KEINS' view of the concept were operationalized and thus the KEINS project utilized three working definitions (or operationalizations) of knowledge intensive entrepreneurship in its analyses: These were: New firms in knowledge intensive sectors (using the common higher education cutoff component); New innovators in a technology or sector; and, Academic inventors (either start-ups or academic patenting activities) (Malerba et al., 2010). Throughout the project, KIE is defined "[concerning] new ventures that introduce innovations in the economic systems and that intensively use knowledge" (Malerba, 2010, p. 4). It analyzed KIE in terms of four varied perspectives: 1.) Industrial dynamics, growth, and development; 2.) New firm creation in knowledge intensive sectors and how they relate to various innovation system perspectives [5]; 3.) New innovating firms across industrial and national borders in Europe; 4.) and academic entrepreneurship and academic patenting.

KEINS also worked to conceptualize KIE on an industry level. Mamede, Mota and Godinho (2010), as well as Hirsch-Kreinsen and Schwinge (2015), as part of AEGIS, adapt sectoral categorizations that fall quite near those established by Hatzichronoglou (1997) and later by the OECD (2002; 2005) in defining low-, medium-, and high-tech sectors and relating them to knowledge intensive activities (European Commission, 2013). They achieved this by using indications of average R&D Intensity ratios in the given sectors. Most documents published as part of KEINS, and later AEGIS, tend to roughly approximate these categorizations, while also including knowledge intensive business services (KIBS)[6]. These business services are specifically defined as "concerned with providing knowledge-intensive inputs to business processes of organizations", and are primarily measured by "educational attainment", usually meaning the percentage of the workforce holding a graduate degree (EMCC, 2005, p. 1).

Other KEINS-based research has attempted to review more established concepts of strategic management and firm theory in a KIE context. Protogerou and Caloghirou (2015) measured how a knowledge intensive entrepreneurial firms' use of certain types of dynamic capabilities impacted its growth and performance by looking at all the firms in the AEGIS survey as if they were KIE firms of varying degrees, and grouped according to a cluster analysis. They looked at firms' market adaptation, new product development, networking, and technological collaboration capabilities, and how these affected performance in terms of firm size, international sales, and radicalness of innovation in products, finding a

[5] Innovation system theory analyzes innovative and technological development as bi-products of complex relationship webs in a given system: be it national, regional, local, technological, or sectoral. See Edquist and McKelvey (2000) for a detailed review of this concept.

[6] KIBS in this case was measured approximately according to the European Monitoring Centre on Change's (EMCC) definition based in NACE 1.1 sectors.

generally positive relationship between the input and output variables, the strongest being on novelty, or radicalness, of innovations produced by the firm.

Overall, the project prescribed deeper analyses of connections between KIE's varying dimensions and how they fit into systems of innovation, the creation of more formalized agent-based models of KIE and industrial dynamics, and quantitative comparison of KIE at different sectoral, regional, and national systemic levels. They emphasized that policy measures must focus on systems approaches to stimulate KIE, competence and capability building are central for aspiring knowledge economies, and that knowledge application to new activities lacks clear focus in policy, which focuses a disproportionate amount of resources on knowledge creation. (Malerba, 2010). The AEGIS study picked up on many of these goals in the years to come.

2.3.2 The AEGIS project

The AEGIS project, in many ways a natural successor to the KEINS project, has, among other aspects, examined:

'[H]ow entrepreneurship is able to foster innovation and economic growth by breaking barriers of various types. The role of entrepreneurship in this process is twofold. On the one hand, the establishing of new firms is often based on new knowledge and on ideas on how to apply it. This means that entrepreneurship, in particular knowledge intensive entrepreneurship, is a direct source of new knowledge and innovation, and thereby it stimulates economic growth.' - Malerba and McKelvey, 2010, p. 29; 2015.

With respect to earlier literature in Innovation and Entrepreneurship studies, this particular perspective on knowledge intensive entrepreneurship as a phenomenon of direct interest evolved largely out of: a Schumpeterian (1942) view of entrepreneurship, where innovation and novelty in firm activity is of course emphasized as the crucial agent of economic growth; and an evolutionary economics approach, popularized by Nelson and Winter (1982) which posits search processes through routine activities as leading to exploration and subsequently exploitation of new and innovative ideas and opportunities, eventually leading to reorganization or change within the firm.

Thus, at the core of the AEGIS project was the argument that entrepreneurship is a chief vehicle or mechanism that translates scientific and technological knowledge into economic growth, innovation, and societal well-being. In addition to this mechanism originating from 'knowledge organizations' such as universities or research and development organizations, it can also stem from joint activities between these actors and the users of new knowledge or innovations and through different spillover-type effects.[7]

By various sectors, Malerba and McKelvey (2010) assert that, building off of rationale used in the KEINS project, knowledge intensive entrepreneurship is not limited only to high tech sectors, as common pre-conceptions in the literature have alluded to. Malerba and McKelvey argue that knowledge intensive entrepreneurship can also occur in low, medium and service sectors. McKelvey and Lassen (2012) conducted a literature review of the KIE concept in the interest of analyzing case studies and studying policy implications in conjunction with the AEGIS project. They later carried out a more detailed study of KIE based on the criteria set out during the project. This dissertation makes great use of the conceptual material devised in conjunction with the AEGIS project. In its recommendations, AEGIS echoed many of the same unresolved prescriptions as KEINS, encouraging policy makers to think more systemically about KIE support schemes, and tailoring them to fit the industrial, national, or regional contexts of the novelty being introduced into the system through new use or application of science and technology. It focused heavily on the implications of KIE on the post-crisis European economy, and the importance of acknowledging the heterogeneity and idiosyncrasy of different types of KIE and KIE firms (Malerba et al., 2015). They point out the need to study pre-entry characteristics of the KIE firm, as well as management and development of the venture in an operational sense, including how it utilizes networks of actors and its existing resources and capabilities to sustain itself and grow.

2.3.3 KIE conceptualized in recent literature through AEGIS and beyond

As noted above, the AEGIS conceptualization of the knowledge intensive entrepreneurial venture was, in part, an attempt to link together different elements surrounding this particular type of company: firm pre-histories,

[7] Holmén et al. (2007) conceptualize innovative opportunities as relying on three core elements: a created economic value for someone; a mobilization of resources; and the ability to appropriate returns or benefits of said opportunity. Additionally, it is a conceptual hybrid of market-based entrepreneurial opportunities (Kirzner, 1997; Shane, 2000), technological opportunities (Scherer, 1965), and production opportunities (Penrose, 1959). None of these concepts as introduced in the literature, as argued by Holmén et al. (2007: 37) "allow for a thorough understanding of the role opportunities play in innovation activities and economic transformation." However, it is these authors' view that an innovative opportunity can manifest as a combination of entrepreneurial, technological, or production-based opportunities.

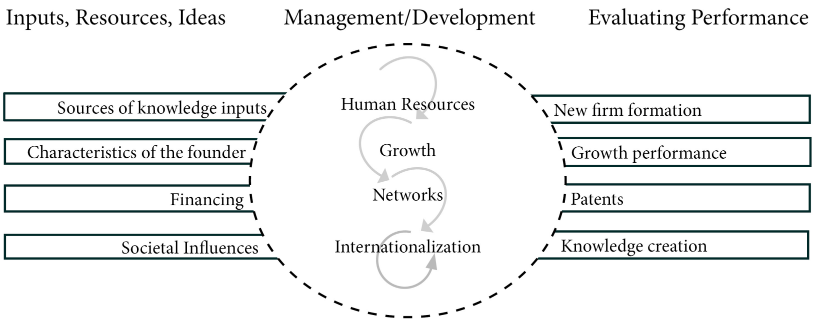
firm level effects such as performance and growth, societal and socioeconomic impacts, and the interface between the firm and the surrounding innovation system (Malerba and McKelvey, 2010). Following AEGIS, more work was done in order to solidify and better describe the KIE firm. The work by McKelvey and Lassen (2013) led to the analysis of three components of KIE: i) Accessing inputs like resources and ideas in the startup phase; ii) development and management of the venture; iii) and the evaluation of performance in KIE, which are illustrated in the KIE creation model shown on the following page. The first column in the figure (Figure 2.1, next page) describes antecedents to firm formation in a KIE context. It concerns accessing internal resources, capabilities, and ideas as well as utilizing networks and relationships to access resources and capabilities external to the firm. The second column describes the management and development process: It takes up aspects of social capital, forming and maintaining network relationships, the growth of the firm in terms of how it allocates resources towards either exploiting opportunities or exploring new opportunities (March, 1990) (this includes the impact of incubation processes, progression from R&D to the market by a product or service, and other dynamic issues) and the international dimension of the firm. The third column describes the evaluation of outputs and performance of the firm, represented by both performance measurement on the firm level as well as measuring output and impact on the economy or society as a whole. Growth performance mainly refers to indicators such as firm survival, sales, turnover, and employees.

A more detailed view of this creation process is illustrated in the systematic literature review conducted by McKelvey and Lassen (2013:179), who argue that:

"By positioning future studies of KIE clearly in relation to the variables described in the conceptual model, it will allow for the development of a more coherent understanding of KIE, how impact is created through KIE, and how policy implication can be targeted to specific aspects of KIE."

Through this detailed review, they identified several popular characterizations of KIE across different industrial settings, as can be seen in Figure 2.1. This classification exercise reemphasizes the fact that what constitutes knowledge intensity is indeed multifaceted, and much depends on the organizational type, the industry, and the relationship between the firm, its innovation system, its external business environment and the knowledge and skills residing in the human minds that drive it in the form of founders, manager, and employees.

Figure 2.1: KIE Creation Model



Adapted from McKelvey and Lassen (2013: 29)

Additional, but related, steps were taken following the AEGIS project to clarify the concept of knowledge intensity in terms of potential empirical indicators. Caloghirou et al. (2015) mapped new firms' knowledge intensity through three different constructs: *knowledge seeking activities*, which relate to linkages that act as sources of knowledge and information for the firm from outside its boundaries; *initial knowledge capital*, or the initial stock of knowledge a founder brings to the venture, measured by average educational attainment and initial funding provided by venture capital; and *human capital and innovation input*, combining employee education and training with R&D Intensity and R&D expenditures. Using these measures, along with others including measures of innovative performance and appropriability regimes, they propose a taxonomy of KIE firms distinguishing highly knowledge intensive firms from those that are less knowledge-intensity reliant based on cluster analysis.

2.4 Reflection on Chapter 2

The objective of this chapter was to identify the conceptual and literary origins of the knowledge intensive entrepreneurial firm. It can be clearly seen that there has been an abundance of work done in the aim of advancing the concept of the knowledge economy as a political and economic tool, and in advancing knowledge intensive entrepreneurship to an empirically researchable phenomenon. However, much remains to be done with respect to deepening an understanding of how and to what extent these firms' inherent knowledge resources and how they use them (i.e. their knowledge intensity) actually relates to affects their performance, whether in operational or economic terms. As Keith Smith (2002) states, it is hard to imagine a time when human society was not based on knowledge, but today's firms' reliance on science and technology for competitive gains has never been more prominent. Many have tried to research the idea of knowledge intensity, but it has proved a difficult topic, demanding often compromise on behalf of the researcher in order to cover all forms of relevant economic activity. The research projects KEINS and AEGIS have been successful in tangibly assessing knowledge intensive entrepreneurship, and serve as a way forward conceptually. Indeed, McKelvey and Lassen's (2013) scheme of input, process and outputs of KIE provide good ground for this.

Both AEGIS and KEINS have called for more in-depth research, and policy, concerning KIE. Important future goals from the former included furthering understanding of pre-history resources and capabilities, as well as how firms manage and draw from networks of external actors, and how this effects venture performance and sustained growth. Broad systemic frameworks have been devised, placing the knowledge intensive entrepreneurial firm at the center of a web of dynamic stage-specific interactions, coupled with the innovation system, regional factors, and external business and social actors of all many types (Malerba and McKelvey, 2015). However, more work is needed in order to measure and assess more specific relationships between inputs and outputs of the KIE firm, and how external interactions influence development over time. What has been shown is that there is yet more to be understood about exactly how internal and external knowledge coexist in KIE firms, and how this drives performance. Here, the work in this dissertation makes a contribution.

From this overview of the KIE area of study, it becomes clear that a distinctive conceptual framework will be beneficial in making further empirical sense of the concepts and constructs used in this thesis. In order to derive this, this report now turns towards key concepts closer to the nature of the firm and delves deeper into different theoretical forms of knowledge intensity in order to refine the scope of the project, at the same time further tracing the gaps in academic knowledge in the literature into which the work as a whole may be placed.

Chapter 3

Explicating knowledge intensity and performance using resource-based theory

This chapter *builds a conceptual framework through which the entrepreneurial firm can be analyzed and understood in terms of how its knowledge intensity impacts performance, and how different types of knowledge intensity and performance may impact each other.* The previous chapter was grounded predominantly in work carried out in regards to knowledge intensity and knowledge intensive entrepreneurship in terms of how they function within different systems of innovation; nationally regionally and sectorally. This mainly looked at KIE phenomena through looking into national policy, economic and industrial classifications, and explorative EU level research projects. This was done in order to provide a brief, but necessary, background about knowledge intensive entrepreneurship. While some theory has been covered, there is more that could aid in analyzing KIE. I now turn to more individual- and firm-centric views of the concept. Several theoretical aspects of knowledge intensity are covered in this chapter to develop some research objectives for investigation, as well as the operationalization of knowledge intensity for the entrepreneurial firm for the purpose of empirical analysis. The last chapter explained that KIE firms intensively use and apply novel knowledge as their chief competitive resource (Malerba and McKelvey, 2015). I will argue that by applying resource-based thinking to the idea of the KIE firm, one can trace how this knowledge intensity is constructed, and that it matters for their performance outcomes. This is done first by reviewing theories of the firm that were instrumental in the derivation of the resource-based view of the firm.[1] The following section, 3.1, reviews these firm level theories that inform the development and direction of the literature review, of the research objectives, and of the subsequent hypotheses. Following this review, the chapter turns to the application of these ideas of knowledge intensity of the entrepreneurial firm: It aims to:

[1]Linking the selection of the theories reviewed here is an underlying reliance on evolutionary economic thinking. The communality expressed here in all the theories covered is based in an evolutionary economics perspective of firms; that firms are heterogeneous, and should be studied as such (Nelson, 1991), and in order to properly evaluate knowledge intensity in entrepreneurial firms, we need models that take into account this heterogeneity.

- Explain notions of how these (and other) firms search for external knowledge.
- Explain how pre-history factors and conditions of the firm such as human capital of founders and employees affect the firm's development and performance.[2]
- Investigate the relationships within the idea of knowledge intensity, specifically, how internal knowledge intensity (firm pre-history; through founder, team and employee human capital) might relate to external knowledge intensity (external search propensity and reliance on specific knowledge sources).
- Cover various performance measures commonly used in studying entrepreneurial firms and how innovative performance may influence business performance of new ventures.

3.1 Theories of the firm: connecting knowledge, resources and capabilities

Theories of the firm are often used to analyze and describe organizational behavior in business, and are abstract methods of addressing sets of behaviors and characteristics exhibited by real businesses (Machlup, 1967). Some also focus more specifically on how firms come into being.[3] Since the emergence of the study of market activity, numerous scholars have attempted to explain why firms exist, and account for how they are distinct from the market in general. These early views include most notably Adam Smith's account of task- or knowledge-specific specialization processes, namely the division of labor (Smith, 1776/1887). In the early neo-classical (and early Austrian) period of economic thought, scholars like William Stanley Jevons, Alfred Marshall, Carl Menger, and Léon Walras pioneered the so-called 'marginal revolution' of economics, through which the characterization of diminishing marginal utility in production functions became a widely heralded tool at the forefront of microeconomics. The firm in these early models was largely reduced to a "set of supply and demand functions" by those that studied them (Penrose, 1985, in Pitelis, 2009). This became more of a controversial point in economic theorizing during the first half of the 20th century (see commentary on this in Machlup, 1967), and new takes on

[2] It should however be noted that due to a lack of literature specifically regarding knowledge intensity in entrepreneurial firms, much of the review has its basis in small enterprise research.

[3] There are of course many more theories of the firm. Here, the scope is delimited to those prominent theories that address the nature of resources, capabilities, and how firms draw on their environments.

firm involvement in economic theory and why they exist at all became a topical issue. In a seminal article for its time, Coase (1937) described the firm as existing due to the variable cost functions present in economic activity, and that by building a firm the costs of doing business are sufficiently minimized to justify the existence of the firm rather than a system built solely on open market transactions. This theory of the firm became known as transaction cost economics (TCE). TCE characterizes firms as constantly making decisions about how to produce goods and perform activities while minimizing the cost of transactions, with this cost minimizing being the chief cause in building an organization (cf. Williamson, 1979; 1981). A related theory is that of industrial organizations (IO), which is intent on describing the firm as a product of its environment, with the structure and dynamics of the industry being the determinants of firm performance (Mason, 1939; Bain, 1968; Porter, 1980; 1985). These theories all have the base assumption of viewing the firm as a profit maximizing entity, working at the margin; often referred to as marginalizing or maximization behavior (Machlup, 1967). While these economic theories of the firm were extremely influential, they received some complementarity as well as some competition from scholars attempting to break open the *black box* of the firm, namely: What exactly happens inside the firm and why? This question is extremely relevant when we want to look at and assess the knowledge intensive entrepreneurial firm, since it is here defined in terms of its *unique* application, transformation, and use of scientific and technological knowledge through their *unique* set of organizational resources and capabilities. More evolutionary branches of economics, based off of groundbreaking work by Joseph Schumpeter, would more directly address this issue. Before delving into this area however, there are a few additional scholars and theories which should be mentioned regarding the study of firm behavior.

During the mid-20th century, different theories began to further question the neoclassicism in firm theorizing, of which the most important was, in business and management science, the following: Cyert and March (1963) as well as Simon (1955) began a discourse about a behavioral view of firm activity. These positions stressed the bounded rationality of individuals in firms, through the idea that decision making cannot always be an act of utility maximizing due to the inherent complexity of real life problems (an idea later embraced by Nelson and Winter (1982)). These positions stressed that firms are not all-knowing, all-seeing entities as economists had commonly postulated in their theories; their knowledge and decision-making power remains imperfect. So, a firm must make the best of its bounded knowledge and rationality, and make uncertain decisions to the best of their current means, knowledge, and resources; a so-called act of *satisficing* (Cyert and March, 1963; Nelson and Winter, 1982). Stemming from the theory of uncertainty management being faced by firms, Cyert and March (1963, p. 95) stressed the notion that organizations, as coalitions of individual stakeholders, have certain organizational expectations about their environment, and these expectations are made up decisions regarding how and when organizations search for, or "prospect" external information and how this is then processed. Thus, the firm uses search as a tool to solve problems (often specific ones), and is a means for achieving prolonged competitiveness. What type of search is employed, Cyert and March have argued, depends on the amount and nature of organizational slack, or, free mobility within the organization of (predominantly) untapped resources and capabilities.

Just before Cyert and March began refining their behavioral theories about firm activity, other influential developments in firm studies were also taking place. Contrasting the views purported by transaction cost economics, and somewhat complementing the industrial organizations approach; another of the most influential theories in management and organizational studies in recent decades has been the *resource-based view of the firm* (RBV) (cf. Penrose, 1959; Wernerfelt, 1984; Barney, 1991). These ideas originated largely with Edith Penrose, in her book *Theory of the Growth of the Firm*[4]. Edith Penrose characterized the firm as a collection of productive resources, both human and non-human, which are coordinated and channeled into the sale of goods and services through the market (Penrose, 1959; Pitelis, 2009). While her ideas put forth in this work were never truly embraced by mainstream economics (something that Penrose did not directly aspire to do, but was criticized for nonetheless by her peers), strategy and management research is heavily indebted to them, and her model was adopted by many influential scholars in these fields.

The basic tenet of the resource-based view is that the firm houses a unique bundle of resources and capabilities, tangible or intangible, that constitute its sustained competitive advantage, and it is the duty of managers to maximize the deployment of these resources and capabilities in the present as well as in long- and short-term time horizons

[4]In addition to theories of the firm that discuss why firms existence, there is an extensive literature on why and how firms grow. One of the more prominent in use today relating to entrepreneurship is the knowledge spillover theory of entrepreneurship (Acs et al., 2009). It discusses how increases in the stock of knowledge in a region has a positive effect on entrepreneurial activity, and that efficient usage of knowledge by incumbents produces a smaller effect on new knowledge resulting from entrepreneurship. This theory, while holding potentially interesting ground for further discussing KIE, is not covered in depth here beyond its mention in Chapter 2.

(Wernerfelt, 1984; Grant, 1996).[5] Thus, if the firm is to achieve this advantage, it must both acquire and control resources and capabilities, and have the organizational competence to both absorb and deploy these resources and capabilities (Kraaijenbrink et al., 2010).

This distinction between absorbing (or picking) resources and deploying them has led scholars to highlight the dynamic and often intangible nature of resources and capabilities, leading to several offshoots of the RBV in recent years. One extension of this view is found in the *knowledge-based view of the firm*, where knowledge is seen as the firm's most important strategic resource (Kogut and Zander, 1992; Grant, 1996), and that the traditional resource based view has overlooked the required collective knowledge and skills needed to coordinate resources into valuable organizational assets (Spender, 1994; 1996). Kogut and Zander (1992) positioned their view of the firm as one where a firm's existence is ratified by the fact that is a better mechanism for sharing and transferring knowledge of individuals and groups than the open market is, choosing to focus on knowledge rather than transactions as the defining efficiency of the firm. They argued that firms can create new skills and knowledge by recombining their resources and capabilities. Grant (1996) focused on the organization as the mechanism of application of knowledge, whereas the *creation* of knowledge occurs on an individual level.[6] The knowledge-based view has been extensively used as a scientific tool in trying to unravel how knowledge flows occur in organizations and their networks, often scrutinizing the difference between tacit knowledge, or in a rough sense, non-codifiable knowledge[7] that we observe that we know only through its use to identify lower-order knowledge, and *explicit knowledge*, which may be taught and/or written down (Polanyi, 1966), and how these types of knowledge might be best harnessed and transferred within and between organizations.[8]

[5]While it has been extensively argued in the literature whether or not the resource- or knowledge- based view of the firm can sufficiently constitute an explanation for the existence of the firm, contra other theories of the firm, specifically TCE theory (Barney 1996; Conner and Prahalad, 1996; Foss, 1996a; 1996b; Liebeskind, 1996), what can and has been agreed upon is that this type of way of viewing and analyzing firm activity in terms of resources and capabilities affecting competitiveness can be extremely useful, and relatively easy to understand for scholars and practitioners (Kraaijenbrink et al., 2010).

[6]Consequently, this individual dimension of knowledge creation harkens back to early roots of innovation and entrepreneurship theory, the Schumpeterian (1934, 1942) inventor-innovator, as well as the evolutionary approach to the firm popularized by Nelson and Winter (1982) in which routine work patterns of individuals change and adapt over time, giving life to organizational novelty and variation.

[7]Polanyi (1966: 24) puts it more eloquently: "Tacit knowing is shown to account (1) for a valid knowledge of a problem, (2) for the scientist's capacity to pursue it, guided by his sense of approaching its solution, and (3) for a valid anticipation of the yet indeterminate implications of the discovery arrived at in the end."

[8]Grant (2002) has also emphasized the differentiation between two distinct types of knowledge-based activity in economic terms: Those that are concerned increasing a given stock

Another recent take on resources, capabilities, and competitive advantage is found in the *dynamic capabilities* literature (Amit and Schoemaker, 1993; Teece et al., 1997; Makadok, 2001), which in some ways evolved as a response to the somewhat 'time invariant' nature of the traditional RBV as portrayed by, among others, Barney (1991). This view focuses more on the building and harnessing of capabilities than the cultivating and picking of resources. It takes a highly Schumpeterian view of capabilities in that they are ever-changing, ever-evolving, and never static if competitive advantage is to be sustained. Dynamic capabilities were later put forth as the mechanism by which firms can derive value from their resources, without which these resources are of no discernable value (Eisenhardt and Martin, 2000).

Yet another recent contribution to this school of firm studies is the problem-solving perspective of knowledge-based firms (Nickerson and Zenger, 2004). This perspective focuses on the efficiency of generating knowledge or capabilities, and argues that firms' ultimate goals are problem-based, since the outcomes cannot always be determined and finding the right problems to solve and matching the required or estimated resources to these problems largely dominates firm strategy (ibid.).

The theories introduced here, drawing from the resource-based ideas of Edith Penrose, and the behaviorist traditions of Simon (1955), Cyert and March (1963) overlap in the sense that they view the firm as a unique product of its resources and capabilities, its assumed sense of rationality, and ultimately its environment. The external and internal environments of the firm are seen through a subjective managerial filter, and the adjustment mechanism between the two is seen as imperfect in both theoretical viewpoints, creating inherent heterogeneity in firms (Pitelis, 2007). All these scholars were influential in placing an emphasis on the relationships between growth and profitability of firms with their organizational structures, their resources and capabilities, and their behavior (Nelson and Winter, 1982). Firms do, and must, differ from one another, especially for analytical purposes. Building theories that embrace firm heterogeneity has gone against much of the neoclassical literature based on general equilibrium theory, which argued that firms

of knowledge, and those that are concerned with deploying knowledge in order to capture value through remuneration of goods and/or services. These concepts exist in dichotomous form in the literature under various monikers; most notably perhaps, exploration and exploitation (March, 1991) and knowledge generation vs. knowledge application (Spender, 1992). The key to a firm unlocking its full potential is then of course to strike a balance between these two fundamental activities, by allowing individuals to apply specialized knowledge to production, while simultaneously or in tandem, preserving their efficiency in acquiring knowledge (Demsetz, 1991).

faced given sets of choices constrained by technology, and how these decisions were actually made were in effect black boxed (Nelson, 1991).

In the context of knowledge-intensive entrepreneurial firms, how the firm differs from the next, and how idiosyncratic resources and capabilities are turned into sustained advantage are paramount. Therefore, the traditional resource-based view and knowledge-based view of the firm constitute an important theoretical pillar upon which this dissertation builds its main operational framework: where *generalized resources and capabilities are taken as input measures and performance and competitiveness indicators are taken as output measures.* Another such pillar involves the idea of evolution of organizations through different types of routines, one of which involves the search for new ideas and knowledge residing outside the boundaries of the firm.[9] The first topic here will be the evolutionary approach to economics alluded to above. This will be covered next.

3.2 Search processes of firms and technological innovation

As introduced above, Cyert and March (1963) utilized the idea of search in their behavioral theory of the firm. Firms conduct search in order to solve specific problems, beginning with search based in problem areas (those closely related to the issue at hand), and moving gradually further away to more distant potential solutions. Relying somewhat on these ideas, Nelson and Winter (1982) applied the concept of *search* in relation to their evolutionary theorizing on firms, whereby the term is used “to denote all those organizational activities which are associated with the evaluation of current routines and which may lead to their modification, to more drastic change, or to their replacement” (ibid., p. 400). In essence, search, and its counterpart, selection, are the means by which organizations change. These means, Nelson and Winter argue, are largely based in routinized processes. These are often undertaken with the goal of enhancing future profit (Zollo and Winter, 2002). Firms perform numerous search activities, ranging from organizational issues to production methods to implementation of new ideas and innovation (Katila and Ahuja, 2002). Additionally, the firm is chiefly interested in utilizing its distinct resources and capabilities in order to achieve some form of sustained competitive advantage, what Penrose referred to as “wide and relatively impregnable bases” from which to adapt and survive in uncertain environments (Penrose, 1959, p. 137). This almost

[9]This framework seems well suited to analyze KIE ventures, with their knowledge intensity conceptualized through resources and capabilities and their sustained competitive advantage in terms of performance outcomes.

always will involve Cyert and March's organizational expectations-driven search activities, or Nelson and Winter's search-driven routines, to change or augment the stock resources and capabilities of the firm.

Indeed, much of the prominent literature in innovation studies deals with the idea of organizations *searching* for new knowledge in order to innovate. Much of this research has differentiated between internal and external search processes, or local and non-local search processes (Fleming and Sorenson, 2001; Katila and Ahuja, 2002; Laursen, 2012), and how these search processes are used for both explorative and exploitative aims (March, 1991). While explorative search focuses on distant or non-local sources of knowledge, exploitative search may refer to internal or local search; utilizing in-house sources of knowledge, or that which lies within the current knowledge base as embodied within the firm (Helfat, 1994; Fleming and Sorenson, 2001). Moreover, there exists a tradeoff between internal and external search strategies in terms of what a firm can gain through search and what it can effectively take advantage of (Laursen, 2012).[10]

These above-named concepts in the literature have the potential to be applied to KIE firm-level analysis; but what research, even if limited, has been carried out in these areas with special regards to KIE firms? While not specifically using the 'search' terminology, KIE theory has touched on this type of entrepreneurship's role in terms of economic value creation, impact and interaction within innovation systems, and that entrepreneurs have a reciprocal relationship with system actors which interact with their origins, growth and performance (Malerba and McKelvey, 2015). Additionally, McKelvey and Lassen conceptualize the KIE firm as drawing from the external environment and specifically by outlining external knowledge and resources as being critical in their influence over the decision-making of the founders and managers of KIE firms, especially in the initial developmental phases.

There has also been some research regarding both pre-historical and search properties of *entrepreneurial* firms in a broader sense, falling into various taxonomies and classifications. The majority of studies have been until very recently directed towards mapping the search strategies of medium to large-sized manufacturing (and more recently, service-oriented) firms and how these patterns affect performance, often measured through different forms of innovation output (Helfat, 1994;

[10]In the context of this dissertation, I view the exploitative and explorative processes of search as complementary, not orthogonal. Exploitation then may involve the firms' drawing on its own pre-entry resources and capabilities, while exploration refers to harnessing external knowledge for innovation and organizational renewal.

Fleming and Sorenson, 2001; 2004; Katila and Ahuja, 2002; Laursen and Salter, 2006, Laursen and Salter, 2014).[11] In innovation studies, the search and selection of new ideas become critically important, and firms' use of search for innovation purposes has become an incremental component of judging how competitive an organization might be in the long term. Indeed, a substantial body of literature has addressed how firms (not only KIE firms) carry out search processes spanning technological and organizational boundaries in the aim of attaining process and product innovation (see Laursen, 2012).

Based on what is known about KIE firms, it can be asserted that the result of these search processes is often closely linked to the desired inflow and introducing of new products and processes, often manifested through novel scientific and technological knowledge, into the firm's current set of routines. Malerba and McKelvey (2015) have pointed out that contextual linkages between other actors in the innovation system play a major role in KIE and how it manifests. Public research organizations, universities, and other academic bodies are active in KIE generation on different planes, including the commercialization of academic research. Users also play a role, stimulating entrepreneurial processes and innovation in different ways and to different extents, often through demand side conditions (ibid.). They convey these context specific interactions as having great influence on the outputs generated by KIE.

3.3 Measuring performance of the entrepreneurial venture

Since we are concerned not only with the knowledge intensive resources and capabilities deployed by entrepreneurial firms, but also with the effect of these artifacts on the firm performance, a brief overview of evaluating performance of entrepreneurial ventures is warranted before continuing. Evaluating the performance of new ventures, not least KIE ventures, is not always a straightforward endeavor. Venkatraman and Ramanujam (1986) distinguish between different layers of new venture performance, ranging from financial measures, which represent a venture's overall economic attainment, to operational measures, which in turn may lead to financial measures. Innovative performance is one such operational measure.[12] Neither are easy to objectively apply to firm studies.

Traditional financial measures often fall flat in this regard, either due to

[11] See Laursen (2012) for a more extensive overview of such studies.

[12] Along with market-share, quality of product, introduction of new products, value added in manufacturing, and other "measures of technological efficiency" (Venkatraman and Ramanujam, 1986, p. 804).

non-availability of data or non-divulgence by the focal firm. Accounting or strictly financial measures have been thus deemed often inappropriate as sole measures for performance in younger firms (Shane and Stuart, 2002; Clarysse et al., 2011). Additionally, growth rates, and thus financial incomes, can differ drastically from venture to venture depending on a wide array of contextual, often industry-specific factors producing many outliers in any sample, making statistical analysis a difficult tool of choice (Chandler and Hanks, 1993).

Objective measures like survival seem like a better alternative at times, but this is also not without drawbacks. To properly assess some sort of survival indicator, one needs an adequate timespan of data from which to draw, which is not always available to the researcher (ibid.). One may also encounter the problem of unobserved heterogeneity, and selection bias, as argued by Heckman (1979).

Chandler and Hanks (1993; 1994) compared a number of different techniques often used in tandem with self-report questionnaires; they looked at analyzing performance via broad categories, some type of owner or manager satisfaction criteria, and via competition relative to competitors. They found that measuring performance via growth (market share, cash flow, sales) and business volume (earnings, sales, and net worth) were more reliable and valid than subjective measures of satisfaction or competitiveness.

Venkatraman and Ramanujam (1986) state that evaluating performance via only either primary or secondary data sources, or only either operational or financial indicators, is often inadvisable. The researcher should aim to, when possible, combine these different types of measures for the best operationalization. More specifically to knowledge intensive entrepreneurial firms; McKelvey and Lassen (2013) conducted an extensive systematic literature review on the concept, and found a series of chiefly employed performance measurements (see conceptual framework in Figure 2.1 in Chapter 2), as well as proposed what they argue to be the most important indicators of KIE performance. Firms' knowledge creation output, patenting, and growth and economic performance, and actual successful formation of the venture, are among these, and are often used to measure success in KIE according to their analysis.

Additionally, looking at the operational performance in terms of innovation performance should be particularly relevant as a performance indicator for more knowledge intensive entrepreneurial firms, since they are proposed to derive their competitive advantage from and impact society through the application and commercialization of new scientific and technological knowledge. Other measures common to entrepreneurial ventures will be additionally enlightening. How a new firm is able to survive market environments and grow and adapt are also especially interesting in a KIE context. Financial indicators as well, while perhaps imperfect as a standalone indicator of new venture performance, when combined with other performance measures as suggested by Venkatraman and Ramanujam (1986), will supplement the overall performance construct to be used in this dissertation.

To sum up, it seems advisable to use a broad set of performance measures, both in terms of volume as well as growth, when possible, to estimate entrepreneurial performance. Combining secondary data with primary, opinion-based survey data, with survival indicators will be the way forward in this dissertation, and when possible, looking at growth (slope), though primarily at volumes (intercepts). More details are to come.

3.4 Resources, capabilities, search, and performance in new firms

Given the ground work laid by the previous section, I will now work towards applying the theories of the firm discussed above to the entrepreneurial firm, and when possible, specifically to KIE firm types. This application is necessary in order to contextualize more specific research objectives exploring how knowledge intensity affects the performance of new ventures. These will be introduced again throughout the chapter, followed by some initial hypothesis generation where appropriate. It is, however, important to note that not all of the relationships to be investigated will be straightforwardly hypothesized based on previous research findings. Given that some of the relationships that I will propose have seen limited coverage in the literature, some will be more akin to working hypotheses. I will denote these specific hypotheses with special coding in this chapter (WH), so that the reader can easily interpret them later on. Much of the literature to be covered does not deal specifically with the KIE concept, due to its relative newness, so there will be some limitations there as well. What I will do, is try to align the hypotheses with the closest representations of KIE firms that I can find, and derive hypotheses in that way.

The view of knowledge as the most important source of competitive advantage in the firm, as put forward in most variants of the knowledge-based view, is a position that is shared by much of the research on knowledge intensive entrepreneurship. However, KIE research is perhaps more interested in the use and application of novel scientific and/or technological knowledge, and how this is harnessed as a resource through the organizational capabilities and skills of the new venture. This application is carried out by firms in order to compete, grow and survive. This then is more crucial to KIE than the more simplified acknowledgment of 'knowledge' as a chief competitive resource purported by knowledge-based theories and views of the firm. This type of application has influenced my choice of structure for the coming pages. Section 3.4.1 focuses on *(largely innovation specific) search processes related to augmenting performance in young firms*, while 3.4.2 will focus on *pre-history resources and capabilities of the firm in the form of human capital and organizational origins and their impact upon business performance.* Later, studies carried out regarding resource-driven performance of different types of entrepreneurial firms will be reviewed, ranging from knowledge intensive entrepreneurship to general studies about new firms. Section 3.4.3 will focus on *both the logical and theoretically established links between internal knowledge intensity and external knowledge intensity* as I interpret them. Finally, Section 3.4.4 take on the assessment of performance of the (knowledge intensive) new venture and proposes some key relationships among two such constructs, innovative performance and business performance.

3.4.1 Technological search processes as affecting firm performance

The research regarding search activities of firms has related heavily to innovation. Since KIE firms, according to the theory outlined in Chapter 2, rely highly on their innovativeness, this dissertation is also most interested in search activities that moderate or amplify a firm's innovativeness.[13] As touched upon earlier, much research has differentiated between search processes either local or non-local to the firm (Dodgson et al., 2014). Additionally, the literature on search for technological or innovative purposes has become heavily tied to the conceptualization of *open innovation* in recent years (Chesbrough, 2003; Laursen and Salter, 2006; Dahlander and Gann, 2010; Arora et al., 2016), and often through more generalized conceptions of openness, whereby firms are measured by the extent to which they draw on relevant non-local knowledge from their environments, i.e. different actors in their innovation systems. Openness is defined in this context as, *the degree to which firms are open to external sources of knowledge in their innovative and entrepreneurial processes.* Much of the literature linking external

[13] Though, it may be acknowledged that search can occur in any number of other areas in which a firm is striving for some form of renewal or change

sources of knowledge to innovation output has focused on manufacturing firms of substantial size through large-scale surveys (cf. Levin et al., 1987; Klevorick et al., 1995; Laursen and Salter, 2006). Some research has begun to focus on the performance benefits (and costs) of openness to external sources of knowledge for small and medium-size (and occasionally, entrepreneurial) firms of differing nature (de Jong and Marsili, 2006; Keupp and Gassmann 2007; Nieto and Santamaría, 2010; van de Vrande et al., 2009; van de Vrande et al. 2010; Forsman and Rantanen, 2011; Vanhaverbeke, 2012; Lasagni, 2013; Love et al. 2014). Generally, the research on entrepreneurial networks and how they enhance firm performance has been plentiful, but considerably less research takes a detailed view of the extent to which small and micro-firms derive value from external sources of knowledge and how this affects performance (cf. Brunswicker and Vanhaverbeke, 2015; Keupp and Gassmann 2007; van de Vrande et al. 2010; Love et al. 2014). Since entrepreneurial firms are often highly embedded in and dependent on networks within systemic contexts, analyzing how they search for new sources of knowledge and what sources benefit what firms is a very pertinent matter. This is especially true for highly knowledge intensive entrepreneurial firms, for which there are very few studies addressing them directly within the context of external knowledge sources and how these are used to enhance performance. Some relate more closely to KIE than others, but since the terminology is not that well developed in the literature as of yet, the researcher must needs draw on different proposed organizational forms. It becomes clear that there are a variety of measures by which one can evaluate different conceptualizations of "innovativeness" or "external knowledge sources". Based on the literature review, this dissertation has selected its measures strongly out of the motivation to advance knowledge around KIE firms along similar lines as has been used in past studies. It focuses on a variant of Laursen and Salter's (2006) well-accepted conceptualization of openness to external knowledge sources to represent *external knowledge intensity.* The following research objective is one subject of the first empirical chapter:

> **RO1**:*Explore the association between external knowledge intensity and innovative performance in the entrepreneurial firm*

Table 3.1 presents a collection of studies addressing performance in terms of innovation by means of external knowledge source utilization on the firm level.

Table 3.1: Key findings about the impact of openness to external knowledge sources on performance in (small and micro) firms

Authors	**Input Concept**	**Input operationalization**	**Output concept**	**Output operationalization**
De Jong and Marsili (2006)	Innovative outputs and inputs; Sources; managerial attitude; planning; external orientation	Product and process innovation; innovation budgeting; suppliers, customers, scientific development; documented plans; consultation of external sources, inter-firm cooperation	Taxonomy,of small innovative firms	Science based firms, Resource intensive firms, Supplier dominate, Specialized suppliers
Laursen and Salter (2006)	Openness to innovation	Breadth, depth, collaboration breadth	Innovative performance	Share of new to the firm, market, and world innovations to sales.
Van de Vrande et al. (2010)	Technology exploitation; Technology exploration	NA	NA	NA
Nieto and Santamaría (2010)	Collaboration; Size	Collaboration with external firms or research orgs (binary); Small, medium, or large firms.	Innovation output	Product and Process innovation (binaries)
Forsman and Rantanen (2011)	Innovation capacity	R&D investment; Degree of innovation capabilities; External input into innovation development through networking	Innovation output	Types of innovations; Degree of radicalness of innovations; Diversity of developed innovations
Leiponen and Helfat (2011)	Centralization of R&D activities, external knowledge sources for innovation	Locations of R&D; Importance of external sources of knowledge	Innovation output	Binary variables representing 1) any innovations; 2) product innovations; 3) new to market; Log variable of innovative sales; Innovative impact (binary)
Keupp and Gassman (2013)	Resource constraints	Financial and knowledge constraints (PCA derived)	Radical Innovation	Share of innovative products to total sales
Lasagni (2013)	External relationships for innovation	Evaluation of external relationships for innovation	Innovation Performance	Range of innovativeness; Turnover from new products.
Love et al. (2014)	Openness	Breadth of linkages; controls	Innovation output	Proportion of innovative products to total sales in 3 year increments; Also, and separately for the last 3 years: Measuring innovative portfolio and value .

3.4.1.1 External sources of knowledge and innovation in entrepreneurial ventures

This section addresses the extent to which both manufacturing- and service-related entrepreneurial firms' innovation performance may be affected by their external search propensities. A natural point of departure is to review the research on external search activities of established firms which have experienced the most attention in the field of innovation studies, namely, large manufacturing firms. The literature has established that innovation in these firms is strongly impacted by external sources of knowledge. In part, the Yales survey of R&D managers demonstrated this (Leven et al., 1987; Klevorick et al., 1995). Additionally, it has demonstrated that these external knowledge sources are vital to a wider spectrum of activities, including how technological opportunities, appropriability regimes, and access to knowledge through different channels could vary by industry. Since these early observations of the linkages between external knowledge and firm innovativeness, the literature on use of collaboration as a means to access and to utilize knowledge from wide ranges of partners has materialized and been refined (Dodgson et al., 2014). While this type of network-based innovation literature has been around for some time, much of the more *recent* work has been carried out in relation to the *Open Innovation* heading. This label is defined by its main progenitor, Henry Chesbrough (2003, p. XXIV) as: "a paradigm that assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as [they] look to advance their technology". In a widely cited review of the open innovation literature to date, Dahlander and Gann (2010) argue that the literature has focused upon the potential positive effects of open innovation, while pointing out the limitations and risks of open innovation to the firm.

Strongly linked to the open innovation paradigm, a substantial body of literature has addressed how firms carry out technological search processes which span boundaries of technologies and organizations, in the aim of attaining process and product innovation (see Laursen, 2012). Indeed, much of the technology and innovation management literature deals with the idea that organization must search for new knowledge in order to innovate. Much of this research has differentiated between search processes which rely on sources that are internal to the firm, and those that rely upon external sources to the firm (Dodgson et al., 2014). A popular analytical categorization for this distinction has been how far away this search is from the existing knowledge base of the firm, known as the difference between local and non-local search processes (Fleming, 2001; Laursen, 2012), and whether these processes are used for explorative and exploitative aims (March, 1991).

Also, the broader technology and innovation management literature includes a series of contributions which examine if and how the firm benefits from being able to collaborate with different types of partners, and how well the firm utilizes its position within a network (Carayannis and Alexander, 1999; Chesbrough, 2003; von Hippel, 1988; Uzzi, 1999). Similarly in this vein, innovation systems literature has examined the actors, linkages, and roles of collaboration and networks amongst firms in a systemic environment: The goal of which being to explain the relative economic competitiveness of different nations and industries (Edquist, 2006).

While the innovation management literature has focused upon networks and collaborations for the large firms, the entrepreneurship literature has focused upon how start-up ventures are dependent upon networks to access resources (see review in McKelvey and Lassen 2013a). Therefore, deriving a good deal of theory from studies of large firms, a few recent pieces of literature have stressed the need to understand how and why networks and collaboration affects innovation within smaller (and sometimes new) ventures (Forsman and Ratanen, 2011; van de Vrande et al., 2008; Vanhaverbeke, 2012)[14]. Smaller firms can use networks and inter-organizational collaborations to increase their overall innovative capacity (Szeto, 2000; Caniels, 2005; Forsman, 2011). Moreover, collaborations around technologies may enhance the innovative capacity of firms (Alonso and Bressan, 2014), and achieving better innovative capacity has been argued to be of critical importance for the smallest of firms (Nieto and Santamaría, 2010). Smaller firms which "are aware of and use external information" (de Jong and Marsili, 2006: 221) were found in one study to be better off in terms of introducing successful innovations than those which do not.

Likely, the learning capabilities of SMEs are enhanced through cooperative networks (Chell and Banes, 2000; Mäkinen, 2002; Reinl and Kelliher, 2010), and there are performance enhancing properties of learning from and interacting with larger organizations (Anderson and Lööf, 2012). Hence, there are likely positive effects of networks and collaboration on the ability of SMEs to gain improved knowledge, acquire new access to markets, and reduce research and development costs (Glaister and Buckley, 1996; Forsman, 2011).

[14] Ranging from 0-50 employees, with micro firms being those with 10 employees or less.

However, small firms will also face challenges in accessing external sources of knowledge. The literature suggests that most types of small firms will face innate challenges and liabilities of newness and smallness, among which, intense inter-firm competition and highly limited resources and experience, which limit their collaboration with external partners (Kotey and Sheridan, 2004; Franco and Haase, 2010; Thorgren et al., 2012; Alonso and Bressan, 2014). Specifically, these types of resource constraints range from management, labor skills, lack of finance and information (Cressy and Olofsson, 1997), to lack of owner-specific and organization-specific resources, all of which are crucial for performance and growth (Brush and Chaganti, 1999). Due to their periphery status in production chains, smaller enterprises often face steep challenges when attempting to use networks to increase their competitiveness (Forsman, 2009; 2011). Additional firm capabilities to manage interactions are required for reaping the benefits of a network. Thus, a firm without these capabilities, or a firm relying on too many sources of knowledge, may experience decreasing returns from collaborations.

Putting these elements together, the technology and innovation management literature argues that there is a tradeoff between internal and external search strategies in terms of what a firm can gain through search and what it can effectively take advantage of (Laursen, 2012), and this may also be applied to entrepreneurial firms. Search that is more explorative likely relies upon external sources of knowledge and is non-local, and hence it involves conscious steps to move beyond or away from current routines and knowledge, and into domains that are new to the firm (Katila and Ahuja, 2002; Laursen, 2012). In contrast, search that is more exploitative in terms of commercialization more likely relies upon utilizing in-house sources of knowledge, and hence that which lies within the current knowledge base as embodied within the firm (Helfat, 1994; Fleming and Sorenson, 2004). In understanding why firms search by utilizing external sources of knowledge for innovations, the literature has also pointed out that the benefits to the firm of searching can be different under different conditions. The notion of over-search is exemplified in Katila and Ahuya's (2002) work on exploitative search processes within the firm in terms of depth and scope, and later in Laursen and Salter's (2006) work with explorative search processes. While search activities in foreign spaces are deemed to be initially healthy, over-searching conveys the notion that there is a limit to what a firm can absorb and reuse for gains. Up until a certain point, additional search activities are beneficial, but after this they decline marginally in their benefit to the firm. The idea of over-searching emphasizes that since search strategies are influenced by past managerial behavior and future expectations, the outcome of carrying out too many search processes could have diminishing returns for the firm and even lead to a detrimental outcome (Laursen and Salter, 2006, p. 136). The same rationale follows for the depth construct: That too deep reliance on partners could lead to decreasing marginal benefits. These constructs were operationalized by Laursen and Salter (2006) using a Community Innovation Survey (CIS) sample of UK manufacturing firms. They found that breadth and depth are curvi-linearly related to innovative performance.

Additionally, openness may associate with the relative degree of novelty of an innovation for entrepreneurial firms. In the literature, innovations are classified according to their degree of radicalness as compared to the current standard of technology or as compared to the existing products (Freeman and Soete, 1997; Fagerberg, 2006). Continuous improvements are referred to as incremental innovations, while the introduction of something truly novel, new, or revolutionary in economic terms is called radical innovation (McKelvey, 1996). This distinction provides a more nuanced understanding of how and when external sources of knowledge impact innovative performance. Laursen and Salter (2006) have shown that for manufacturing firms, increased depth is more beneficial than increased breadth for companies' radical innovation turnover rates, due to patterns of narrowed source reliance in product life-cycle innovation (cf. von Hippel, 1988). Because of this association of radical innovation with a discontinuous reliance on knowledge sources, and the argument that some forms of knowledge may become obsolete as an innovative process narrows in focus (Utterback and Abernathy, 1975); increased breadth by a firm may produce smaller gains than depth. Incremental innovations often become more important after a dominant design has emerged late in the product life cycle (ibid.)[15], thus, Laursen and Salter (2006) argued that breadth becomes more relevant with product maturation, market expansion, and knowledge of the design becoming more widespread. This has led to the assertion that breadth of search is more strongly associated with incremental innovations for manufacturing firms.

Increased breadth of sources of knowledge could help the SME to increase the novelty of its innovation. Love et al. (2014) find that organizations that build off their prior innovation linkages in order to more effectively utilize their present day search breadth will tend to "experience higher

[15]This premise is also related in the theory of industry life cycles put forth by Gort and Klepper (1982) and Klepper (1996), which specifies how an industry's innovative character goes through a series of stages: First many innovations are commercialized; Second, many producers compete and drive down real prices; Third, there is a 'shakeout' process in which the number of producers declines markedly, with a dominant innovation emerging and more focus being laid towards process (incremental) rather than product (radical) innovation (Agarwal and Braguinsky, 2015).

innovation returns", and this may be characteristic of certain SMEs. Lee et al. (2010) argue that SMEs may be quite capable when it comes to invention, but lack the resources to properly innovate and commercialize invention. This notion of innovations requiring complementary assets (Teece, 1986) is not new, but it could be nonetheless relevant for explaining radicalness of innovations in small and micro firms. *Viz.*, this could mean that small and micro firms, while radical in their inventions, require extensive collaboration to properly reach innovation.

The extent to which new ventures rely upon external sources of knowledge can be seen as a form of openness. Following the definition in (Laursen and Salter, 2006: 134), search breadth is here defined as the "number of external sources or search channels that firms rely upon", where search depth is the "extent to which firms draw deeply from the different external sources or search channels". While their focus was on manufacturing firms, this dissertation focuses upon these concepts for entrepreneurial ventures. All in all, it seems highly likely that a broad assortment of external sources of knowledge, and thereby high degrees of external knowledge intensity, should positively contribute to an entrepreneurial firm's innovative performance. Therefore, I will propose the same relationships hold as in the original Laursen and Salter (2006) paper, even though they focused upon larger manufacturing firms, but for the different reasons given above.

When it comes to understanding KIE ventures in terms of the impacts of external sources of knowledge, the KIE-specific literature specifically proposes that they extensively use networks and external sources of knowledge to overcome resource limitations (Malerba et al., 2015). Still, due to the resource constraints specified above for SMEs in general, we would expect that an excessive breadth in sources of knowledge as well as excessive depth of collaboration with these sources should eventually result in a negative relationship. *In other words, if the entrepreneurial venture has a large number of deep (or relatively highly important) external sources of knowledge, this search depth will eventually lead to diminishing returns in terms of manufacturing innovations, producing an inverted U-shaped relationships as the number of partners increases.*

There is also a good deal of literature regarding how openness affects service-based outcomes, as they can sometimes manifest different circumstances regarding connectivity with different actors within an innovation space or system. It is important to bring this to light, in part due to the fact that many sectors that are proposed to encapsulate knowledge intensive activity lie within the definitional boundaries of business services[16]. I will now argue that the same relationships outlined above will hold for service-based entrepreneurial firms, but for slightly different reasons: Many knowledge intensive firms are more heavily involved in service sectors than in those traditionally categorized as manufacturing-based. Although studies of innovations have traditionally studied large firms in manufacturing industries, more recent work has also focused upon service innovations, and service industries, and the importance of collaborations (Tether, 2014). Results have shown that there are often vast differences between the two categories of innovation, including how they address users, customers, networking, and the innovation process itself. According to Tether (2014: 604-605,), four characteristics of classical services, which distinguish them from goods and manufacturing industries are: 1) Intangibility; 2) Inseparability between what is provided and who is providing it; 3) Temporal and perishable, that is, they exist in time; and 4) Heterogeneous depending upon the context of service deliverable, rather than standardized.

Hipp and Grupp (2005) were among those to study the utilization of external knowledge sources concerning service-oriented firms, using samples of firms in Germany. They categorized their results based on Pavitt's (1984) taxonomy, and emphasize the importance of knowledge intensive business services (KIBS) as a group which supplies a large number of economic actors with new knowledge. The OECD (2012) has also relied on this concept of KIBS to describe the drivers of service innovations in many knowledge-intensive organizations. These organizations tend towards generation of *ad hoc* and highly customized solutions to problems, with a high reliance on professional skills (Sundbo and Gallouj, 2000). Information regarding these innovations are likely to "flow through professional networks and associations, or other communities of practice" (Miles, 2012: 11). External collaborations can thus generally be seen as beneficial for service-based innovation, especially for knowledge intensive business services. Den Hertog et al. (2010: 494) argue that service innovations are more and more the result of a realization of opportunities to create and appropriate value within a wide network of actors, including providers, value chain partners, and others, and that new and improved services are often generated within large communities through linked platforms and business relationships.

Hence, when developing service innovations, firms activities involve processing and evaluating user needs as signaled to them by user groups, as well as recognizing and sorting between technological options (den

[16]The conceptualization of service innovation employed here assumes that "existing instruments will work effectively to describe the service economy" i.e. an assimilative approach (Miles, 2012: 11).

Hertog et al., 2010). Technological options provide opportunities for new paths of innovation, and the firm should remain open to external (as well as internal) sources of information, because that knowledge is crucial to translate potential technological options into innovations, including new service innovations (Teece, 2007; Wang and Ahmed, 2007; den Hertog et al., 2010). Bruni and Verona (2009: 107) similarly attribute these abilities to a firm's "dynamic marketing capabilities". A key variable in a firm's ability to generate innovations in services is thought to be user interaction as (Kindström et al., 2013), given the importance of co-creating services.

In terms of the degree of innovativeness, entrepreneurial firms may produce service innovations that are technological in nature or that are reliant on new business models which are based on a radical innovation. After a while, that firm may begin engaging in less radical forms and instead focus upon incremental and process improvements (Sundbo and Gallouj 2000).

This review holds entrepreneurial service firms in a somewhat different light than their manufacturing-based counterparts, however, I will argue that *service-focused firms meet with the same limitations in terms of liabilities of size, resources, networks, and experience: And that while a certain degree of openness to outside sources of knowledge can be beneficial, too much may have an adverse effect on their innovative output of services.* Additionally, because service innovations may rely extensively upon external sources of knowledge due to the inherent characteristics of services, the same direction of effects is predicted for both manufacturing and service firms regarding their external knowledge reliance, that is, the degree to which a firm attributes importance to external knowledge sources for their innovative processes and new opportunity identification.

I will argue that the more open an entrepreneurial venture becomes, in terms of both breadth and depth, the more likely they are to realize the process of invention through radical innovations. Keupp and Gassmann (2013) find strong support for the hypothesis that knowledge constraints on a firm spur radical innovations (defined as new to the firm innovations), arguing that resource scarcity of firms can trigger an increased propensity towards explorative activities and recombining of resources both internal and external to the firm in order to innovate. Hence, utilizing external sources of knowledge can be seen as being a reactive attempt of SMEs to overcome their own resource constraints in order to seek out new combinations of their own resources as well as that of others. *Therefore, we expect the greater breadth and depth of external sources of knowledge that a KIE venture has, the more radical innovations that the firm will have, when we analyze manufacturing and service innovations together. But, too much reliance on external knowledge* *sources may lead to a lessened capability to manage and implement more radical innovations.*

I use the preceding logic to formulate the following hypotheses regarding breadth and depth of search on innovative performance in both manufacturing and services firms using the 3 operationalizations detailed above, manufacturing innovation, service innovation, and degree of radicalness of innovations. Breadth and depth should be curvi-linearly related to innovative performance:

> **Hypothesis 1.1a:** *Search breadth is positively related to innovative performance in KIE ventures.*
>
> **Hypothesis 1.1b:** *At higher levels of search breadth, the positive relationship will experience diminishing marginal returns, resulting in a marginally declining positive association.*
>
> **Hypothesis 1.2a:** *Search depth is curvi-linearly related to innovative performance in KIE ventures.*
>
> **Hypothesis 1.2b:** *At higher levels of search depth, the positive relationship will experience diminishing marginal returns, resulting in a marginally declining positive association.*
>
> **Hypothesis 1.3:** *Search breadth should have a stronger effect than search depth on innovative performance*

Additionally, though admittedly based on less empirical evidence, I argue that overall measures of external knowledge intensity in the firm's reliance on different sources of knowledge will, as argued above, also be observed to be inverse-curvilinearly related to the innovative performance of entrepreneurial ventures. Here *reliance* refers to a reliance on external knowledge sources for identification of new business opportunities and innovation possibilities

The literature has presented concrete links between R&D Intensity, innovativeness and reliance on specialist knowledge providers (Laursen and Salter, 2004; Tether and Tajar, 2008). For firms with higher external knowledge intensity towards these types of sources, it is plausible that a similar link will be found. Additionally, given the strong links between more knowledge intensive service firms, high tech manufacturers, and low- and mid-tech knowledge appliers, and their customers, suppliers and clients when concerning the development and commercialization of their innovative products and services, an association between these firms' reliance on intra-industry knowledge sources and overall innovative performance is expected. Moreover, literature investigating co-opetition, or, the simultaneity of cooperation and competition between (small) firms, has shown that new technologies and products are often the result of *co-opetition*, that is, the increased technological diversity and new combinations of complementary resources between rival firms (Harbison and Pekar, 1998; Gnyawali and Park, 2009; Quintana-García and Benavides-Velasco, 2004). Additionally, competitors often collaborate in order to appropriate benefits from achieving scale economies, improved risk mitigation, and heightened resource-leveraging capacity (Morris, Kocak, and Özer, 2007).

Therefore, several more working hypotheses[17] (coded as WH throughout the text) may be generated as a result of this theoretical overview of search activities and their implications for innovative performance of new ventures and small firms, and used to investigate these effects as present in knowledge intensive entrepreneurial firms. These have less theoretical and empirical support in the literature, but can be traced to some related research. Based on the above discussion, I will investigate the following additional working hypothesis:

> **Working hypothesis 1.4:** *Reliance on external sources of knowledge is positively related to innovative performance in entrepreneurial ventures.*

> **Working hypothesis 1.5:** *Too much reliance on external knowledge sources will result in diminished marginal benefits in terms of innovative performance.*

As a result, I argue that external knowledge intensity in new ventures is positively associated with innovative performance. The next relationship to be investigated is more based on working hypotheses, that is, the effect of internal knowledge intensity as I will characterized it here on external knowledge intensity, which is based on the idea of firm openness from the previous section. I will attempt to establish a link by reviewing relevant theory, and will propose several working hypotheses to guide the analysis to come.

[17]In the pages to follow, I generate a number of hypotheses with varying empirical grounding. This is due to the fact that many relationships have not been previously investigated in the same setting, context, or relational directions as I am doing in this dissertation. In the interest of clarity of purpose, I will summarize all of them again in conjunction with each empirical sections with which they are specifically related.

3.4.2 Pre-history, founders, and organizational conditions and their effect on a firm's external knowledge reliance

As touched upon in Chapter 1, the concept of internal knowledge intensity is defined as *the knowledge intensity that is largely inherent in a firm when it comes into being, rooted in different types of human capital investments and outcomes, as well as other knowledge-based factors have driven the firm to formation.* In this section I will propose a series of working hypotheses in order to explore the relationship between firm pre-history rooted in the above concepts, and how these associate with the external knowledge intensity of, when possible, entrepreneurial ventures. Otherwise, literature on SMEs or firms in general will be reviewed due to lack of theoretical coverage regarding the former. While on average under-researched, and consequently not yielding many direct studies or propositions regarding the link between concepts lying near internal and external knowledge intensity as I define them, the literature has actually often indirectly implied such a relationship. Huber (1991:91) argued already in 1991 that *"[w]hat an organization knows at its birth will determine what it searches for, what it experiences, and how it interprets what it encounters",* an argument that may indirectly be found in much of entrepreneurship literature to date. This section will review this literature in the attempt to expand on the following research objective:

> **RO2:** *Explore the association between internal knowledge intensity and external knowledge intensity in the entrepreneurial firm.*

In order to generate some working hypotheses that address this objective in more detail, I will first motivate why each type of external knowledge source which we have already identified might be of interest for study in terms of how their external knowledge source reliance is affected by different types of internal knowledge intensity, which I will begin to discuss here, and will elaborate on further in the following section regarding business performance.

3.4.2.1 External knowledge in the form of specialized knowledge providers

That knowledge flows through relationships is one of the core assumptions of many branches of organization and business theory, including Network Theory (Schrader, 1991; Powell and Grodal, 2005; McKelvey and Rake, 2015), Innovation Systems (Edquist and McKelvey, 2000; Nelson, 1993; Shariff, 2006) , and Open Innovation approaches (see Dahlander and Gann, 2010). One highly influential development in understanding how external sources of knowledge interact with both the innovation process and firm level characteristics happened through the combined efforts of the researchers involved in the Yale Survey (Levin et al., 1987; Klevorick et al., 1995). Klevorick et al. (1995) explored inter-industry variation of technological opportunities through the examination of how different non-industry knowledge sources interacted with industry knowledge sources, and that firms across industries draw on a wide and varied range of knowledge inputs (Salter and McKelvey, 2016). There has moreover been some work conducted on the connectivity between founders and employees of new ventures' diversity, experience and education and how they affect search choices and search strategy. Much of this is based on the idea that an organization that is better able to absorb, assimilate, and exploit new knowledge given that it has the adequate prior knowledge that is sufficiently related (Cohen and Levinthal, 1990).

Since this early beginning, many have sought to better understand how firm-level factors influence (both directly and indirectly) external knowledge sourcing and reliance. Laursen and Salter (2004) emphasized the effects of various firm level and behavioral factors that could influence the extent of the firm's interaction with sources of knowledge such as universities. They found that the more a firm interacts on an overall level with various knowledge sources, the more important university interaction in this capacity may become. Building on this argument, Tether and Tajar (2008) proposed that the role of universities in this context was just the tip of the iceberg in a deeper sea of specialist knowledge providers, including a wider variety of non-business related sources of knowledge. They argued for the positive effect of social capital at the firm level (including surrounding networks) on the value placed on knowledge stemming from these specialist knowledge providers for both high and medium tech manufacturing entities, as well as knowledge intensive business service providers. I am in agreement with this, as will be seen in the hypotheses to follow, but focus more on human capital and pre-history ambitions and goals of the venture and its originators. In social capital they included variables such as employment of graduates, the firms' own R&D commitments, and innovation financing. Knowledge intensive entrepreneurial firms, I argue, also benefit from many forms of human capital, in terms of an increased reliance on knowledge from Tether and Tajar's specialist knowledge providers. Namely, the internal knowledge indicators including composition of the founding team, the background of the entrepreneurs and employees, as well as the organizational origins of the venture, should have some distinct effects on knowledge reliance of different groupings of SKP sources.

3.4.2.2 Intra-industry knowledge

Classen et al. (2012) point out that small and medium sized enterprises are often highly reliant on external information and resources, which more often than not come from the networks of their founders and employees. These firms, then, gladly and often out of necessity link with customers and suppliers in terms of reliance on them as a source of crucial business knowledge (Fann and Schmeltzer, 1989). Firms whose founders' previous networks and experience within a given industry had a strong role in shaping the firm itself, have then, a higher likelihood of being strongly influenced by industry knowledge sources that are highly related to those valued by its founders in previous careers and work experience. In terms of qualities among the new ventures that should positively impact knowledge reliance on types of sources such as clients, customers, and competitors, it may be the case that more internally knowledge intensive firms, that is, firms that have higher diversity among founders in terms of functional background, education and experience, will rely less on these 'normal' channels of search, and more on non-industry sources. This is because a wider functional diversity puts them in a better position to search outside their own industry in order to gather new valuable business knowledge. Additionally, I argue that more educated and experienced founders may be more exposed to less industry-tied activities and actors, due to the fact that over time they may acquire a broader perspective, through formal education or entrepreneurial experience, on what constitutes valid areas for expanding their business.

New firms in an industry that values client and customer input in new product development, or has close relationships with their competitors could benefit greatly by being aware of new emergent opportunities resulting from technical change and shifting market needs. Awareness of new developments could lead to heightened network collaboration to pursue the necessary resources and capabilities needed to exploit the opportunity, as often small new firms will not have all the required resources at their disposal, and must reach out to other firms within their chain of value added activities or competitors. Also, potential founders that are able to recognize new technical and regulatory trends may require additional resources from intra-industry sources, driving a heightened reliance on these sources for developing the opportunities recognized.

3.4.2.3 A summary of proposed effects of internal- on external knowledge intensity

Boeker (1987) examined how strategy origination in organizations occurs as a consequence of prior entrepreneur experience (and place and time of founding the firm). He found links between adopted strategies of the organization and prior 'functional experience' on the part of the entrepreneur as well as age and level of education. As others, Boeker also relies upon the *imprinting* language developed by Stinchcombe (1965), and moreover emphasizes that organizational routines and early strategic decision making are a product of both the entrepreneur and the external environment, and founding a new firm enables entrepreneurs to infuse it with their own assumptions, views, and means to carry out strategy. Relating this strategy origination to KIE firms (or to a concept at least near to it), previous studies have found that for young new technology based firms, enhanced human and social capital can lead to strengthened network ties, increased access to larger pools of external information, the ability to recognize information as inherently strategic, and the facilitation of knowledge transfer where dynamism is valued over stasis or imitation (Yli-Renko et al., 2001; Thorpe et al., 2005).

This means that the pre-history of a firm and its founder(s) will to some extent guide its choices of what types of external knowledge they may rely on and how heavily, for their own strategic intent. This idea fuels the line of inquiry to follow.

The discussion can also be linked to routines and behavioral theory. Beckman (2006) argues that since founding team's overall composition, and it members' prior affiliations, have an impact on the establishing of new firm behavior and routines, exploitation (local search) activities may occur more commonly when members of a prior organization are collaborating in founding a new venture, while founders with disparate backgrounds and employment history encourages a more explorative (non-local search) environment. Strategic initiatives of the firm to cooperate with diverse external actors in innovative environments, that is, search activities, has also been to some degree linked with behavioral characteristics of organizational leaders and decision makers (Classen et al., 2012). Further, diversifying teams may be better able to pick up on new search directions: Kim and Kogut (1996) postulated that a firm's diversification is linked to the development of the technological trajectory pursued by the firm. They argue that push factors like current capabilities and pull factors like environmental signaling shape the search directions of firms; again, implying a relationship between inceptive capabilities and search activities.

Moreover, there is a transitivity of strategy that flows between organizations via the entrepreneur. Actions considered by founders in newly arisen business situations often have been manifested earlier on through past experience, and are carried from venture to venture (Boeker, 1997; Beckman, 2006). Firms possessing the virtues of a combination of founders, and thus potentially multiple past strategies and types of experience to draw from, have growth advantages, suggesting team composition's impact on both explorative and exploitative behavior[18]. Furthermore, Basu et al. (2015) argues that bridging multiple areas of knowledge within an organizational decision making unit can be the result of greater knowledge diversity within this unit, while related area opportunities may also be exploited through cohesion of knowledge and expertise in that specific domain of knowledge. With venture growth, additional employees may modify and extend the technological trajectory of the venture as it exists at that time (ibid.). Thus, different founding team compositions have an effect on types of search activities, and further, which types of external sources are most highly valued by the firm.

Additionally, types of opportunities recognized, characteristics, past affiliations and working history of the team will ultimately color the development of new ideas and new identified opportunities and business areas (Beckman, 2006). Thus it could be proposed that teams with broader diversity will access and potentially use broader sources of information. Also, new ventures whose formation was highly contingent upon certain uncovered opportunities of a technical, market failure, or regulatory nature, should exhibit a heightened reliance on specialist knowledge providers, especially given that knowledge intensive activities of entrepreneurial firms often draw on these types of sources.

In addition to diversity of experience, strategic orientation, and education brought about by a heterogeneous blend of founders and employees in a newly formed venture, it is also worth considering how more traditional human capital inputs like education might relate to this discussion. Often, since new ventures are so inherently small in size, their employees levels of education and experience can be also argued to play a role in the "team effect" of diversity of in overall levels of education. Dahlin and

[18]As we have seen, exploration and exploitation are easily superimposed into a search schema, with more radical, boundary spanning search often being equated with exploration and competence harnessing, and otherwise predominantly local search, being equated with exploitation.

colleagues have argued that educational diversity enhances the use of diverse information in firms (Dahlin et al., 2005), which in turn affects and shapes their decision making practices. It has been argued that founders, as well as employees, with higher levels of education have greater cognitive diversity (Hitt and Tyler, 1991; Wally and Baum, 1994: both in Classen, et al., 2012) and will have a better ability to absorb relevant external knowledge (Barker and Mueller, 2002, in Classen et al., 2012). As the exploitation of an opportunity hinges on, some say, its prior discovery, new venture employees that are better educated may be more apt to pick up on new business and innovation opportunities (Agarwal et al., 2004), and human capital within the workforce can and often does contribute to performance on the organizational level (Smith et al., 2005; Siepal et al, 2016). This has been argued in recent studies, which have suggested that levels of human capital of employees at early stages in the entrepreneurial venture's lifespan can affect firm performance in terms of growth and survival in the long run (Siepal et al., 2016). Additionally, founders with more satisfactory levels of education should be more likely to span search boundaries when they utilize external knowledge (Dollinger, 1984 in Classen et al., 2012). I therefore propose that:

> **Working hypothesis 2.1:** *For entrepreneurial firms: Higher levels of stocks of human capital in the form of education will positively influence the reliance on external knowledge stemming from all types of external knowledge source.*

Beyond solely focusing on the agglomerated effects of the degree of *functional diversification* or *education* of founders and employees, and how this might affect reliance on and reliance on external sources of knowledge, the specific nature of the background or previous career or occupation of founders may also have a strong influence. Some claim that more 'academic' founders have the potential to amplify the amount of distinctive resources in a new venture (Stuart and Ding, 2005 in Basu et al., 2015). Relevant prior industry employment and experience also should play a role: Agarwal et al. (2004) argue that prior employment affiliations impact not only new venture creation, but also the trajectory in terms of product focus and strategy employed by these newborn ventures. Prior same industry experience of the founder also facilitates the ventures' ability to satisfy unique customer demands, harness tacit knowledge about the industry, and utilize ties to suppliers, distributors and other network actors that have existed previously (Campbell, 1992; Delmar and Shane, 2006)[19]. So, I propose that:

[19] Other more specific types of firm activities create a different reliance scheme. For instance, regarding R&D partnering of firms, different partners fill different functions (Teece, 1980;

Working hypothesis 2.2a: *Founders' previous work experience at a university or research institute will be positively associated with the reliance on knowledge stemming from non-industry sources of knowledge such as specialist knowledge providers.*

Working hypothesis 2.2b: *Founders' previous work experience in the same industry or having entrepreneurial experience will be negatively associated with the reliance on knowledge stemming from these types of knowledge sources.*

Working hypothesis 2.3a: *Founders' previous work experience at a university/research institute will be negatively associated with the reliance on intra-industry sources of knowledge.*

Working hypothesis 2.3b: *Founders' previous work experience in the same industry, or previous entrepreneurial experience, will positively associated with the reliance on intra-industry sources of knowledge.*

Research on founding team effects often involves analyzing a team's functional heterogeneity, or the diversity or lack of diversity of the backgrounds of the founders (Hambrick and Mason, 1984). Commonly, when looking at functional heterogeneity, entropy-based[20] measures are often relied upon to measure diversity among a group of individuals. Often, researchers adapt variations on the Blau index (Blau, 1977), or the Shannon measure of entropy (Shannon and Weaver, 1948). Recently, Cantner, Goethner and Stuetzer (2010) went beyond traditional measures of team functional heterogeneity to craft a two dimensional model, in contrast with more accepted models of heterogeneity capturing mainly net or average effects of a team's composition (i.e. Shannon and Weaver, 1948; Attneave, 1959; Blau, 1977). Their two dimensional measure is proposed to capture both the positive and negative forces at work within the composition; *knowledge scope*, beneficial effects stemming from cognitive resource breadth, and *knowledge disparity*, non-beneficial effects stemming mainly from social categorization due to team role, and other such factors (Cantner et al., 2010).

There may also be implications on innovative direction of a venture based

Classen et al., 2012). These may range from helping the firm cope with uncertainty avoidance or reduction by way of user interaction (von Hippel, 1988), or, managing quality or cost efficiencies through supplier relations (Hagedoorn, 1993). This may be done in the interest of building synergistic relationships with the competition (Das and Teng, 2000). Radical innovation might also be facilitated through public funding and building relationships with universities and research units (Tether, 2002).

[20] This concept has its basis in Shannon and Weaver's (1948) work about information systems, where the system entropy is the expected value of an average amount of information in a system.

on these founding team characteristics, which would in turn affect what types of knowledge sources become highly valued for more knowledge intensive new ventures. According to Shane (2000) individual team members stemming from different expertise areas will differ in their strategic direction of exploiting a singular technological innovation, and Cliff et al. (2006) argue that individuals with weaker core area experience and stronger peripheral experience have higher likelihood to innovate, and that the degree of novelty is dependent upon these experiential domains of founders. It can be argued then that wider experienced founding teams, all else equal, will innovate more broadly than those lacking such diversity. So, firms with a broad functional heterogeneity positively impact the reliance on knowledge sources external to the firm. Naturally then, firms with more limited functional heterogeneity, and higher founding team knowledge disparity, that is, non-beneficial overlap of competences, or lacking key backgrounds within the team, might have the opposite effect on external knowledge source reliance. With all this in mind, I put forth the following working hypothesis:

> **Working hypothesis 2.4a:** *In terms of functional heterogeneity of the founding team, the level of knowledge scope should positively associated with the level of reliance on external knowledge from all categories.*
>
> **Working hypothesis 2.4b** *The level of knowledge disparity should negatively associate with the level of reliance on external knowledge from all categories.*

Previous startup experience has also been raised as a contributing factor to what types of external search patterns may be undertaken by new ventures: Delmar and Shane (2006) argued that startup experience was a conduit for learning about opportunity identification, evaluation, and exploitation for founders, and others have confirmed this link with an impact on performance (Jones-Evans, 1996; Shane and Stuart, 2002). This experience also links the entrepreneurial team to networks of employees, suppliers, investors, customers, and other channels of vital information for the growth and survival of the venture. Founding experience thus contributes to reduced liability of newness (Shane and Khurana, 2003). This rationale leads me to formulate the next propositional hypothesis:

> **Working hypothesis 2.5:** *Stemming from a previous organization, i.e. being a corporate spinoff, will positively associate with the reliance on knowledge stemming from all types of external knowledge sources.*

Opportunity recognition/creation also has a role to play in the reliance on external sources of knowledge. Firms who are active within networks and have previous branch experience may be predisposed to insider knowledge about shifting directions and trends in a given industry, which may give rise to better disposition towards exploiting an opportunity and solidifying a new venture. The sensitivity to identify or create new opportunities based on market conditions signals often a competent entrepreneurial force behind a venture, be it a sole founder or a team. Given that the opportunity identified proves to be one of validity and commercial potential, the ability to recognize and exploit these types of opportunities, be they technical developments in surrounding fields or industries, government initiatives, or newly emerging market needs, should be quite strongly related to the level of reliance on external knowledge.

> **Working hypothesis 2.6:** *The extent to which a firm's formation was based on novel opportunities should be positively associated with the reliance on knowledge stemming from all types of external knowledge sources.*

3.4.3 Internal knowledge intensity as affecting performance in new firms

In this section, I make use of previous research on pre-entry endowments and organizational origins, founder characteristics, resources and capabilities, as well as more established human capital indicators like education, as proxies of the latent internal knowledge intensity concept as present in entrepreneurial firms, and attempt to link it with business performance. While there is limited research on KIE in general, researchers have begun to use variations of resource- and knowledge-based firm theory to characterize and explain phenomena related to entrepreneurial startups, and to some extent, to the KIE concept directly. [21] Numerous studies have spoken generally about entrepreneurial firms' origins in the form of pre-entry endowments as enhancing long-run performance and growth (Brüderl et al., 1992; Gimeno et al., 1997; Klepper, 2002; Shane, 2003; Parker, 2004). Helfat and Lieberman (2002) utilized an amalgamation of resource-based perspectives in order to

[21] McKelvey and Lassen (2013), for instance include different types of resource-based indicators in their categorization of 'accessing resources and ideas' in their KIE-based studies.

analyze the development of resources and capabilities at the time when a new firm is born and enters a market. They find that pre-entry firm resources and capabilities greatly influence the likelihood of market entry, and subsequent survival and prosperity, depending largely on the match between said resources and capabilities of a firm and those required in an industry. Gaps in resources were also found to affect entry modes by new firms. Their research constituted important steps taken towards building an understanding of how pre-entry resources and capabilities, including founder characteristics, affect the performance outcomes of entrepreneurial firms. This perspective has been largely influential here, in viewing resources and capabilities as inputs in the knowledge intensive entrepreneurial model. Other studies have provided similar evidence regarding the crucial role founders play in shaping the trajectory of their organization, in both established and newly formed ventures (cf. Klepper, 2002; Shane and Stuart, 2002; Delmar and Shane, 2004; Dencker et al., 2009; Unger et al., 2011; Baptista et al., 2014; Dencker and Gruber, 2014). In this section the following research objective is proposed:

> **RO3:** *Explore the association between internal knowledge intensity and business performance of the entrepreneurial firm.*

3.4.3.1 Founder and firm human capital

Founder characteristics are instrumental in determining the pre-entry resource endowments of a firm and how it performs after it has started business activities. There is a large literature reviewing the conceptualization of human capital of founders, where human capital represents education, experience, knowledge, and skills, which in turn increases the founder's capability to discover and to exploit business opportunities (Unger et al., 2011). Human capital theory was originally devised as a tool to analyze employee income distribution and efficacy in the workplace (Becker, 1964), but has since become largely influential in management and entrepreneurship research (Unger et al., 2011). Moreover, human capital has been long argued to play a crucial role in the potential success of a new venture (Pfeffer, 1994; Florin et al., 2003), as well as in the decision making process leading up to venture formation (Evans and Jovanovic, 1989; Campbell, 1995) and ultimately market exit (Gimeno et al., 1997). In economics it is common that the concept is divided between two related measures: human capital investments, or, education and work experience; and human capital outcomes, or the knowledge and skills that are the result of a combination of the investments (Becker, 1964). This study operationalizes both of these types of measures. Most commonly perhaps, human capital has been measured by quality or years of schooling (Hitt et al., 2001, Kor and Leblebici, 2005), but later research has issued recommendations against *solely* using such simplified proxies to measure it (Ployhart and Moliterno, 2011; Unger et al., 2011)[22].With this in mind I have relied on several measures in approximating human capital, not just quality of schooling.

Relating more directly to founder influence on venture creation, it has been argued that founder human capital increases venture success probability through increasing capabilities, entrepreneurial alertness, and exploitation of opportunities (Shane, 2000; Shane and Venkatraman, 2000); as well as attracting other sources of capital investment (Brush et al., 2001), and creating a base level of learning capability that can be used in accumulating new related knowledge and skills (Unger, et al., 2011). Human capital has also been used as a firm level indicator in some studies, emphasizing the firm-specific nature of the resource (Ployhart and Moliterno, 2011). In this sense, it is dealt with as an aggregated form of knowledge, skill or organizational experience. Since knowledge, skills and human resources in general are contributing to a firm's pool of resources and capabilities which in turn help shape its competitiveness (Coff, 2002), human capital aggregated to the firm level can be useful in determining their performance.

Research on individual human capital as a component of (internal) knowledge intensity has not been plentiful thus far, especially when tied directly to knowledge intensive entrepreneurship. Most comes as a result of the KEINS/AEGIS legacy. Both McKelvey and Lassen (2013) as well as Malerba and McKelvey (2015) include aspects of individual and founder-level human capital in their analyses of KIE, though more in-depth work is needed to examine the underlying relationships[23]. By analyzing related topics we can learn a bit about what might also be applicable in KIE-related scenarios. For instance, concerning potential relevance to knowledge intensive entrepreneurship and its presence in diverse sectors, the human capital of founders and founding teams has recently shown as being influential for the success of a new venture in both high and low technology oriented industries, and more effective for younger as opposed to older firms; also, outputs of human capital

[22] Ployhart and Moliterno (2011, p. 129) do not however condemn this approach entirely, they state that:"*Although we realize that practical constraints may frequently necessitate the need for proxy measures and that some disciplines (such as economics) may have different views about whether proxy measures are problematic, we believe that, whenever possible, it is preferable to use measures of [knowledge, skills, abilities and other characteristics] and then aggregate them as appropriate for the emergence model theorized."*

[23] Caloghirou et al. (2015) used cluster analysis to categorize KIE firms based on a number of criteria. Their analysis yielded different categories and levels of KIE firms, in which *world class KIE* was made up of firms with high levels of human capital in founders and employees

(knowledge and skills) have a larger effect on success (in terms of both operational/innovative and financial performance) than investments (mainly education) in human capital (Unger et al., 2011). Additionally, human capital appears to have a higher moderating effect on new venture performance, and Unger et al. (2011) have called for further research with a moderator approach to human capital in entrepreneurial startup processes. Another way to attribute human capital to the entrepreneurial firm involves discussing attributes of founders, founding teams, as well as employees.

3.4.3.2 Education, experience, opportunities, and performance outcomes

Table 3.2 consists of an overview of these and other resource-based studies of the pre-history of entrepreneurial firms and how this affects performance. Often traits of founders and founding teams regarding previous experience, last source of employment, and education are used to represent human capital origins of firms; additionally, firm level constructs such as corporate vs academic origin (and other taxonomies), types of chiefly employed resources, extent of planning present in company strategy, types of aggregated firm level factors driving formation (such as opportunity identification, technological skills of founders, or networks and experience from previous careers), and other indicators are used to capture pre-history qualities of entrepreneurial ventures. It can be seen that most are concerned with more than economic output, but that economic output does constitute an important performance component. Growth in employees, or size of the firm each year, is also a valid and commonly used indicator. When financial indicators are used, they are often transformed to conform to normality principles of regression analysis. Venture survival conditions are particularly common as dependent variables in analyses as well.

Table 3.2: Instrumental studies about the impact of pre-entry and pre-history factors of entrepreneurial firms on performance

Authors	**Input Concept**	**Input operationalization**	**Output concept**	**Output operationalization**
Almus and Nerlinger (1999)	NTBF firm	Different indicators capturing NTBF (corporate structure/origin, tech input, founder specific traits). Differing cohorts of firms (High, Mid, Low-tech)	Firm growth	Growth model of logarithmic size Change between periods.
Klepper (2002)	Time of entry Pre-entry experience	Year of entry R&D Productivity *Type of firm:* (Inexperienced firms, spinoffs, experienced entrepreneurs, experienced firms)	Survival	Hazard function of firm exit (moderated by firm age).
Helfat and Lieberman (2002)	Resources and capabilities of startups	*Type of entrant*: (Diversified, parent company venture, *de novo* entrant Entry Mode:(JV, acquisition, merger, franchise, Greenfield, etc. *Types of resources*: (Core vs. complementary, Specialized vs. generalized)	Firm growth	Growth model of logarithmic size Change between periods.
Colombo and Grilli (2005)	Founder human capital	Years of (industry specific) education Years of prior work experience Prior experience in management or entrepreneurship (Binary)	Post-entry growth	Log transformed employee growth
Dencker et al. (2009)	Business planning Pre-entry knowledge and experience	Extent of early stage business planning Importance of pre-entry experience to current occupation	Firm growth	Growth model of logarithmic size Change between periods.
Criaco et al (2013)	Founder human capital	Entrepreneurship human capital Industry-based human capital University-based human capital	Survival	University startup activity after 3 years

Taking into account these findings on education, opportunity recognition, and experience of founders, I hypothesize the following:

> **Hypothesis 3.1:** *Higher levels of human capital in the form of education levels within the firm are positively associated with business performance in entrepreneurial firms.*
>
> **Hypothesis 3.2:** *Higher levels of human capital in the form of entrepreneurial, industrial, and academic experience within the firm are positively associated with business performance in entrepreneurial firms.*

3.4.3.3 Founding team effects

There has been a substantial amount of research into the effect of homogeneity vs. heterogeneity present in management teams. (examples include Hambrick and Mason, 1984; Eisenhardt and Schoonhoven, 1990; Carroll and Hannan, 2000)

While singular traits, backgrounds, and motivations of founders are seen as influential in shaping new venture performance, the aggregated effects of founding teams also plays a focal role. Studies have shown that team-based startups will generally perform better than sole entrepreneurs who start a new venture (Cantner et al. 2010; Chandler et al., 2005; Chowdhury, 2005; Ucbasaran et al., 2003). The issue of assembly of a successful team composition permeates the literature on entrepreneurial founding teams, and most agree that diverse characteristics of these teams (including functional background, career experience, education, age, and financial position) tends to positively affect new venture performance both directly and through its effect on strategic decision making by the firm (Hambrick and Mason, 1984). While many statistically significant associations in both positive and negative directions have been recorded (Cantner et al., 2010), the entrepreneurial context is often not in focus, and the explicit use of the knowledge intensive entrepreneurial concept in conjunction with firm pre-history is again limited to the output of the AEGIS project itself (Malerba et al., 2015), which has thus far not gone in-depth into the issue, and has been mainly focused on network and cluster analysis, along with confirmatory factor analysis (Fontana et al., 2015; Caloghirou et al., 2015) in order to better categorize KIE firms.

The literature on founding team functional heterogeneity (see also section 3.4.2.3) (Cantner et al., 2010) has argued that knowledge scope and disparity have differing effects on new venture survival and growth: with knowledge scope being positively associated with venture growth and knowledge disparity being negatively associated with venture survival. They stress that ventures need an optimal trade-off in order to optimize the benefits, and minimize the drawbacks, of heterogeneity within founding teams. Additionally, having too broad a scope in terms of founding team knowledge could begin to diminish the positive gains from heterogeneity, in that the new founded venture cannot effectively allocate resources to all areas in which it has competences of value. Therefore in the empirical models to follow, I expect a curvilinear effect by knowledge scope on new venture business performance. The following hypothesis states the expectations of this relationship:

> **Hypothesis 3.3a:** *In terms of founding team functional heterogeneity: Knowledge scope is positively associated with business performance.*
>
> **Hypothesis 3.3b:** *At higher levels of knowledge scope, it will have a diminished marginal association with business performance for entrepreneurial firms.*
>
> **Hypothesis 3.4:** *In terms of founding team functional heterogeneity: Knowledge disparity is negatively associated with business performance for entrepreneurial firms.*

Recent findings have also pointed out that apart from the human capital inherent in founder's general knowledge, skills, and education, there is an important contribution to venture performance made by such factors as earlier experience in starting a business, or specific experience that can be gained working in a larger organization, or a previous startup (Bosma et al., 2004). We now turn to an overview of how different types of firm level organizational origins affect performance of new ventures.

3.4.3.4 Organizational origins of new firms

Generalized studies (that is, using rather broad and over-arching categories) of pre-history and pre-entry knowledge of firms and how this affects venture performance have become a common way of analyzing firm origins. For instance, Dencker and colleagues (2009) investigated the mechanisms underlying the influence of pre-entry knowledge of new firms on their chance of survival in the long term, and that pre-entry knowledge and experience increases survival rates through moderating the effects of early stage business planning. Different types of entrepreneurship have also been analyzed using similar methods and theory. Baptista et al. (2014) studies the effect of founder's backgrounds on new firm survival with special focus on necessity vs. opportunity based entrepreneurship, or, those that are unemployed and must become entrepreneurs compared to those that are acting on opportunity recognition in the sense of Shane (2000).

Regarding forms of KIE, some research in this area has been carried out, but is often split by differing foci in terms of what characteristics the KIE firm has. Studies focusing on categorizations such as new technology based firms and similar variants have been much more common, so they are worth mentioning here as there are some parallels to be drawn. Some such studies focused on linking specific skills generated prior to entry (education in specifically relevant fields and same industry experience, mainly) to improved post-entry performance outcomes (measured in rate of growth or changes in firm size over time) in NTBFs (Almus and Nerlinger, 1999; Colombo and Grilli, 2005). Additional research has tested hypotheses regarding the effect of founder human capital on growth of these ventures, finding that founder human capital has both direct effect, as well as indirect effects as a conduit for obtaining venture capital, on new venture growth (Colombo et al., 2010). Other KIE-related research has looked more at the organizational origins of entrepreneurial firms:

Specifically, how different types of spinoffs perform based on their pre-entry factors in different industries. Klepper (2001) reviewed much of the theory about spinoff nature, heritage and performance: Different factors like strategy concerning innovation direction or market focus vs. those of the parent firm, as well as the overall relationship with the parent firm and its characteristics (including strategy and innovativeness) were instrumental in determining how spinoffs structured their own activities and how they performed. Klepper based his findings around themes such as: a spinoff's high propensity to base initial strategy off of experiences of their parents, the high number of spinoffs stemming from innovative, more successful firms; and that the ability to draw on past experience of founders gives spinoffs a competitive advantage not shared by other types of new ventures. He argues that in general "it is not technologies they appropriate from their parents but the broad experiences of their founders that seem to determine their performance" (Klepper, 2001, p. 662). As far as what sort of lessons spinoffs learn from their parents, founders appear to draw in a more limited sense, largely based on specific training and areas of expertise within which they may be involved at the parent company.

Klepper (2002) later investigated the dawn of the American automobile industry, tracing the industrial origins of different categories of entrants to the market. He found that having founders previously employed in other companies that were active in related industries significantly impacted the performance (survival) of new ventures in the auto industry, with spinoffs generally outperforming *de novo* entrants. However, it seems these spinoff companies performed better only if the expertise of the 'parent firm' was built upon in the new venture, or if the parent was a leader in its industry. Klepper corroborates with Boeker (1988) in suggesting that *founder experience may become imprinted on organizations*, shaping their performance and routines for many years following inception. Also, the importance of industrial context in the roll of evaluating new venture performance based on pre-entry elements is emphasized (Klepper, 2002). While not specifically relating to knowledge-intensive entrepreneurship, Klepper's findings provide strong guidance for what should be expected in KIE-rich sectors. Some questions remain however. For instance, while it was been established that founder's traits can and will influence their ventures' performance in a variety of ways, the topic of how founders of spinoffs benefit from their employees and what they learn from them is a topic that at the time of Klepper's study was largely unexplored.

Additional research has focused on academic vs. non-academic spinoffs as the focal firms, looking into pre-entry factors such as configuration of knowledge and skills, technological capabilities, and organizational heritage. For example, Criaco et al. (2013) found in their sample of Catalonian university spinoff NTBFs, that having non-university human capital endowments negatively affected firm survival, while human capital endowments stemming from university environments or entrepreneurship education endowments had a positive effect. This might mean that KIE firms could be diversely effected by different types of human capital stemming from academic vs. non-academic sources. Additionally, Zahra et al. (2007) studied pre-entry knowledge conversion capabilities, a three stage concept including conceptualization, configuration and integration of new knowledge, and how these affected the post-entry performance of corporate vs. academic spinoffs. They found that corporate spinoffs benefit more from prior experience and connections, including resources transferred through founders and employees who previously worked at the parent organization, than academic spinoffs did. Clarysse et al. (2011) investigated the technological knowledge base possessed by corporate vs. academic spinoffs at inception and how this predicts firm growth, and found that the impact of narrowly focused technology is important for corporate spinoffs, while broadly focused technology can have benefits for academic spinoffs, given that they can appropriately transfer the technology from the university to the firm by some means. Wennberg, Wiklund and Wright (2011) looked at performance of university vs corporate spinoffs in knowledge intensive sectors (read: according to

OECD and Eurostat's classifications, where R&D Intensity is above the mean of the overall economic level) based on the subjects' last vocation before starting the venture, in Sweden. They found that commercial knowledge attained through industry experience was more valuable to entrepreneurs than academic knowledge, measuring firms' performance through employee growth and firm survival rates. Others have looked specifically at university startups, and how this particular group of firms benefits or does not benefit from certain types of human capital endowments.

While there may well be differing effects for different types of spinoffs, it seems that being some type of spinoff will have an effect on the performance of the entrepreneurial firm. Based on this review of the spinoff literature, both academic and non-academic, I propose the following:

> **Hypothesis 3.5:** *Having the organizational origin of being a corporate spinoff entrant is positively associated with business performance for entrepreneurial ventures*

Research on opportunity identification or creation has pointed towards entrepreneurial capability being somewhat gauged by an entrepreneur's ability to sense and exploit opportunities in business and markets more effectively than the average person. This type of logic can also be linked to knowledge intensive entrepreneurial firms[24]. Firms that have a high sensitivity to technological progress, new market needs, or new public finance and regulatory opportunities may be predisposed to instigate activity in certain business sectors or industries, and those having a high likelihood of containing knowledge intensive entrepreneurship as put forward in Chapter 2 could be fruitful ground for this type of opportunity creation. The question is whether or not this type of action on the part of the founder or founding team will enhance performance in the firm's future. Based on the ideas that more effective and successful entrepreneurs sense and seize opportunities based on their alertness, awareness, or entrepreneurial orientation, this can be deemed as a type of internal firm knowledge (intensity) that should positively impact growth and profitability if the opportunity proves, though there is an apparent risk component associated with an opportunity not reaching fruition. In sectors with a high potential for KIE, these types of developments often drive new market segments and growth, so it seems reasonable to assert that the effect should be positive on the individual new venture's business performance.

[24] Opportunities constitute one of the external factors of Malerba and McKelvey's (2015) KIE model, where they describe the dynamics of knowledge intensive entrepreneurship

> **Hypothesis 3.6:** *The extent to which a firm's formation was based on novel opportunities should be positively associated with business performance in entrepreneurial firms.*

Based on the above review, I have fashioned indicators of internal knowledge intensity based on different measures of pre-entry qualities of the KIE firm:

- Human capital inputs such as founder education and employee education.
- Human capital outputs like founders' previous experience.
- The two dimensional measure of knowledge scope and disparity measuring functional heterogeneity of the founders.
- Organizational origins in the form of being a corporate spinoff
- Opportunity-based formation factors.

The next section provides some rationale for studying the relationship between innovative performance and business performance in entrepreneurial firms, and develops some working hypotheses on what these relationships might be, while highlighting the knowledge intensive character of entrepreneurial firms.

3.4.4 Innovative performance as a driver of growth, volume, and development of firms

I began this work by commenting on the proposed relationship between new ventures and economic growth, and especially that of knowledge intensive new ventures and their influence on national economies. All signs point to that this relationship is deemed to exist by many organizational bodies in political and well as scientific areas of interest (OECD, 2003; OECD, 2008; Malerba, 2010; McKelvey and Lassen, 2013; OECD, 2013; Malerba et al., 2015). But in order for this to be the case, that is, in order for a knowledge intensive new venture to contribute positively to economic development and growth, it must itself achieve sufficient growth and development. Should there not be a tangible relationship between the measures of innovativeness and those of economic growth and performance on the firm level when speaking about these types of firms that intensively use, create, and apply new knowledge to solve their problems and to build new firms? With this in mind, I introduce the final research objective to be investigated in this chapter:

> **RO4:** *Explore the association between innovative performance and business performance of entrepreneurial firms.*

Traditional accounts of innovation convey it as an incredibly influential variable in shaping firm's long term profit and survival (Schumpeter, 1934; Penrose, 1959) in which "innovation depends upon the generation of feasible new capabilities, the operation of which adds new value to the existing circular stream of income, and thereby creates new profits and higher income" (Cantwell, 2002: 216). Many early innovation scholars were quite vocal regarding the innovative capabilities of established firms and their impact on competitiveness and profitability. For instance, in Utterback and Abernathy's (1975: 638) seminal treatment of innovation processes in different developmental stages of industries and products over time, they implied "strong and important relationships ... among the capability of a firm to innovate, its competitive strategy and the posture of its production resources". Later, Abernathy and Clark (1984) argued for different categories of innovation giving different weights to competitiveness in different environments, with incremental technical change being a strong shaping mechanism in increased returns.

More recently, studies have sought and found some link between the two relating both to large firms and SMEs that exhibit goods-based (Hughes, 2001; Van Auken et al., 2008; Gunday et al., 2011) and service-based innovativeness (Cainelli et al, 2006; Ariana Mansury and Love, 2008), and in both cases have yielded positive relationships. Others have taken a more cautious approach, pointing to contextual factors and moderators such as age of the firm, specific types of innovation, and corporate and national culture surrounding the venture (Rosenbusch et al., 2011).

Indeed there is a strong trend in the recent conceptual modeling of innovative and entrepreneurial firms which puts innovation at the forefront of competitiveness Both the dynamic capabilities literature (Teece et al., 1997; Eisenhardt and Martin, 2000, etc.) and the proponents of the knowledge based view of the firm (Grant, 1996) argue heavily that competitive advantage is the result of novel combinations of resources and capabilities. And that this, moreover, is a necessary but not sufficient component of innovation (in a broad sense). This viewpoint should easily extend to SMEs, and often does, as Rosenbusch and colleagues (2011) analyze at the meta-level. And since knowledge intensity in entrepreneurial firms is often claimed as having a profoundly positive effect on economic performance of regions, countries, and other systemic units of analysis (OECD, AEGIS, KEINS, and many more have claimed this), it is interesting to perhaps begin by measuring whether the innovative performance of firms with a potentially high level of knowledge intensity has a discernable influence on their own business performance.

Studies have often attempted to link together new firm innovativeness and economic well-being and prosperity on the macro-economic level, but it is interesting to assess whether these firms have been successful at themselves becoming profitable, healthy, and strong in terms of growth and development. In an advanced treatment of adaptive economic growth by Metcalfe and colleagues (2006:29), one of their primary postulates that can be taken away is the following:

"Growth, technical progress and the competitive process are inseparable; they are genuinely adaptive evolutionary processes driven by microeconomic diversity and coordinated by market and other institutions to generate emerging, ever-changing patterns of economic structure."

Metcalfe and colleagues succinctly convey the fact that aggregate growth is dependent on what is being aggregated, which is the growth of individual enterprise, and that increasing returns must be first measured on the firm level, from the bottom up, in order to assess evolutionary economic change as a whole.

Research more specifically addressing firms that may overlap with the KIE classification, YICs (young innovative companies), that addresses innovation's impact on growth has also emerged. Measuring innovation activity as R&D spending per employee, Coad et al. (2016) argue that this indicator positively affects firm growth for larger new firms, and negatively affects firm growth for smaller new firms, suggesting that R&D investments, while risky and potentially unsuccessful for young firms, can have a strong positive effect if the investment pays off, otherwise resulting in a fast decline in performance.

Additionally R&D investment becomes more stable the older the young firm becomes. This means that for young innovative companies, volatility is expected in the innovation-growth relationship, but that this relationship should have a positive trend. It is fruitful in the present study to analyze similarly how innovation affects growth, as the classification KIE extends beyond that of YIC to include firms in sectors that are not inertly high tech in nature, though they may also possess qualities that drive macro-economic growth in their innovative and knowledge intensive character.

Radicalness of the innovating firms is also as interest, as there remains much work to be done regarding how patterns of incremental vs radical innovation materialize in young firms and how this affects growth (Criscuolo et al., 2012). At the risk of overstating the point, it is necessary to evaluate the effectiveness of KIE firms' innovativeness in driving their own individual performance. Based on this review of the relationships stated above, I construct the following working hypotheses:

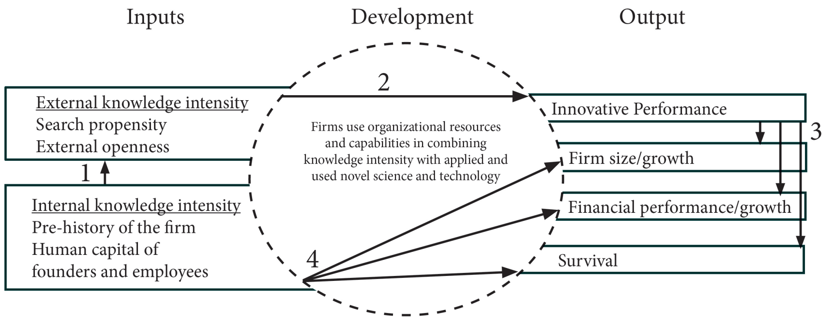
> **Working Hypothesis 4.1:** *Higher innovative performance in goods sales will positively associate with business performance for entrepreneurial firms.*
>
> **Working Hypothesis 4.2:** *Higher innovative performance in service sales will positively associate with business performance for entrepreneurial firms.*
>
> **Working Hypothesis 4.3:** *A higher degree of radicalness in products and services will positively associate with business performance for entrepreneurial firms.*

3.5 Summing up the research objectives and moving forward

This concludes the theoretical review, and derivation of both full and working hypotheses for subsequent analysis. Figure 3.1 outlines the conceptual model to be employed in the subsequent analyses. It is loosely based on McKelvey and Lassen's (2013) Knowledge Intensive Entrepreneurship creation model (see Chapter 2, Section 2.3.3 for the original model). In my modified version, external and internal knowledge intensity constitute inputs, while performance of the firm constitutes the output. New firms hold then both a stock of current internal knowledge at the time of, and leading up to, the moment of founding, and subsequently aspire to acquire and assimilate new external knowledge. They seek to 'synthesize and apply' both of these types of knowledge (Kogut and Zander, 1992, p. 384). This can be argued to manifest in external and internal knowledge intensity. The figure above shows the conceptual framework. Through organization specific capabilities combined with the use and application of new scientific and/or technical knowledge (which is not the main focus of analysis, and are included mainly to show the sort of "black boxed" effect that happens in the mid-stages of the process), firms utilize the two different types of knowledge intensity to augment their performance. The main variables of interest (shown in bullet points below) for each type of knowledge intensity will be derived and supported separately in the chapters to follow:

A firm's knowledge intensity is a complex, interwoven latent object, consisting of both external and internal components. This refers to the

Figure 3.1: KIE modeled after McKelvey and Lassen (2013)



Adapted from McKelvey and Lassen (2013: 29)

location from which the knowledge is derived in a physical sense. While it is obvious that when the venture itself is the unit of analysis, at inception, all resources and capabilities relating to knowledge intensity are "entering" the venture from the external environment, I refer to knowledge-based resources and capabilities housed inside the firm, including its formation factors, human capital indicators, etc., as internal constructs. External then, means that knowledge that stems from outside the firms' boundaries. I posit that the internal dimension influences the external dimension in a profound way worth investigation. These two dimensions together influence the performance outcomes located on the right side of the figure above. Based on the selection criteria of the survey and sample, to be discussed in the research method chapters, it can be assumed that there is at least a high likelihood that the firms studied are participating in these black boxed activities. This makes the empirical data to follow ideal for assessing knowledge intensity in entrepreneurial firms

KIE firms are proposed to draw on both internal and external knowledge, and intensely use it for competitive gains through heterogeneous combinations of organizational capabilities, with a proposed effect on performance as seen above. This serves as the conceptual framework around which the dissertation will base its empirical analysis. I will now briefly re-state all research objectives and hypotheses derived, and assemble them into 4 models. These models will serve as the basis of the empirical chapters to follow:

RO1: *Explore the association between external knowledge intensity and innovative performance in the entrepreneurial firm.*

H-1.1a: Search breadth is positively associated with innovative performance in new ventures.

H-1.1b: At higher levels of search breadth, the positive relationship will experience diminishing marginal returns, resulting in a marginally declining positive association.

H-1.2a: Search depth is positively associated with innovative performance in new ventures.

H-1.2b: At higher levels of search depth, the positive relationship will experience diminishing marginal returns, resulting in a marginally declining positive association.

H1.3 Search breadth should have a stronger effect than search depth on innovative performance.

WH1.4 Reliance on external sources of knowledge is positively related to innovative performance in entrepreneurial ventures.

WH1.5 Too much reliance on external knowledge sources will result in diminished marginal benefits in terms of innovative performance.

RO2 *Explore the association between internal knowledge intensity and external knowledge intensity in the entrepreneurial firm.*

WH-2.1: Higher levels of stocks of human capital in the form of education will positively influence the reliance on external knowledge of the firm

WH-2.2a: Founders having higher previous work experience at a university or research institute will be positively associated with the reliance on knowledge stemming from non-industry sources of knowledge such as specialist knowledge providers.

WH-2.2b: Founders having higher previous work experience in the same industry or having entrepreneurial experience will be negatively associated with the reliance on knowledge stemming from non-industry sources of knowledge such as specialist knowledge providers.

WH-2.3a: Founders' previous work experience at a university/research institute will be negatively associated with the reliance on intra-industry sources of knowledge

WH-2.3b: Founders' previous work experience in the same industry, or previous entrepreneurial experience, will positively associated with the reliance on intra-industry sources of knowledge

WH-2.4a: In terms of functional heterogeneity of the founding team, the level of knowledge scope should positively associated with the level of reliance on external knowledge from all categories.

WH-2.4b: The level of knowledge disparity should negatively associate with the level of reliance on external knowledge from all categories.

WH-2.5: For entrepreneurial firms, stemming from a previous organization (spinoff) will positively impact the reliance on knowledge stemming from all types of external knowledge sources.

WH-2.6: The extent to which a firm's formation was based on novel opportunities should positively influence the reliance on knowledge stemming from all types of external knowledge sources.

RO3: *Explore the association between internal knowledge intensity and business performance of the entrepreneurial firm.*

H-3.1: Higher levels of human capital in the form of education of founders and employees are positively associated with business performance in entrepreneurial firms.

H-3.2: Higher levels of human capital in the form of entrepreneurial, industrial, and academic experience of the founding team are positively associated with business performance in entrepreneurial firms

H-3.3a: In terms of founding team functional heterogeneity: Knowledge scope is positively associated with business performance.

H-3.3b: At higher levels of knowledge scope, it will have a diminished marginal association with business performance for entrepreneurial firms.

H-3.4: In terms of founding team functional heterogeneity: Knowledge disparity is negatively associated with business performance for entrepreneurial firms.

H-3.5: Having the organizational origin of being a corporate spinoff entrant is positively associated with business performance for entrepreneurial ventures

H-3.6: The extent to which a firm's formation was based on novel opportunities should be positively associated with business performance in entrepreneurial firms.

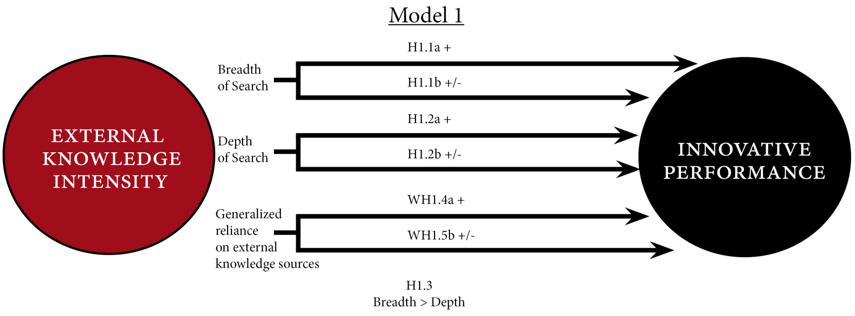
RO4: *Explore the association between innovative performance and business performance of entrepreneurial firms.*

> WH-4.1 Higher innovative performance in goods sales will positively associate with business performance for entrepreneurial firms.
>
> WH-4.2 Higher innovative performance in service sales will positively associate with business performance for entrepreneurial firms.
>
> WH-4.3 A higher degree of radicalness in products and services will positively associate with business performance for entrepreneurial firms. performance will positively associate with business performance for entrepreneurial firms.

The next chapters will outline the research design for achieving the research objectives derived here, as well as testing the hypotheses developed throughout the chapter, which will be summarized again when applied to each specific model. Chapter 4 will analyze the relationship between external knowledge intensity and innovative performance, depicted in Figure 3.1 above as Model 1, before moving on to investigate the link between internal knowledge intensity on external knowledge intensity, or Model 2. Chapter 5 will begin by looking into the association of internal knowledge intensity with economic/business performance, Model 3; following this, the relationship between innovative and business/economic performance will be empirically investigated, Model 4.

The following figures visually illustrate how the hypotheses are structured relative to the overarching concepts used in the research objectives for each model. Each construct is connected the outcome concept (or in some cases output construct) through hypotheses. As is the convention, '+' signs signify a positive relationship while ' - ' signifies a negative. A hypothesis that is followed by the sign '+/-' denotes an inverse quadratic relationship, or declining marginal positive relationship. I will re-apply this type of visual aid later in the thesis in the hope that it aids in interpretation of what are otherwise complex and multi-layered results.

Figure 3.2: Hypothesis Map Models 1 and 2



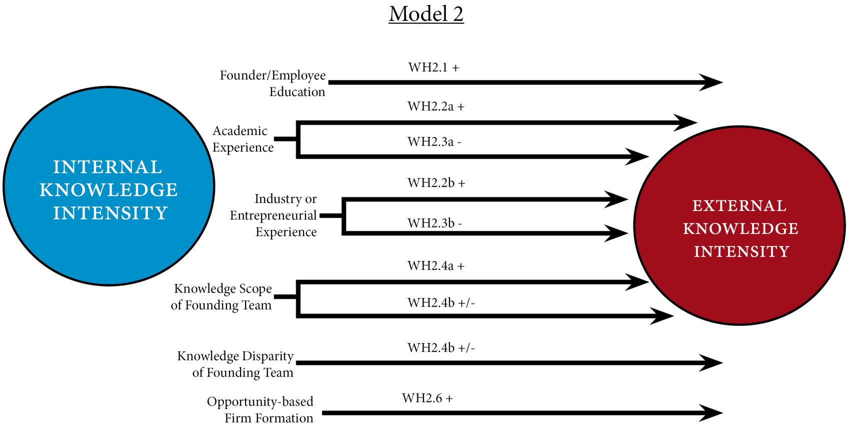
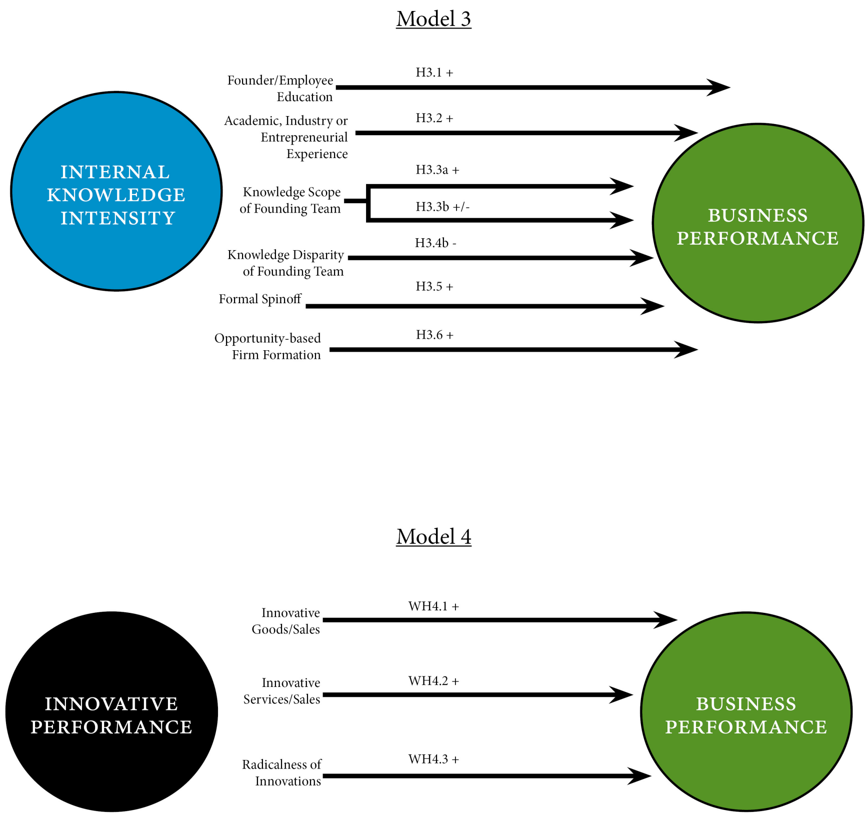


Figure 3.3: Hypothesis Map Models 3 and 4



Chapter 4

Research method, data, descriptives, and analysis: Part 1 - Models 1 and 2

This chapter of the dissertation outlines the research methods, techniques, and results of Models 1 and 2, those dealing solely with the data collected from the AEGIS survey. It models the relationship between external knowledge intensity (operationalized as search for and reliance on external knowledge sources), and innovative performance (operationalized as innovative goods and services proportional to sales and degree of radicalness of said innovations), as well as the relationship between internal knowledge intensity (operationalized as pre-entry founder, employee, and organization resources and capabilities) and external knowledge intensity. It gives an overview of the research methods and research process of this portion of the report, and includes explanations about the empirical techniques used, modeling, diagnostics and results. It also gives a slightly more in depth look at the survey used as the base of much of the analyses, the AEGIS survey, including its design and administration to the targeted sample.

Note for the reader: After restating the hypothesis, introducing the data and variables, and setting up the models, I will proceed to diagnose and provide limited interpretation of results and hypotheses.The reader will notice that I am quite conservative in my hypothesis confirmation strategy. I will say that a hypothesis is fully confirmed if it holds for all response variables which are used to represent a given concept, and partially confirmed not all hold. Since this is a PhD dissertation and not an article submitted for publication, I found that this holistic approach helped in illustrating the scope of the work that was undertaken, and in tracing the steps of the empirical analysis. For diagnostics I rely mainly on the use of residual analysis that fits the particulars of the statistical method being applied. Robustness checks of the different models are either included in the specification tables (see for instance the OLS regression of the logit transformed response variables in Model 1), or in the Appendix to this dissertation. Marginal model plots, quantile comparison plots, variance inflation factor, and residual plots when applicable. Model 2 was done using ordinary least squares estimation, so the diagnostics are slightly more extensive given that OLS allows for much more developed testing and assumptions. Also, I must forwarn the reader that I make great use of graphical and visual display material, since I feel that this can give a more in-depth perspective of the choices made throughout the research process. The diagnostics sections in general, especially for Model 2, might be skipped without losing any continuity, and are mainly included to showcase the depth of the analysis that I have gone to. Also, more statistically-inclined readers are sure to appreciate the detail included here. All of the hypotheses will be in any case reviewed and revisited in Chapter 6.

In order to investigate the knowledge intensive entrepreneurial firm, I use measures for external and internal knowledge intensity, in accordance with the previous chapter's theoretical grounding, and try to assess how these constructs interact, as well as affect performance of the newly founded venture that may be a KIE firm. To do this, a quantitative analysis strategy will be the main one employed. The choice of empirical model in each section rests on the dependent variable or phenomenon of interest and on the research setting. This chapter broadly aims to model the following research objectives and hypotheses derived previously:

> **RO1:** *Explore the association between external knowledge intensity and innovative performance in the entrepreneurial firm*
>
> **RO2** *Explore the association between internal knowledge intensity and external knowledge intensity in the entrepreneurial firm.*

As explained in the previous chapter, external knowledge intensity here is conceptualized as external search activities, that is, the degree of reliance placed on external sources of knowledge for innovation by the firm. This has been commonly referred to in the literature as *openness*, which I use as a proxy for external knowledge intensity. We generated the following hypotheses (and working hypotheses) for Model 1 based on an extensive literature review:

> H-1.1a: Search breadth is positively associated with innovative performance in new ventures.
>
> H-1.1b: At higher levels of search breadth, the positive relationship will experience diminishing marginal returns, resulting in a marginally declining positive association.
>
> H-1.2a: Search depth is positively associated with innovative performance in new ventures.
>
> H1-2b: At higher levels of search depth, the positive relationship will experience diminishing marginal returns, resulting in a marginally declining positive association.

H1.3 Search breadth should have a stronger effect than search depth on innovative performance.

WH1.4 Reliance on external sources of knowledge is positively related to innovative performance in entrepreneurial ventures.

WH1.5 Too much reliance on external knowledge sources will result in diminished marginal benefits in terms of innovative performance.

With such a large amount of preliminary hypotheses in an analysis, it may be of worth to briefly restate the claims of them. The first two two-part hypotheses (H-1.1 and H-1.2) are proposed in accordance to much of the innovation search theory on large firms, while taking into account the constraints in terms of networks, networking, and resources on small entrepreneurial firms; and the nature of knowledge intensive activities in potential KIE sectors, with knowledge intensity being largely a systemic phenomenon with a lot of emphasis on external connections and reliance. I expect that, though for different reasons in some regards, the direction of effects for breadth and depth of search will be largely the same as what has previously been found in the literature on larger firms, and to some degree, SMEs, for this sample of potentially knowledge intensive entrepreneurial firms. Moreover, as previously stated in Chapter 3, due to constrained resources, network limitations, and other related factors detailed above, breadth should be more effective in than depth in producing gains from openness for degree of radicalness of innovations produced. Successful commercialization of radical innovations may be more tied to the amount of sources more so than to the deep interaction with those sources, thus, H-1.3 is proposed.

Addressing the underlying areas of search into which these firms might delve, WH-1.4 and WH-1.5 are concerning the effects of broader 'areas' of reliance on innovative performance. Put plainly, the limited literature tied to search for external knowledge residing different specifications of types of sources leads to the assertion that a higher reliance on all of these types of sources leads to higher innovative gains by small new firms.

Later in Chapter 3, I developed multiple hypotheses regarding the association between internal knowledge intensity and external knowledge intensity in the given empirical setting. The following hypotheses were proposed in conjunction with the theoretical overview of this relationship, along with some summary of their derivation:

WH-2.1: Higher levels of stocks of human capital in the form of education will positively influence the reliance on knowledge stemming from all types of external knowledge sources.

For this I have drawn on the human capital arguments, especially that of human capital investments such as education and training. The higher the levels of education of both the founders as well as the employees of the firms are expected to positively impact the reliance on all types of external knowledge sources for these potentially KIE firms.

> WH-2.2a: Founders having higher previous work experience at a university or research institute will be positively associated with the reliance on knowledge stemming from non-industry sources of knowledge such as specialist knowledge providers.
>
> WH-2.2b: Founders having higher previous work experience in the same industry or having entrepreneurial experience will be negatively associated with the reliance on knowledge stemming from non-industry sources of knowledge such as specialist knowledge providers.
>
> WH-2.3a: Founders' previous work experience at a university/research institute will be negatively associated with the reliance on intra-industry sources of knowledge
>
> WH-2.3b: Founders' previous work experience in the same industry, or previous entrepreneurial experience, will positively associated with the reliance on intra-industry sources of knowledge

Prior experience of the founders should have differing effects, based on both the nature of the previous experience or occupation, as well as the source category of external knowledge that the effect will be mapped on. I expect firms founded by persons with prior experience from universities or research institutes to have a higher degree of reliance on these types of external knowledge. Conversely these firms should be comparably less reliant on intra-industry sources for knowledge. This effect is partly due to the fact that university-employed founders of a firm is a common gauge for that firm being an 'academic spinoff', and the literature on this type of firm shows strong connective links in terms of knowledge flows with these types of actors (Perkmann et al., 2013). On the other hand, firms with founders who have previous industrial or entrepreneurial experience within the sector of the current business should experience to some extent an opposite effect. I expect these firms to benefit more from intra-industry knowledge sources than from the other two, which I expect to be *negatively* related with the presence of industrial or entrepreneurial experience in the current sector.

> WH-2.4a: In terms of functional heterogeneity of the founding team, the level of knowledge scope should positively associated with the level of reliance on external knowledge from all categories.

> WH-2.4b: The level of knowledge disparity should negatively associate with the level of reliance on external knowledge from all categories.

I assert that the broader the scope of functional background knowledge encompassed by the founding team, the higher the reliance on external knowledge stemming from specialist knowledge providers. However, increased knowledge disparity among the founding team's functional backgrounds should have the reverse effect. Firms that do not overlap in their knowledge enough and/or are too dissimilar will have difficulties focusing on what to search for as a group. This will have an effect on knowledge search in all categories of sources.

> WH-2.5: For entrepreneurial firms, stemming from a previous organization, i.e. being a corporate spinoff, will positively impact the reliance on knowledge stemming from all types of external knowledge sources.

The literature on spinoffs suggests that they will be more likely to utilize external sources that stems from previous business, that is, those relationships and contacts that the founders acquired during previous careers or employment at the parent firm. I expect that the firm being a formal spinoff will increase its reliance on external knowledge sources of all types, due to heightened network awareness stemming from the parent firm, as well as more established trajectories within which the firm may search for new knowledge.

> WH-2.6: The extent to which a firm's formation was based on novel opportunities should positively influence the reliance on knowledge stemming from all 3 types of external knowledge sources.

I will investigate these two sets of hypotheses in two distinct empirical sets of models: The first outlines the effect of external knowledge intensity on innovative performance, and will be referred to broadly as Model 1. The other explores the underlying factors of external knowledge intensity as explained by different forms of internal knowledge intensity; namely, through founder, employee, and organizational resources and capabilities leading up to and during the inception period of the venture, and will be referred to as Model 2.

4.1 Data

In order to test these hypotheses, I draw on firm level data collected that is taken from the AEGIS survey, based on a sample collected during the

AEGIS project in 2011,. The survey was an exploratory attempt to map the activity and characteristics of knowledge intensive entrepreneurship in Europe (review Chapter 2 for more details about this survey and its precursors). The construction of the sampling frame used in the survey drew largely from the Amadeus database, with supplementation from a few other databases. Amadeus is a massive firm database owned by Bureau van Dijk, a privately owned business intelligence conglomerate. It covers over 12 million European-based business entities. Originally the sectoral query used in identifying the sampling frame, which was targeted to maximize the frequency of KIE firms, returned 547,678 companies, after cleaning, dropped down to 338,725 firms[1]. Contact information was found for 180,215 firms, and in order to retrieve the target sample from each country and each sectoral grouping (High tech, Medium and Low-tech, KIBS) the dataset was complemented by a few other databases (Dun and Bradstreet, Kompass, and others). This resulted in a final sampling frame of 202,286 firms. The survey team explained their rationale as follows:

"The fact that more companies had to be approached, also implied that more sample was needed than ex-ante forecasted. The sample from Amadeus database was not sufficient in order to achieve the desired number of interviews. Additional sample was purchased from other sources (Kompass/D&B) in order to achieve the targets per country/sector combination." (Caloghirou et al., 2011, p. 23). A target response rate was set at 4000 firms, and the sample was randomized with stratified sampling occurring in each distinct country (Croatia, Czech Republic, Denmark, France, Germany, Greece, Italy, Portugal, Sweden, and the United Kingdom). At the survey's completion, 4004 firms had been interviewed and surveyed. Their distribution across countries and sectors can be seen in table 4.1 below. The KIE definitions employed by the study served as a screening mechanism in order to reach firms that were most likely to be KIE. They needed to be: at the time less than 8 years of age; involved in market activities (exploiting remunerative opportunities), not subsidiaries or simply changed status existing firms, and not corporate venturing projects or corporate entrepreneurship.

The survey was carried out using telephone interviews: subcontracted by the research team through Global Data Collection Company. Since most of the variables in the survey are built on summated rating scales, one must be mindful that the answers collected are based largely on the respondent's subjective interpretation of the question, despite the fact that the survey

[1]For more extensive descriptive statistics, I refer the reader to the wealth of details available in Caloghirou et al., 2011; Malerba et al., 2015

Table 4.1: The sampling of the AEGIS survey - Origin of the AEGIS database*

Country	Amadeus	Dun and Bradstreet	Kompass	Other sources	Total	Completed Interviews	Screenouts or non-qualifiers
Croatia	660	202	1535	0	2397	200	201
Czech Rep.	1029	1995	0	22	3046	200	321
Denmark	5834	147	1909	0	7890	330	787
France	56503	639	0	0	57142	570	1906
Germany	34149	1539	1336	0	37024	557	2883
Greece	1367	277	0	2536	4180	331	369
Italy	49836	1999	0	0	51835	580	445
Portugal	4203	982	274	0	5459	331	235
Sweden	18727	2159	0	0	20886	334	501
UK	7907	1175	3222	123	12427	571	2934
Total	180215	11115	8276	2681	202286	4004	10581

*adapted from Caloghirou et al., 2011

was administered via telephone in the native language of the respondent, which likely afforded some degree of clarification for any ambiguity in the survey construction. Table 4.1 details the composition of firms included in the survey at each consecutive stage of the sampling process, while Tables 4.2 and 4.3 show and contrast the planned and achieved sampling of each country by sector.

The survey aimed to gather data in all sectors that could be classified as potential containing a high amount of knowledge intensive entrepreneurial firms. Table 4.4 shows the classifications used.

Since the AEGIS survey represents a unique set of respondents in that they were randomly sampled from sectors containing a high potential frequency of KIE firms (Malerba et al., 2015), it is a fitting dataset by which to analyze these types of new firms in terms of knowledge intensity and performance. Also, micro firms constitute the majority of the firms sampled in the AEGIS survey (64%), a population that receives as of yet little empirical attention in large scale innovation and entrepreneurship surveys. Since the CIS is known to exclude firms with personnel amounting to less than 10 employees (de Jong and Marsili, 2006), this survey presented an opportunity to assess a population of firms that has received considerably less attention from innovation surveys and surveyors, as well as exploring the concept of KIE firms by collecting data from what could potentially be KIE firms. In terms of formulation of survey questions regarding innovation processes and knowledge sources, much of the AEGIS survey was originally modeled after the CIS, including similarly subject oriented data and questions, with additional influence from established entrepreneurship and competitiveness gaging surveys like the General Entrepreneurship Monitor and the Kauffman survey finding its way into the design. This was carried out in order to make the survey more likely and able to capture the concept of knowledge intensive entrepreneurship (Caloghirou et al., 2011).

Table 4.2: Planned sample size by country and sector

	HTMS	LTMS	KIBS	Total
Croatia	40	108	52	200
Czech Rep.	40	76	84	200
Denmark	40	69	221	330
France	70	195	305	570
Germany	70	170	330	570
Greece	40	171	119	330
Italy	70	293	207	570
Portugal	40	163	127	330
Sweden	40	105	185	330
UK	60	177	333	570
Total	510	1527	1963	4000

Table 4.3: Achieved sample size by country and sector

	HTMS	LTMS	KIBS	Total
Croatia	35	115	50	200
Czech Rep.	25	92	83	200
Denmark	34	69	227	330
France	68	196	306	570
Germany	67	160	330	557
Greece	22	184	125	331
Italy	57	316	207	580
Portugal	31	170	130	331
Sweden	34	108	192	334
UK	47	192	332	571
Total	420	1602	1982	4004

Table 4.4: Selected sectors of the AEGIS survey

	NACE 1.1.
High-technology manufacturing sectors	
Aerospace	35.3
Computers and office machinery	30
Radio-television and communication equipment	32
Manufacture of medical, precision and optical instruments (scientific instruments)	33
Pharmaceuticals	24.4
Medium to high technology sectors	
Manufacture of electrical machinery and apparatus	31
Manufacture of machinery and equipment	29
Chemical industry (except Pharmaceuticals	24 (except 24.4)
Medium to low technology manufacturing sectors	
Wood/Furniture	36
Basic metals	27
Fabricated metal products	28
Low technology manufacturing sectors	
Paper and printing	21, 22
Textiles and clothing	17, 18, 19
Food, beverages, and tobacco	15, 16
Knowledge intensive business services (KIBS)	
Telecommunications	64.2
Computer and related activities	72
Research and experimental development	73
Other business services: (Legal/accounting, technical consulting including architectural and engineering activities, technical testing and analysis, labor recruitment and personnel provisioning, other mis. business activities	74.1 - 74.5, 74.8

In terms of content, the survey contains a range of variables detailing information on the following broad indicators:[2]

Section 1: General information about the firm

Section 2a: General information about the founder or the founding team

Section 2b: Firm's formation process

[2]The complete AEGIS survey questionnaire can be found in the appendix, section 8.4

Section 3: Market environment

Section 4: Strategy

Section 5: Innovation and business models

Section 6: Firms' performance and impact of the economic crisis

In order to answer the research objectives and hypotheses denoted above, I have drawn from all aspects of the survey, with the least amount of analysis occurring in regards to the 3rd and 6th sections.

4.2 A critical view of the data

The data and sample used is also not without caveats: The AEGIS survey included a variety of industries (both goods and services) as well as 10 different EU countries that differ on a number of indicators (including the macro-economic landscape, science policy, and demography). It was conducted by a professional data collection agency that used telephone (CATI) interviews to administer the questionnaire, targeting first founders, and if they were unavailable; partners, a CEO, or a managing director. It remains unclear while perusing the data who was actually reached based on position, and if they were the targeted body, for each individual firm. Also, recruitment and background information on and of personnel to perform the calls in the native languages of the respondents is not thoroughly detailed in the survey methodology, additionally, a lack of documentation exists concerning the translation of terminology to the most comparable wordings in order to retain meaning to the highest degree. Therefore it cannot be excluded that different response items may have been interpreted in different ways by either those conducting the survey or those participating in the survey during and after language translation.

General limitations of the sampling strategy included a high amount of potential firms being screened out due to ineligibility, over- and underrepresentation on the sectoral and country level. Additionally, the sampling procedure drew predominantly from the Amadeus/Orbis database by Bureau Van Dijk, with around 10% of the total sample being firms added from other sources as detailed in the empirical chapters. However, I was unable to find any documentation about how specifically these databases where combined, and using what parameters. It also is unclear whether the same criteria for selection were utilized in the supplemental sampling as in the Amadeus sampling. The random sampling component of the survey is also somewhat problematic. Though the sampling is reported as being random (Caloghirou et al., 2015), documentation of this was not acquired., and the methods employed suggest the possibility of some degree of convenience sampling may have occurred. Nonetheless, the statistical models operate under the assumption of randomness in the sampling and data collection process. This is also a reason for the detailed approach that I have taken concerning the variables, including assessing their approximate normality, as well as how summated rating scales were constructed, how they measure (or likely measure) the 'true scores' underlying the questionnaire, and how the component alignment of the rating scales alludes to underlying latent variables.

Also, the inherent difficulty in using an exploratory survey to capture a meaningful sample of a population of firms whose categorization is neither established nor unanimously agreed upon should be addressed. The sampling process makes use of OECD classifications, and previous theoretical discussions of what types of industries 'may' include knowledge intensive activities. Sometimes, exclusion of certain industries seems arbitrary, and the industry selection did not always match more widely used strata in intergovernmental organization publications, such as UN and OECD reports through the years. Like any scientific undertaking, the survey was the result of a vast number of decisions that needed to be taken by the design team, as well as those administering it, some of which have no clearcut right choice. Therefore, I acknowledge the limitations of the 'at times' imperfect data in the interpretation of the results as well as their economic implications.

In short, the data has many limitations, but a thorough and careful examination of it, and some data scaling, transformation, and modeling methods that have not seen widespread use in innovation and entrepreneurship studies provided a solid toolkit for analytical work on my part. A few words on the validity and reliability of the data should be noted upon. Since the AEGIS project was explorative in nature, and its sampling procedure for the survey component was based on a model that could maximize the occurrence of knowledge intensive entrepreneurial firms in the sampling frame, there may be some validity problems. It also only covers 10 EU countries, which have differing economic stabilities and variabilities. Hence, it can be difficult to generalize far beyond the sample. However, since so many sectors of different character are included, and since the process was professionally managed, it seems plausible that it is at least a predominantly representative sample of entrepreneurial firms in these sectors. The reliability of the data is supported by the random sampling procedure within the sampling frame, although since the survey was on such a broad scale, there was some degree of targeted sampling for filling a quota of representation for different industry types and countries. Since the operational measures taken here were built from well established variable coding techniques, many with roots in previous wide scale innovation/entrepreneurship surveys, construct validity and reliability should be strong, the latter of which has been assessed throughout using Cronbach's alpha and similar measures.

Regardless of these drawbacks, many of which are not uncommon for large-scale survey projects, the AEGIS survey represents the first step in gathering detailed information about knowledge intensive entrepreneurship from firms that have a high potential to be such. The data is rich in meaningful concepts and constructs, and provides an as of now unparalleled resource in investigating KIE as a firm-level phenomenon.

4.3 Operationalization of variables and descriptive statistics

I will now discuss my operationalization of the concepts put forth until now that are to be analyzed. The variables derived here will also be used in Chapter 5, when I will combine the AEGIS survey with up to date information from the Orbis database. In conjunction with this I will review relevant descriptive statistics of the variables used in these first two models, including their construction through different means.

4.3.1 External Knowledge Intensity variables - Importance of sources of external knowledge

In order to approximate the concept of external knowledge intensity, I have employed the concepts of Breadth and Depth of External Search (Laursen and Salter, 2006), as well as derived principal components via principal axis factoring. I use these two separately, producing two comparative sets of sub-models. In the AEGIS survey, firms were asked to rate the relative important of 11 different sources of knowledge for exploring new business opportunities, ranging from 1 being not important and 5 being extremely important. Table 4.5 presents the results of the (n = 4004) firms on this indicator in terms of percentage of the sample:

`Breadth`: This represents the combination of the 10 external sources of knowledge (see Table 4.5 above) expressed in regards to exploring new

Table 4.5: External sources of knowledge reliance by firms (%)

Sources of Knowledge (1 is not important, 5 is very important)	**1**	**2**	**3**	**4**	**5**
Q24_1 Clients or customers	2	2	10	24	62
Q24_2 Suppliers	13	13	23	25	26
Q24_3 Competitors	9	15	33	26	17
Q24_4 Public research institutes	44	22	20	9	5
Q24_5 Universities	47	20	18	10	6
Q24_6 External commercial labs/ R&D firms/Technical institutes	48	20	18	10	4
Q24_7 In house (know how, R&D)	24	7	16	25	27
Q24_8 Trade fairs, conferences and exhibitions	18	18	30	21	13
Q24_ 9 Scientific journals and other trade or technical publications	20	18	29	21	12
Q24_10 Participation in nationally funded research programs	58	16	13	9	5
Q24_11 Participation in EU funded research programs (Framework Programs)	62	13	11	8	6

business opportunities in the AEGIS survey questionnaire (while omitting the response representing in-house R&D activities, as the variable is only concerned with external-to-the-firm knowledge). The value 0 is assigned if the observation indicated the source was not important (a score of 1), the value 1 is assigned if the observation indicated that it was anything greater than not important (a score between 2 and 5). The number external sources are then summed for each firm to create the `Breadth` variable (min = 0, max = 10; Cronbach's alpha coefficient = 0.81).

`Depth`: This variable represents the deepness of collaboration of the sources of knowledge from the questionnaire. Built from the summated rating scale also used to construct `Breadth`, it was coded 0 if the observation was coded 1, 2 or 3; assigning a 1 if the observation was coded 4 or 5 (Thus only those sources ranked 4 or 5 are deemed to be very or extremely important). As `Breadth`, this variable takes on values between 0 and 10, where firms getting a score of 10 deeply collaborate with all external sources of knowledge listed in the questionnaire (Cronbach's alpha coefficient = 0.69) This binary coding approximates very nearly the method employed by Laursen and Salter (2006). These two variables are also used to test the hypothesis that there is also a curvilinear, or inverse quadratic, relationship between the concepts of interest. This is done by including the quadratic interaction effect of both variables in the regressions. This interaction is carried out using orthogonal polynomials generated using the `car` package in R (Chambers and Hastie, 1992; Fox and Weisberg, 2015). This is in order to remove the collinearity problems of including a 'raw' polynomial in the regression equation. I use orthogonal polynomials in most occurrences of terms involving powers in at least one model specification to avoid these problems. Sadly, including these terms in regressions renders their coefficients difficult to interpret (Fox and Andersen, 2011), which is one reason why I will often rely on more graphical interpretation of data in the analyses.

Additionally, principal components were extracted from the external knowledge sources conveyed in the survey questions for use in separate models. Measuring external knowledge intensity via a company's breadth and depth of external knowledge sourcing is in the first stage of this operationalization, similar to the operationalization of Openness, which has become a fairly well established construct in the literature. Nonetheless, the summation of binary outcomes used to construct the `Breadth` and `Depth` variables may fail to capture the more nuanced underlying variance conveyed by the summated rating scale for reliance on different external sources of knowledge. To counteract this in a more exploratory attempt to account for more of the variance in the summated rating scale used to derive the `Breadth` and `Depth` variables, a principal components analysis (PCA) of this summated rating scale was carried out, and its interpreted components run as independent variables (and later, as dependent variables). Due to the strong substantive weight of the non-rotated categories, no rotation method was employed. Of the ten components retrieved, three principal components were derived and retained:

- `EXPC1`: External, non-industry sources of knowledge (or specialized scientific and technological knowledge providers (Tether and Tajar, 2008)).
- `EXPC2`: Intra-industry sources of knowledge: Business and operations-based relationships (made up of clients, customers, suppliers and competitors)[3]
- `EXPC3`: Sources of (informal) codified knowledge stemming directly from academia and related communities.

Principal components analysis is "a statistical technique that linearly transforms an original set of variables that represents most of the information in the original set of variables [in order to] reduce the dimensionality of the original data set", for use in subsequent analyses (Dunteman, 1989: 7). These derived variables are orthogonal with one

[3] This category is also similar to Bengtsson et al.'s (2015) *Value chain partners*.

another, and maximize the variance accounted for in the original set of variables (ibid.). This technique can be extremely useful in understanding the underlying dimensions which account for the variation in a set of correlated variables. Here, it is of interest to model the preceding regressions with principal components added in. The fact that the components are per-construction orthogonal with one another facilitates their interpretation as regression coefficients. Using the scree plot[4] comparing the Eigenvalues generated by the PCA with the number of components (See Table 4.6 below), it can be seen that 3 principal components account for a cumulative 63.45 percent of the total variance of the 10 variables used to estimate `Depth`. While the Kaiser-Guttman criterion (Kaiser, 1960; Guttman, 1954) commonly applied to principal components analysis suggests retaining only those with an eigenvalue of $\Lambda > 1$, I retain the 3rd component, as it is extremely close to 1, and constitutes the end point in the bend in the scree plot, as is also often practiced.[5] Upon closer inspection through a bivariate correlation matrix comparing the components with the original variables, a pattern begins to emerge: The right-side table in Figure 4.1 below shows the bivariate correlations of principal components of external knowledge sources with the original set of variables.

How high the components' bivariate correlations must be in order to be interpretable by the researcher is highly discretionary, as no reliable guidelines exist. Though it is generally thought that patterns in the correlations must be readily identifiable in order for substantive interpretation of the components to follow (Dunteman, 1989). Hence, of most interest are the correlations of components with certain variables relative to the other components. Cronbach's alpha coefficient of the scale itself was reported as 0.81.

To aid additional interpretation, a biplot[6] was constructed to map the singular value decomposition (the first three singular values in this case) of the rating scale as a whole. Seen in Figure 4.2:

[4] Introduced by Cattell (1966), the scree plot compares the number of components with their *eigenvalues*, and is a commonly used visual aid in determining the number of components to retain from a PCA.

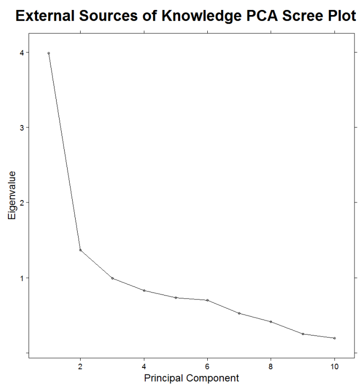
[5] I thank Professor Bill Jacoby for this recommendation. Additional OLS regressions on the dependent variables were also carried out using all 10 components, and while a few of the lower order components were significant in the regressions, they were not substantively interpretable. Also important to note is that there is a lack of consensus regarding the most 'effective' method to decide what to retain in a PCA.

[6] A biplot, which jointly plots row and column effects in a matrix, is a commonly applied statistical tool in visual appraisal of large matrix data structures. It is often used to supplement principal components analysis by showing inter-unit distance, unit clustering, and visual representations of variances and correlations between vectors or variables in a summated rating scale (Gabriel, 1971).

Table 4.6: Principal component importance for reliance on external knowledge sources

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Standard deviation	1.99	1.16	0.99	0.91	0.85	0.83	0.72	0.64	0.50	0.44
Proportion of Variance	0.39	0.13	0.09	0.08	0.07	0.07	0.05	0.04	0.02	0.01
Cumulative Proportion	0.39	0.53	0.63	0.71	0.79	0.86	0.91	0.95	0.98	1.00

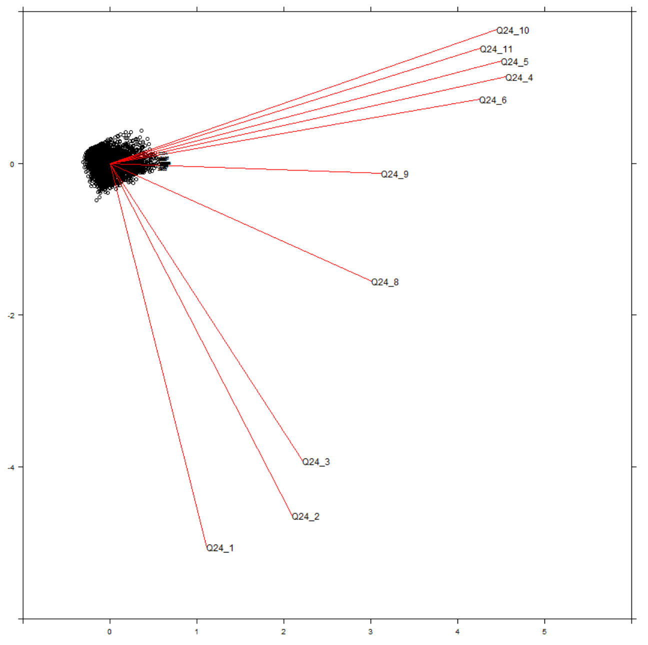
Figure 4.1: Scree Plot of External Knowledge Source PCA & PCA-Variable correlations



	PC1	PC2	PC3
Q24_1	0.20	-0.69	0.22
Q24_2	0.37	-0.63	0.15
Q24_3	0.39	-0.53	0.06
Q24_4	0.81	0.15	0.16
Q24_5	0.80	0.18	0.17
Q24_6	0.76	0.11	0.14
Q24_8	0.54	-0.21	-0.63
Q24_9	0.55	-0.02	-0.64
Q24_10	0.79	0.24	0.13
Q24_11	0.76	0.20	0.15

The biplot gives further evidence to the underlying components we have observed in the unrotated correlation matrix. The first principle component seems to be most highly correlated with Q24_4, Q24_5, Q24_6, Q24_10, and Q24_11. Comparing this coding with the labels assigned to the questions by the AEGIS survey, one can see that the first principal component (`EXPC1`) is highly correlated with these *external, non-industry sources of knowledge*, mainly related to the collaboration with state, national, or regional research-based or academic entities: Roughly equivalent to Tether and Tajar's (2008) specialist knowledge providers (SKPs). Conversely, the second principal component is most correlated with Q24_1 - Q24_3: Representing clients or customers; suppliers; and competitors. This second component (`EXPC2`) can thus be interpreted as explaining the shared variation of sources of knowledge through *business and operations-based relationships*. The third principal component is most correlated with Q24_8 and Q24_9: Trade fairs, conferences and exhibitions; and scientific journals and other trade or technical publications. This last component (`EXPC3`) can be seen as *sources of informal codified knowledge stemming directly from academia and related communities.* These three components will be seen as approximations of these concepts, although they are only mathematical transformations of the variables. In testing hypotheses and working hypotheses regarding reliance on all types of sources of external knowledge, I use these 3 components to proxy.

Figure 4.2: Biplot of external knowledge sources in vector space using singular value decomposition



4.3.2 Innovative performance variables

Innovative performance has commonly been measured through turnover of new or substantially improved products or services over a relatively recent time period, usually three years (Caloghirou et al., 2004; Jantunen, 2005). Laursen and Salter (2006) measured innovative performance in terms of ability to produce radical and incremental innovations in terms of turnover. Here I take a combinative approach of these two methods in order to measure innovative performance:

In terms of research design, the concept of interest is *innovativeness*, the conceptualized construct of which is *innovative performance* at the firm level. This variable is measured by 3 indicator variables, second order factors, or measures on the construct: Thus, The latent dependent variable is approximated by 3 dependent variables, which were built using the following questions from the AEGIS survey: First, the respondent was asked if their company introduced any new or significantly improved (i.e. innovative) goods or services in the past three years; if yes, they were asked to estimate the share of both innovative goods and innovative services within total sales. They were also asked if the innovative goods *and/or* services were new to the firm, new to the market, or new to the world.

The first and second dependent variables combined uses the proportion of innovative sales of goods/services to the total sales generated by the firm over the past three years, approximating the amount of sales generated by this recent innovation activity:

- `InnoGoods`: The proportion of innovative (new or improved) goods to that of total sales.
- `InnoServ`: The proportion of innovative (new or improved) services to that of total sales

The third dependent variable measures the degree of radicalness of the innovation practices of the firm during the past three years, approximating the degree of radicalness of innovation in that firm in general.[7] Originally the respondent was given the opportunity to indicate up to three different products or services that were new or significantly improved, and rank them according to the provided scale. In order to construct a meaningful model from the data, a new variable was constructed using conditional indicators of the highest achieved level of novelty for each firm:

- `RadInn`: The Degree of Radicalness of innovations (goods or services) introduced to the market by the firm over the past 3 years.

    0 = No new innovations introduced.

    1 = Up to the "new to the firm" level innovations introduced.

    2 = Up to and including "new to the market" level innovations introduced.

    3 = Up to and including "new to the world" level innovations introduced.

[7]Note: Well I am aware that this terminology might provoke, as radicalness may draw up connotations of radical vs. incremental innovations, and even incremental innovations may well be 'new to the world' etc., I speak here of the level of radicalness in terms of degrees of 'subjective' radicalness to the setting or the environment, be it the firm, market or world; and not how 'objectively' radical the innovation itself may be in its own right

4.3.3 Internal knowledge intensity variables

To construct the variables for internal knowledge intensity, several different groupings were employed based on the survey data:

Employee level human capital indicators: Employee education: Two variables representing education levels of employees were created:

- `EmpEdu`, the first, conveys the percentage of full- and part-time employees who have obtained a degree at the tertiary level of education.
- `EmpHiEdu`, the second, conveys the percentage of full- and part-time employees who have obtained a postgraduate degree, namely a Master's degree or equivalent or a PhD.

Founder level human capital indicators

`FoundEdu` - Founder educational attainment: The average level of educational attainment across the founding team. This variable is interval and bound between 0 and 5, with the original categories ranging from:

- Elementary Education
- Secondary Education
- Bachelor degree
- Postgraduate degree
- PhD

`FoundEnt` – Founder Entrepreneurial Experience: A binary variable taking the value 1 if one or more of the founders had prior experience either owning an existing firm, owning a firm that has ceased operations, or was self-employed, and taking the value 0 otherwise.

`FoundUni` – Founder University Experience: A binary variable taking the value 1 if one or more of the founders of the venture had prior experience working as a University or research institute employee, and 0 otherwise. This serves as a commonly applied proxy for the firm being an academic spinoff (Perkmann et al., 2013; Perkmann and Walsh, 2007).

`FoundInd` – Founder Industry Experience: The highest number of years of professional experience attained by any member of the founding team, a positive discrete value.

`AgeMax` – The age (recorded at the time of the survey) of the oldest member of the founding team, a positive discrete value, bound in ordinal categories:

1. Aged 18-29
2. Aged 30-39
3. Aged 40-49
4. Aged > 50

`Spinoff` - A binary variable measuring whether (1) or not (0) the focal firm formally materialized out of another prior organization.

Additionally, a number of indices were constructed in order to evaluate the effects of team heterogeneity within the following areas of founders' expertise when the firm was founded, which were deemed relevant for the operations of the company by the respondents:

- Technical and engineering management
- General management
- Product design
- Marketing
- Finance

This was achieved by constructing the following indices for the above-named functional categories: First, Blau's (1977) traditional index of functional heterogeneity (`Blau` in the data) was built, where p_i is the proportion of team members who have experience in the ith functional category:

$$1 - (\sum_{i=1}^{n} p_i^2) \tag{4.1}$$

This measure cannot be readily applied to consider more than one separate dimensions simultaneously in one team member; they are assumed mutually exclusive. This index calculates an overall net effect of both positive and negative forces of team heterogeneity (Cantner et al., 2010).

In a first attempt to correct of this lack of multidimensionality in the functional categories, I have also constructed the Attneave (1959) entropy based index of functional background diversity (See also Boone and Hendricks, 2009; Bunderson and Sutcliffe, 2002). This index takes into account the fact that people gain experience on more than one dimension, but in a sense it still attempts to average across categories and team members, and is purported to measure the diversity of the functional backgrounds between team members, while subtracting the intrapersonal functional diversity (ibid.). This index is composed of three separate measures: The first is the marginal entropy measure, or Shannon measure:

$$H_x = \sum_{n}^{i} p_i log \frac{1}{p_i} \tag{4.2}$$

. . . where i stands for each functional category. The second is the marginal entropy between team members:

$$H_y = \sum_{m}^{j} p_j log \frac{1}{p_j} \tag{4.3}$$

. . . where j stands for each member of the founding team, and the combination, or total entropy of the frequency table:

$$H_x y = \sum_{nm}^{ij} p_{ij} log \frac{1}{p_{ij}} \tag{4.4}$$

Finally, the Attneave's transmission measure T_{xy} (`Attneave` in the data) is calculated as:

$$H_x + H_y - H_{xy} \tag{4.5}$$

Cantner et al.'s (2010: 4) two-dimensional firm heterogeneity index was employed to obtain an additional and more comprehensive measure of founding team competences. It may be constructed using two components: *Knowledge scope*, or, "the breadth of a new venture team's knowledge stock"; and *knowledge disparity*, or "the deviation in the knowledge stocks of the individual team members". Knowledge scope is calculated by combining two separate measures, team *variety* and team *diversity*, while knowledge diversity is achieved through combining team *dissimilarity* and team *non-redundancy*. Knowledge scope is, like

Attneave's measure, an entropy based one. It is calculated here as done in Cantner et al. (2010): From the entropy measure[8]:

$$v_a(s) = \left\{ \left(\sum_{i=1}^{n} s_1^a \right) \frac{1}{1-a}; a \geq 0, a \neq 1 \right\} \tag{4.6}$$

$$v_a(s) = \left\{ \lim_{a \to 1} \left(\sum_{i=1}^{n} s_1^a \right) \frac{1}{1-a}; a = 1 \right\} \tag{4.7}$$

... s_i is the weighted probability that founding members are experienced in the functional category i. The number of *team experiences* in each functional category I is weighted against the total number of functional categories within which the team of founders is has at least some experience in, thus:

$$s_i = \frac{\sum_{j=1}^{m} x_{ij}}{\sum_{i}^{n} \sum_{j=1}^{m} x_{ij}} \tag{4.8}$$

Where n is the total number of categories, and m is the total number of team members. x_{ij} is a binary measure taking the value 1 if team member j is experienced in category I, and 0 otherwise. From this it follows that variations in the parameter a is prioritized as an absolute value of variety in functional experience (taking low values), or evenness of distribution of functional experience (taking high values). Variety and diversity are formed with two disparate values of a, namely 0 and the limit of a as it asymptotically approaches $+\infty$. Variety is then equal to

$$v_0(s) = \sum_{i=1}^{n} s_1^0 = z \leq n \tag{4.9}$$

Where z is the amount of functional categories where at least one member of the team is experienced. Cantner et al. (2010) normalize this index, and here the same procedure is followed.

Diversity is equal to:

$$v_{+\infty}(s) = \frac{1}{max(s_i)} \tag{4.10}$$

[8]To maintain the clarity of Cantner and colleagues' equations, I retain the notation used in that paper.

... and is the weighted probability of the functional categories within which the team is most experienced, constituting an indicator of *de-concentration* of prior experience in various functional categories (Cantner et al., 2010). It is also normalized before use.

Now *knowledge scope*, or the variable, `KScope` may be derived by taking the mean of both the variety and the diversity indices that have been constructed. Higher values may indicate a broader knowledge base, but a less concentrated one, present in a founding team.

In deriving *knowledge disparity*, the first step is to calculate the *dissimilarity* measure for each pair of founders in the team:

$$f_i^{A,B} = \left\{1\, if\; x_i^A \cap x_i^B = 1\; for\; all\; i \in N \cap i \in [1,...,n]; 0\; otherwise\right\} \tag{4.11}$$

This may subsequently be summed over all functional categories, yielding ...

$$F^{A,B} = \sum_{i=1}^{n} f_i^{A,B} \tag{4.12}$$

... as the number of categories in which founders A and B share prior functional experience. The dissimilarity measure is obtained by contrasting the overlap of experience with the total combined individual functional backgrounds (Cantner et al., 2010):

$$Dissim^{A,B} = 1 - \frac{F^{A,B}}{(\sum_{i=1}^{n} x_i^A + \sum_{i=1}^{n} x_i^B)/2} \tag{4.13}$$

Ranging between 0 and 1, 0 constitutes complete overlap of functional backgrounds, while 1 indicates the opposite. The team *dissimilarity* measure used in subsequent analysis is the mean of all these pairwise measures, taken for every reported founding member combination in the AEGIS survey. The more dissimilar, the greater degree of dispersion within the founding team's functional knowledge base. Finally, the *non-redundancy* measure is calculated by dividing the number of functional categories z in which the team is experienced by the total functional experience possessed by the team in all categories; or:

$$z/\left(\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij}\right) \tag{4.14}$$

The *knowledge disparity* indicator, or the variable `KDisp` used in subsequent analysis is obtained by taking the mean of both the *dissimilarity* and *non-redundancy* indices, again following Cantner et al. (2010).

Firm Formation Factors

AEGIS's survey included a summated rating scale in which recipients were to rate the importance of a series of factors for the formation of the company. To gauge the importance of these factors without including all questions, a principal components analysis was used to extract the most variance captured possible out of the scale. For theoretical details, please see the variables detailing External Knowledge Source reliance above. The factors (Table 4.7) were denoted and accordingly distributed in their answers:

Table 4.7: Formation Factors Descriptive Statistics by percentage*

	1	2	3	4	5
Q13_1 Work experience in current activity field	4.60	3.22	9.22	21.85	60.81
Q13_2 Technical/engineering knowledge in field	10.94	6.42	14.36	25.10	42.81
Q13_3 Design knowledge	20.93	14.84	23.38	20.23	20.08
Q13_4 Knowledge of the market	2.62	4.40	19.43	32.29	41.06
Q13_5 Networks built during previous career	7.37	8.44	21.28	27.62	34.77
Q13_6 Availability of finance	12.41	14.79	25.20	21.08	25.82
Q13_7 Opportunities in a public procurement initiative	47.88	16.38	18.13	9.67	6.57
Q13_8 Existence of a large enough customer	16.13	13.04	23.48	23.95	23.00
Q13_9 Opportunity deriving from technological change	20.58	14.76	25.87	22.53	15.51
Q13_10 Opportunity deriving from a new market need	12.79	11.66	28.92	27.67	18.28
Q13_11 Opportunity deriving from new regulations or institutional requirements	31.74	19.46	24.08	13.91	9.54

* 1 = Not Important, 5 = Very Important

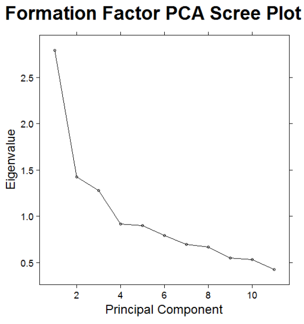
Similarly to the preceding principal components procedure for external knowledge sources, I returned component to variable correlations, and retained those that occur just after the first "bend" in the data, as observed using the scree plot method below. While the descent is somewhat less smooth than is often the case, one can clearly see that the variance explained declines strongly after the 3rd component, which accounts for more than 1 unit variance. Thus, retaining 3 components seems an acceptable decision (see Table 4.8).

Table 4.8: Principal component importance for firm formation factors

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
Std dev.	1.67	1.19	1.13	0.95	0.95	0.89	0.83	0.81	0.74	0.73	0.65
Prop. of Var.	0.25	0.12	0.11	0.08	0.08	0.07	0.06	0.06	0.05	0.04	0.03
Cum. Prop.	0.25	0.38	0.50	0.58	0.66	0.73	0.80	0.86	0.91	0.96	1.00

Retaining the first 3 components for regression allows us to capture 50% of the variance explained by the rating scale. While not ideal, it will suffice, as the alpha coefficient of the scale itself is lower than desired (around 0.70,

Figure 4.3: Principal components analysis of formation factors



	PC1	PC2	PC3
Q13_1	-0.30	0.71	-0.01
Q13_2	-0.38	0.47	-0.56
Q13_3	-0.47	0.21	-0.47
Q13_4	-0.45	0.31	0.46
Q13_5	-0.39	0.44	0.44
Q13_6	-0.43	-0.10	0.38
Q13_7	-0.57	-0.26	0.17
Q13_8	-0.45	-0.01	0.21
Q13_9	-0.67	-0.25	-0.33
Q13_10	-0.65	-0.38	-0.09
Q13_11	-0.64	-0.30	-0.02

suggesting that some variables fail to map onto our 3 retained components. The correlation matrix between the components and the original variables give us our component classifications.

Based on the correlations we identify the following component "values" which are extracted to become variables in the dataset:

- `FF1`: Opportunity based factors (Q13_11, Q13_10, Q13_9, Q13_7)
- `FF2`: Experiential and network based factors (Q13_1, Q13_4, Q13_5)
- `FF3`: Specialized knowledge-based factors (Q13_2, Q13_3)

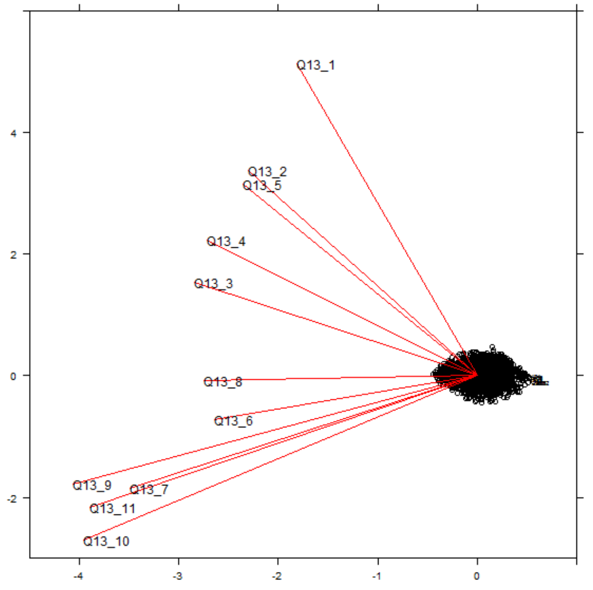
The biplot (see Figure 4.4 below) confirms the directional grouping that we suspected from our PCA, and reassures us that the components make sense from a visual point of view as well as from correlation analysis. The signs of the extracted (by observation) components `FF1` and `FF3` are reversed in the variable generation, since the correlations between the components and the Q13 variables of interest were assessed in the negative portion of the data space relative to the `FF2` component[9].

4.3.4 Control Variables

We also include a set of control variables for all regression in different combinations:

[9]In general, the direction of the effects does not matter in the derivation of component-variable correlations, and can be reversed, as long as all are reversed for the given component.

Figure 4.4: Biplot of firm formation factors in vector space using singular value decomposition

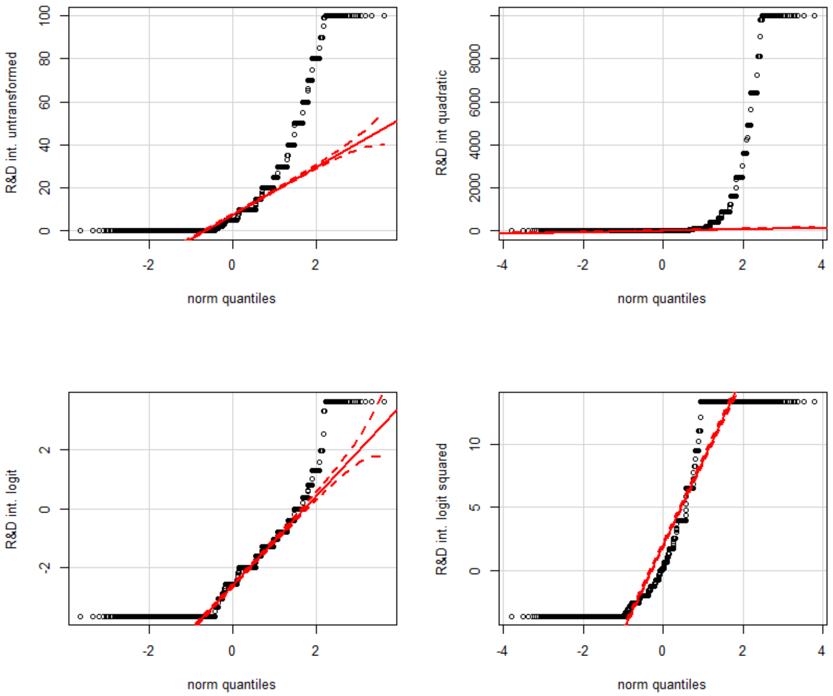


FirmAge: Based on the year the venture was established subtracted from year of collecting data: 2015 (screened for change in legal status of existing firm).

IntlSales: Percentage of international sales estimated by survey respondents. The logit, or log-odds, transformation was employed on this variable following graphical interpretation of the distribution of the data (following Fox, 2016).

R&DInt: R&D Intensity (by % of sales invested) estimated by survey respondents, in order to control for the effect of R&D on our variables of interest (Hagedoorn and Cloodt, 2003) and the effect of absorptive capacity of the firm (Cohen and Levinthal, 1990; Tsai, 2001). I have included a quadratic term for this variable as well in most regressions, in order to assess the presence of declining marginal returns to R&D Intensity for performance. Also, based on residual plotting and quantile plotting, the quadratic function fit the data rather well. Consequently, depending on the regression application, this variable is transformed into either a logit (log-odds) function, or an orthogonal polynomial function of the logit transformed variable, which is generally advisable with power transformations of discrete variables (Fox and Weisberg, 2011: 218). The quantile comparison plot below (Figure 4.5) compares the variable before and after transformation: from untransformed, to logit, to quadratic logit:

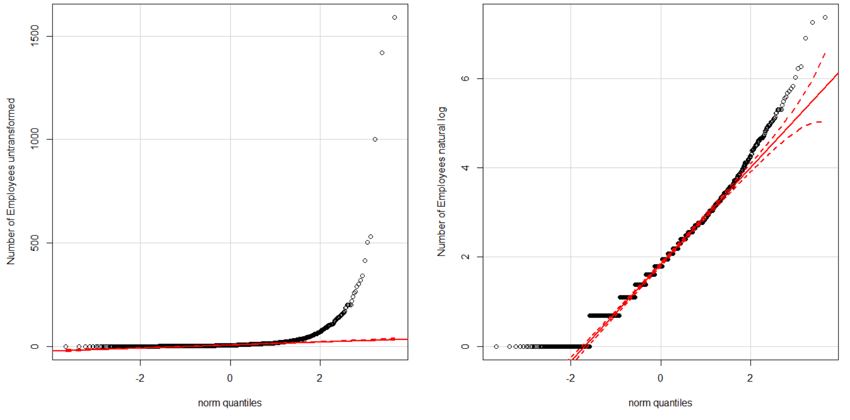
Figure 4.5: Transformations of R&D Intensity variable plotted against normal distribution quantiles



One can see that the fit to the quantile normal conditional distribution (the red line) improves markedly with subsequent transformations, while the tail of the distributions remain outside the confidence envelope, the fit is adequate for control purposes, and has approximated normality passing well. Additionally, I reviewed the residuals of this variable from trial regressions, which suggested that a quadratic fit may better fit most models. I have often included such a term. It stands to reason that R&D Intensity (or, if one desires, absorptive capacity, for which R&D Intensity is often a proxy in innovation studies) might have decreasing marginal benefits at higher levels.

`Emp`: The number of employees (full + part-time) is used to control for the size of the firm. By using this measure in our controls, we also indirectly account for other measures related to firm size, this control variable is the survey reported value, taken in 2010/2011. To correct for the occurrence of some firms having reported having no employees, I add 1 and take the natural logarithm. The graph on the left in Figure 4.6 plots the normal quantiles of the untransformed variable. The graph on the right, the natural log + 1 version, is clearly a better representation of a normally distributed variable. With the outliers at the top of the distribution being reigned in considerably towards normality.

Figure 4.6: Number of Employees plotted against normal quantiles



`Regions`: Additionally, the EU region of the firm (which was compared with the distribution of companies across countries, and found to be a better indicator of explained variance) was included in the regression to control for regional and national differences. Viewing simple regression plots (Figure 4.7) of the distribution of each variable on `InnoGoods` lends some rationale for this transformation: That is, not much variance appears lost in rescaling.

`SectorClass`: HTMS (High-tech manufacturing sectors) ¸ LTMS (Low-tech manufacturing sectors); KIBS (Knowledge intensive business services); OBS (Other business services). These were constructed based on the sector selection of the AEGIS survey itself (for reference, see Table 2.1). Similarly to the country identifiers, the sectors were derived from combining categorically the different sampled sectors. This was in order to smooth out the effects of the control variable, and to match the categories assigned to the sectors in the AEGIS project and survey. Figure 4.8 also shows the summarizing of the variance on one response variable used in the models. Other results proved to be similar enough in pattern not to warrant inclusion in the text[10].

Figure 4.7: Distribution of country and regions on Innovative goods/sales

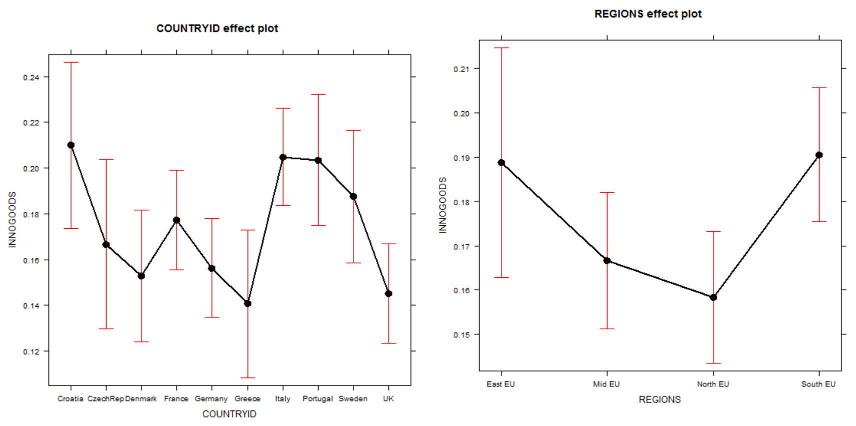
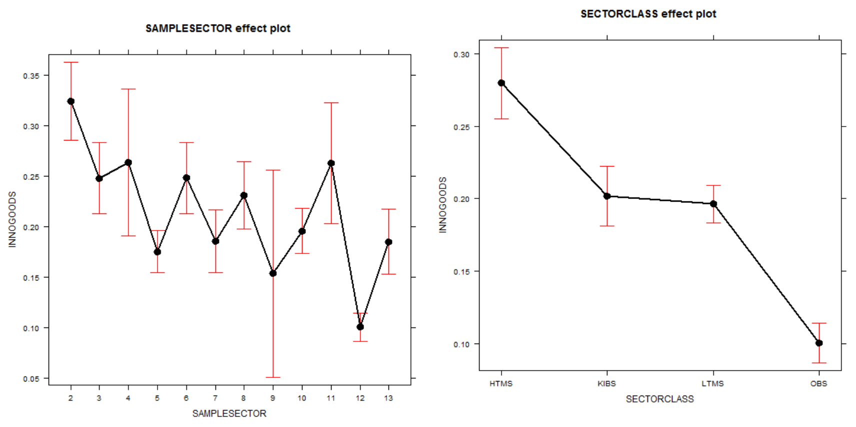


Figure 4.8: Distribution of sample sector and sector class on Innovative goods/sales



[10]Sectors: ICT (2), Machinery (3), Chemicals (4), Paper/printing (5), Textiles (6), Food products (7), Wood (8), Telecom (9), Computers (10), R&D (11), Other business services (12), Metals (13).

4.4 Summary statistics

Tables 4.9 and 4.10 present the number of firms sampled in each sector in terms of the variables used in the Model 1 regressions, the degree of radicalness of innovation, average R&D Intensity of the sector, and the calculated means of the Breadth and Depth indicators. Lastly, the percentage of innovative goods and innovative services as a proportion of total sales by sector is displayed. As expected, many firms, between 20% and 45% depending on sector[11], have no innovation at all, and new to the world innovations are scarce in the data, while new to firm are much more prominent. R&D Intensity is scarce in low tech, and semi-prevalent in high tech fields. Telecom seems strongest in innovative services while ICT, R&D development, and Chemicals lead in innovative goods.

Table 4.9: Model 1 variables by Industry/Sector, n = 3659

Sample Sector	# firms	% no Inno	% New to Firm	% New to Mkt	% New to World	Avg R&D Int	Breadth mean	Depth mean
ICT manufacture	159	27	43	18	13	19.7	7.6	3.8
Machinery and equipment	193	37	39	16	8	11.5	7.0	3.1
Chemical industry (ex. pharma)	46	24	48	15	13	18.6	7.5	4.0
Paper and printing	559	40	38	17	5	10.7	6.7	6.8
Textile and clothing	187	44	35	14	7	10.7	6.8	3.6
Food, beverages, tobacco	251	40	38	15	7	8.1	7.2	3.6
Wood and furniture	216	43	36	15	8	9.1	7.0	3.3
Telecommunications	23	26	57	9	9	13.5	7.3	3.4
Computers	477	30	38	24	8	18.3	6.7	2.9
R&D	65	32	28	21	19	42.7	8.3	4.1
Other business services	1252	46	34	16	4	10.6	6.5	3.0
Manufacture of metals	231	37	42	16	5	9.6	7.0	3.0
Total	3659	39	37	17	6	12.4	6.8	3.2

Table 4.10: Model 1 variables continued

Sample Sector	% of Inno. Goods/Sales	% of Inno. Services/Sales	% of Intl Sales
ICT manufacture	32.1	16.4	23.1
Machinery and equipment	24.3	9.9	24.2
Chemical industry (ex. pharma)	26.3	12.6	24.0
Paper and printing	16.9	17.1	10.3
Textile and clothing	24.1	9.9	22.2
Food, beverages, tobacco	18.7	7.7	9.5
Wood and furniture	22.7	9.0	12.9
Telecommunications	15.3	34.3	7.4
Computers	19.1	24.4	12.5
R&D	22.9	24.1	38.8
Other business services	9.9	19.4	12.0
Manufacture of metals	18.4	12.9	17.0
Total	17.0	16.9	14.3

[11] The sampled sector Aerospace, with only 1 firm, was excluded following listwise deletion.

Tables 4.11 gives an overview of the Model 2 variables of interest in descriptive terms. Many of the sectoral results of the descriptives are not unexpected, with more traditional knowledge intensity variables measuring higher in high tech and KIBS industries than in low and mid tech ones. Moreover, indicators of functional heterogeneity seem largely consistent across industries. Most firms are not spinoffs from previous organizations, though the highest frequency comes from low tech sectors like Wood and Furniture and Textiles.

Table 4.11: Model 2 variable means by Industry/Sector, n = 3331

	N	% Spinoff	R&DInt	FoundEdu	FoundEnt
ICT Manufacturing	144	9.03	20.60	3.38	0.31
Manufacture of machinery and equipment	169	14.20	11.75	2.70	0.33
Chemical industry (including Pharmaceuticals	42	4.76	18.12	3.36	0.38
Paper and printing	499	10.22	10.84	2.96	0.34
Textile and Clothing	175	21.71	11.68	2.54	0.34
Food, beverages, and tobacco	234	17.95	7.89	2.68	0.37
Wood and furniture	194	23.71	9.78	2.42	0.37
Telecommunications	21	9.52	14.52	3.10	0.38
Computer and related activities	447	11.41	18.94	3.44	0.36
Research and experimental development	56	14.29	40.86	4.41	0.32
Other business service activities	1146	12.57	10.56	3.46	0.29

	FoundUni	FoundInd	AgeMax	EmpEdu	EmpHiEdu
ICT Manufacturing	0.06	16.26	3.28	0.27	0.16
Manufacture of machinery and equipment	0.01	18.34	3.31	0.17	0.10
Chemical industry (including Pharmaceuticals	0.05	16.07	3.62	0.22	0.14
Paper and printing	0.03	13.94	3.11	0.22	0.13
Textile and Clothing	0.01	16.13	3.18	0.09	0.05
Food, beverages, and tobacco	0.01	14.48	3.16	0.13	0.07
Wood and furniture	0.00	16.68	3.16	0.08	0.04
Telecommunications	0.05	11.62	2.95	0.21	0.10
Computer and related activities	0.04	12.76	2.99	0.35	0.22
Research and experimental development	0.27	16.66	3.41	0.52	0.38
Other business service activities	0.02	15.64	3.28	0.33	0.22

	KScope	KDisp	Attneave	Blau	IntlSales %
ICT Manufacturing	0.41	0.49	-0.12	0.71	23.03
Manufacture of machinery and equipment	0.39	0.48	-0.09	0.73	22.53
Chemical industry (including Pharmaceuticals	0.40	0.47	-0.10	0.73	24.76
Paper and printing	0.38	0.48	-0.11	0.73	9.42
Textile and Clothing	0.39	0.52	-0.09	0.72	23.88
Food, beverages, and tobacco	0.39	0.51	-0.10	0.73	10.08
Wood and furniture	0.45	0.52	-0.11	0.75	13.07
Telecommunications	0.38	0.46	-0.07	0.71	6.71
Computer and related activities	0.40	0.49	-0.11	0.72	13.50
Research and experimental development	0.38	0.48	-0.08	0.70	40.57
Other business service activities	0.37	0.47	-0.11	0.68	12.15

It can also be observed by looking at the number of observations per sector what was detailed above about the AEGIS survey procedure. There is a distinct over-representation of 'Other business service activities'. 'Paper and printing' as well as 'Computer related activities' list quite highly as well. The average level of R&D Intensity is at just below 8% in the least intense sector (Food and beverages), and above 40% in Research and experimental development (unsurprisingly). While education levels are somewhat of a middle of the road construct, with most founders having at least a tertiary education (3 out of 5) in high tech and service sectors, but not in low/mid tech sectors. Moreover, average values of entrepreneurial experience across industries does not seem to overly vary, with all at around 20 to 30 percent. University experience is quite low for founding teams across all sectors. Also, though many of these sectors being ’knowledge intensive’ according to the OECD, tertiary employee education levels only exceed 30% in a few industries. At their lowest, they are around 8-9% for Wood and Furniture as well as Textiles. Somewhat surprising is that the industries with the most relative corporate spinoffs are in low tech sectors: Wood and Textiles.

Table 4.12: Descriptive statistics for Model 1 averaged across industries

	vars	n	mean	sd	med	trim	mad	min	max	rng	skew	kurt	se
InnoGoods	1	3659	0.17	0.26	0	0.12	0	0	0.99	0.99	1.54	1.56	0
InnoServ	2	3659	0.17	0.25	0	0.12	0	0	0.99	0.99	1.55	1.77	0
RadInn	3	3659	1.90	0.90	2	1.79	1.48	1	4	3	0.72	-0.35	0.01
RadInnOS	4	3659	0.88	0.91	1.39	0.84	0.23	-0.16	2.75	2.90	0.07	-1.09	0.01
Breadth	5	3659	6.80	2.54	7	6.94	2.97	0	10	10	-0.21	-1.07	0.04
Depth	6	3659	3.18	2.12	3	2.93	1.48	0	10	10	1.04	0.96	0.04
FirmAge	7	3659	11.10	2.17	11	11.12	2.97	8	15	7	0.11	-1.42	0.04
Intl Sales	9	3659	14.25	26.29	0	7.43	0	0	100	100	2.04	3.11	0.43
R&D Int	10	3659	12.44	19.25	5	8	7.41	0	100	100	2.57	7.27	0.32
Emp	11	3602	12.76	46.83	5	7	5.93	0	1590	1590	22.32	647.99	0.78

Table 4.12 shows descriptive summary statistics for variables in Model 1. Across all industries the averages are quite low. About 17% of sales come from innovative products/services across the whole sample (where the variables are reported), while the average level of radicalness is just below the new to the market level. The average size of firms reads at about 12 persons, though we can clearly see that the variable might be heavily skewed, with on firm having nearly 1600 employees. This further justifies a logarithmic transformation.

Table 4.13: Covariate correlations of Model 1*

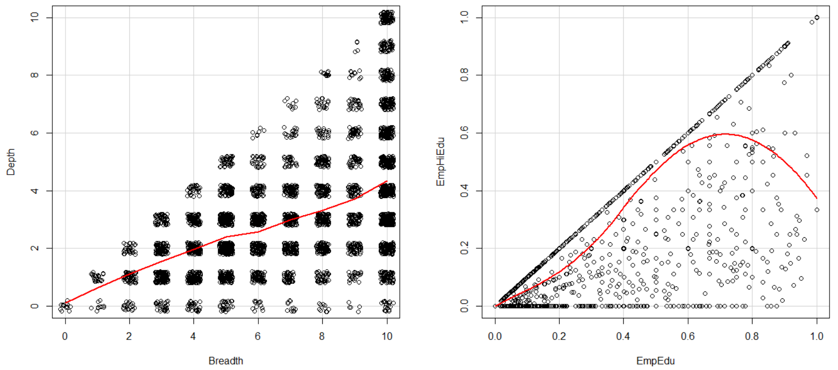
	1	2	3	4	5	6	7	8	9
1. InnoGoods									
2. InnoServ	0.10								
3. RadInnOS	0.53	0.51							
4. Breadth	0.13	0.08	0.19						
5. Depth	0.13	0.08	0.16	0.56					
6. FirmAge	-0.02	-0.03	0.02	0.03	0.01				
7. Intl Sales	0.15	0.00	0.12	0.08	0.06	0.06			
8. R&D Int	0.29	0.20	0.24	0.21	0.19	-0.01	0.17		
9. Emp	0.01	-0.02	0.03	0.08	0.02	0.06	0.08	-0.01	
10. Sector	0.02	0.02	0.09	0.15	0.13	-0.01	-0.03	0.02	0.04

*Values significant ($p < 0.05$) if > 0.02

Looking at the pairwise correlations for Model 1 in Table 4.13, two things are apparent; there is potentially high correlation between the radicalness indicator and the other two innovativeness response variables. However these are not in the same regression models so it is more an indication that they may represent a similar latent construct, innovative performance. `Breadth` and `Depth` are also somewhat correlated with each other (0.56).

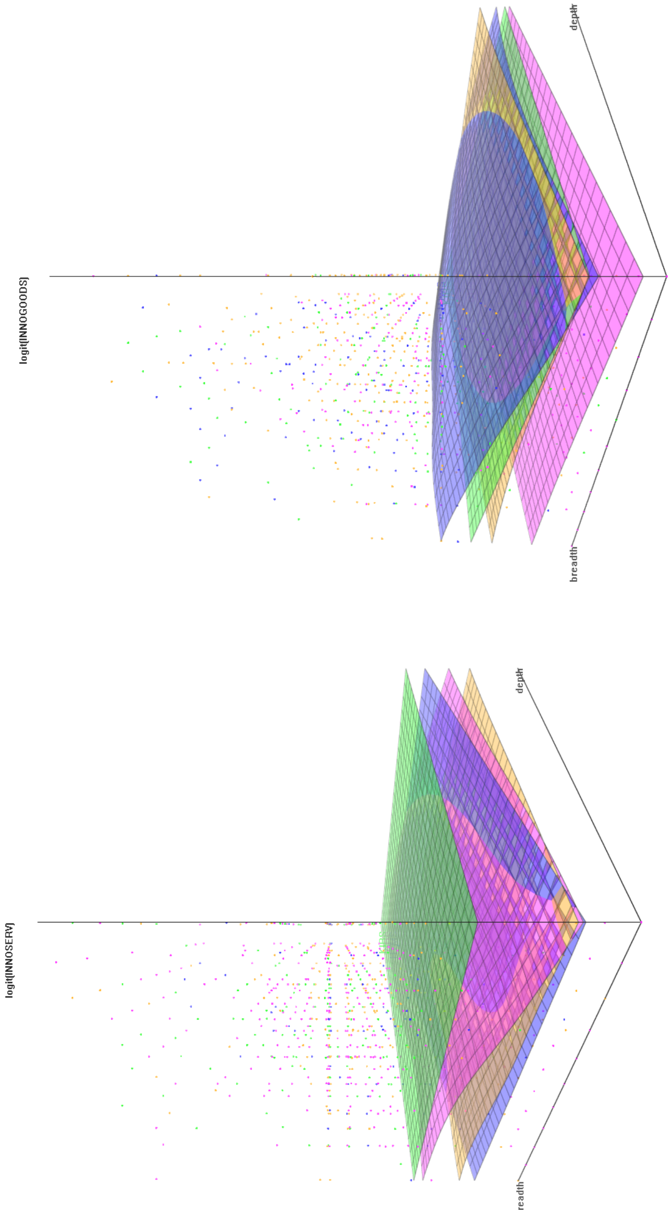
This is likely due to the fact that they are similarly derived from the same summated rating scale. The following graph (Figure 4.9, left pane) plots the two variables together in a jittered scatterplot with a loess smoother. One can plainly see that `Depth` is in fact a subset of `Breadth`, and can never take a value that is higher than its corresponding value of `Breadth`.

Figure 4.9: Key variables in jittered scatterplot with loess smoother



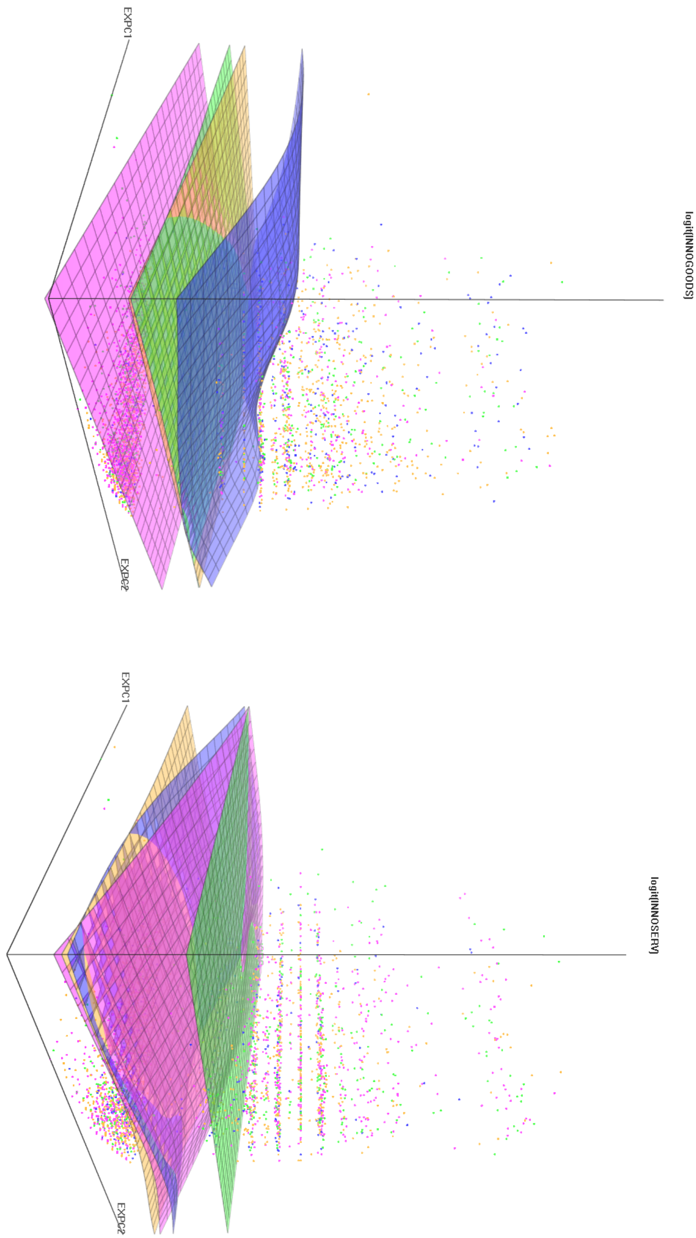
To counteract this potential problem, and for other reasons as well related to the inclusion of squared terms in the regressions, I convert these two variables to orthogonal polynomials in the subsequent analyses for select model specifications. This is achieved by using contrast matrices in R to generate the polynomials (Venables and Ripley, 2002). Fox and Weisberg (2011) recommend using this procedure for second-order or higher polynomials of regression terms when the values of the variable are discrete in order to correct for collinearity problems. The next plot is a 3D scatterplotting of Breadth and Depth on Innogoods and Innoserv, grouped by sector class with loess surfaces (Figure 4.10). One can see that there are potential linear relationships (different magnitudes for different sectors) between the x-and z-axis variables with the y-axis variables. Note again though how the space in the graph where values of `Depth` > `Breadth` is empty of observations. Similarly, I plot `EXPC1` and `EXPC2`, the highest variance explaining principle components, against the dependent variables. Smoothed non-parametric regression planes show that levels for both variables differ quite substantially depending on sector, but most show a predominantly positive association between the y-axis and x- and z-axis variables. One exception is HTMS in both graphs (the blue plane) which takes on more exponential patterns of association in the loess regression.

Figure 4.10: 3D scatterplot of `Breadth` and `Depth` variables by sector with `InnoGoods` and `InnoServ`:



Color Key for Figure: Green =KIBS, Magenta = OBS, Cyan = HTMS, Yellow = LTMS

Figure 4.11: 3D scatterplot of `EXPC1` and `EXPC2` variables by sector with `InnoGoods` and `InnoServ`:



Color Key for Figure: Green =KIBS, Magenta = OBS, Cyan = HTMS, Yellow = LTMS

Table 4.14: Descriptive statistics for Model 2 following listwise deletion

	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
IntlSales	3331.00	14.36	26.41	0.00	7.52	0.00	0.00	100.00	100.00	2.03	3.02	0.46
R&DInt	3331.00	12.60	19.34	5.00	8.19	7.41	0.00	100.00	100.00	2.59	7.46	0.34
FoundEdu	3331.00	3.14	1.07	3.00	3.13	1.48	1.00	5.00	4.00	-0.10	-0.94	0.02
FoundEnt	3331.00	0.33	0.47	0.00	0.28	0.00	0.00	1.00	1.00	0.74	-1.46	0.01
FoundUni	3331.00	0.03	0.16	0.00	0.00	0.00	0.00	1.00	1.00	5.76	31.21	0.00
FoundInd	3331.00	15.25	10.72	15.00	14.58	10.38	0.00	60.00	60.00	0.62	0.08	0.19
AgeMax	3331.00	3.20	0.83	3.00	3.29	1.48	1.00	4.00	3.00	-0.66	-0.53	0.01
EmpEdu	3331.00	0.25	0.26	0.17	0.22	0.25	0.00	1.00	1.00	0.73	-0.61	0.00
EmpHiEdu	3331.00	0.16	0.22	0.00	0.11	0.00	0.00	1.00	1.00	1.40	1.00	0.00
Emp	3331.00	11.65	37.21	5.00	6.81	5.93	0.00	1420.00	1420.00	23.60	774.69	0.64
Spinoff	3331.00	1.14	0.35	1.00	1.05	0.00	1.00	2.00	1.00	2.10	2.40	0.01
KScope	3331.00	0.39	0.22	0.34	0.38	0.25	0.00	0.93	0.93	0.34	-0.76	0.00
KDisp	3331.00	0.49	0.18	0.50	0.51	0.09	0.00	0.92	0.92	-1.26	1.78	0.00
FF1	3331.00	-0.02	1.65	0.02	-0.01	1.69	-5.65	4.50	10.15	-0.10	-0.06	0.03
FF2	3331.00	0.02	1.18	0.08	0.08	1.12	-5.14	3.07	8.21	-0.48	0.56	0.02
FF3	3331.00	0.01	1.13	0.04	0.03	1.09	-3.93	3.65	7.59	-0.08	0.06	0.02
EXPC1	3331.00	-0.02	1.98	-0.38	-0.19	2.03	-3.50	6.24	9.74	0.72	-0.08	0.03
EXPC2	3331.00	0.01	1.15	0.08	0.06	1.12	-5.98	2.64	8.62	-0.55	0.88	0.02
EXPC3	3331.00	0.01	1.00	-0.03	-0.01	1.01	-3.41	3.90	7.30	0.12	-0.05	0.02
FirmAge	3331.00	11.07	2.17	11.00	11.08	2.97	8.00	15.00	7.00	0.13	-1.42	0.04

Table 4.15: Pairwise Correlation matrix for Model 2

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.
1. Firm_age																				
2. IntlSales	0.05																			
3. R&DInt	-0.03	0.17																		
4. FoundEdu	-0.03	0.11	0.16																	
5. FoundEnt	-0.01	0.03	0.06	0.02																
6. FoundUni	0.01	0.08	0.16	0.20	-0.05															
7. FoundInd	-0.01	0.04	0.02	-0.04	0.10	0.00														
8. AgeMax	0.12	0.08	-0.02	0.06	0.13	0.04	0.51													
9. EmpEdu	0.02	0.12	0.17	0.48	-0.07	0.15	-0.06	0.00												
10. EmpHiEdu	-0.03	0.09	0.17	0.52	-0.04	0.18	-0.05	0.00	0.73											
11. Sector	0.00	-0.03	0.02	-0.12	0.13	-0.04	0.07	0.02	-0.05	-0.14										
12. Emp	0.04	0.09	-0.01	0.04	0.01	0.00	0.02	0.05	-0.01	-0.03	0.03									
13. Spinoff	0.00	0.01	0.03	-0.03	0.13	0.02	0.08	0.07	-0.08	-0.05	0.16	0.07								
14. Attneave	0.00	0.01	-0.02	0.02	0.08	0.04	0.03	0.04	-0.05	-0.06	0.07	0.03	0.01							
15. KScope	-0.01	0.02	0.08	0.09	0.08	-0.01	0.05	0.06	0.01	0.03	0.05	0.02	0.04	-0.43						
16. KDisp	-0.07	0.02	0.03	-0.02	0.02	-0.04	0.01	0.02	-0.01	-0.01	0.02	-0.01	-0.02	0.16	0.44					
17. Blau	0.02	0.00	0.05	0.03	0.06	0.02	-0.01	-0.03	-0.03	0.01	0.04	0.01	0.03	-0.28	0.45	-0.15				
18. FF1	0.02	0.04	0.18	-0.07	0.04	0.02	0.14	0.07	-0.02	-0.03	0.17	0.06	0.04	-0.03	0.11	0.04	0.08			
19. FF2	-0.03	0.05	-0.03	0.05	0.00	-0.02	0.27	0.10	0.05	0.07	0.00	-0.03	0.03	-0.02	0.05	0.04	-0.03	-0.01		
20. FF3	0.03	0.05	0.18	0.03	-0.02	0.06	-0.03	-0.06	0.02	0.01	0.06	-0.04	0.00	0.01	0.06	0.08	0.04	0.00	0.00	
21. EXPC1	0.01	0.07	0.23	0.00	0.05	0.09	0.03	0.03	0.04	0.00	0.18	0.07	0.06	0.03	0.05	0.03	0.05	0.44	-0.09	0.03
22. EXPC2	0.00	-0.03	-0.09	-0.18	0.01	-0.10	-0.05	-0.05	-0.14	-0.15	0.03	0.01	-0.03	-0.01	0.04	0.05	0.05	0.11	0.04	-0.11
23. EXPC3	0.00	0.05	0.03	0.08	0.03	0.00	-0.02	-0.03	0.06	0.04	-0.13	-0.02	-0.03	-0.01	0.02	0.03	0.01	-0.02	0.03	0.02

Correlations significant at $p<0.05$ if $x > 0.03$

Table 4.14 shows the mean values of each variable in Model 2 across all sectors. On average, about a third of founding teams have entrepreneurial experience, while only 3% have university employment or research experience. Employees with tertiary degrees seem just below the OECD knowledge intensive activity sectoral definition at 25%. Only about 14% of firms are formally spun out of previous organizations. On average, founding teams have at most 15 years of industry experience.

Pairwise correlations for Model 2 (Table 4.15) variables reveal a moderately high correlation between `EmpEdu` and `EmpHiEd`, again stemming from the fact that the latter is a subset of the former, since if one has at least a Master's or PhD level education, one has per default at least a tertiary degree level education. As a result, the models to follow are quite careful to track the effects of these two variables, along with the `FoundEdu` variable, when used in the same model specification, as they may interfere with each others' effects and associations through multicollinearity. Again, graphical interpretation confirms this (Figure 4.9, right pane).

The 3 plots in figure 4.12 show R&D intensity plotted against the 3 response variables in Model 1 representing the firm's innovative performance. It can be seen that while associations vary, and this plot does not control for other factors of influence, there are clear trends viewable. Innovativeness in goods has a near linear association with the HTMS sectors, while it has a more inverse curvilinear shape of association with KIBS sectors. The other two are less pronounced, but it is discernable that all experience positive gains from R&D up to (or after) a certain point. For innovativeness in services, the trend is more constant between sectors, with all 4 having at least a slight inverse curvature (KIBS and OBS are unsurprisingly more associated and the relationships are more distinct). Finally, radicalness of innovation (overall level) is has an overall positive association with all 4 sectors with varying curvature. This strengthens the argument that R&D intensity is a meaningful control across the empirical models.

Additional plotting using loess curves is used to investigate a few different data relationships preliminary to regression analysis of Model 2: First `EmpEdu` is plotted against `FF3` by `FoundUni`. Since Founder's university experience, tertiary education of employees, and the importance of technical and design knowledge for forming a firm all represent the broad concept of education, and are all used in the regression model, I wanted to ensure no strange relationships existed between them. Figure 4.13, left pane, confirms this, with a more or less null relationship plot, though founders having a University background on average relied more on technical and design knowledge. The right pane in the same figure shows `FoundInd` plotting against `FF2` by `FoundEnt`.

Figure 4.12: R&D intensity and Innovativeness: Scatterplots

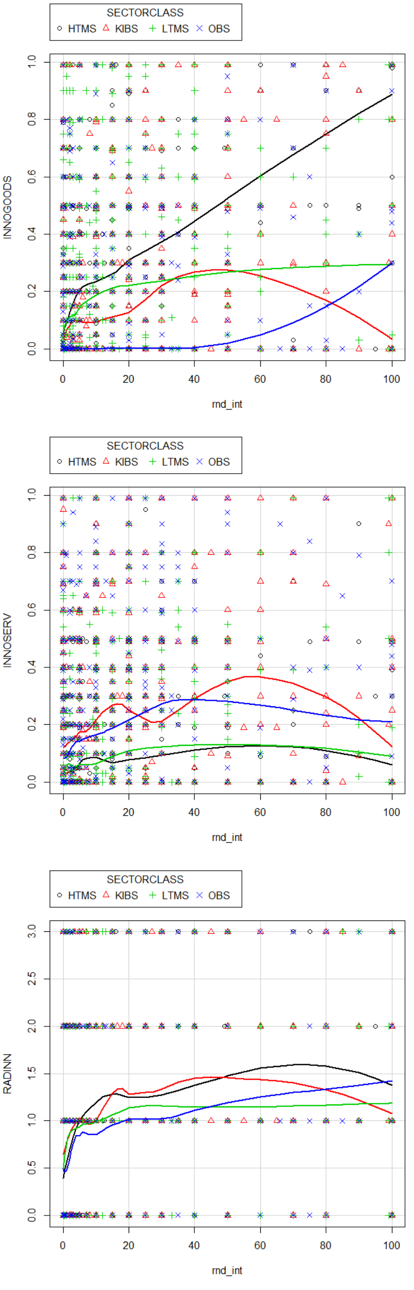


Figure 4.13: Key variables in Model 2 scatterplots

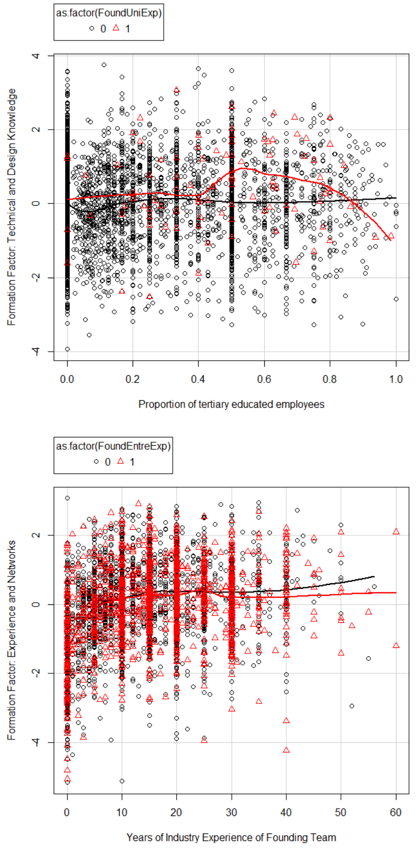
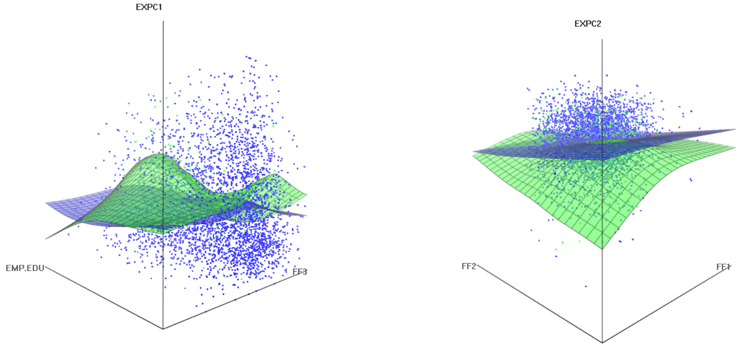


Figure 4.14: Model 2 3D scatterplots



Left figure: Green = University experience, Blue = No such experience; Right figure: Green = Spinoff, Blue = Not a spinoff

Here again, potentially interrelated variables show no signs of interference. The two plots in 4.14 show 3D scatterplots against `EXPC1` and `EXPC2`. The left pane shows a potential relationship between `FF3` and `EXPC1`, and a weak relationship between `EmpEdu` and `EXPC1`. It seems that University experience is influential at least in locally weighted average regression for `EXPC1`. The right pane a potential relationship between both `FF1` and `FF2` with `EXPC2`, here spinoffs actually seem to rely less on `EXPC2`. Many more relationships might be revealed in this way, but to conserve space I will not do so. However, we can already now see some potentially interesting relationships.

4.5 Estimation Methods

For the empirics of these two sets of models drawing from the AEGIS survey dataset, regression techniques have been chosen based on the set of variables in the survey in its final form. In the case of Model 1, techniques initially considered were Ordinal Logit, and Tobit, regression models. These however proved to be poorly matched to the data for a variety of reasons; therefore, more advanced techniques were applied[12]: An alternating least squares optimal scaling (ALSOS) routine was undertaken in order to rescale an ordinal dependent variable into one that could be used in ordinary least squares (OLS) regression. Additionally, a quasi-maximum likelihood estimation (QMLE) fractional logit model (Papke and Wooldridge, 1996) was chosen over Tobit due to intricacies in the data that demanded this treatment included the distribution of the residuals in the data. Model 2 was carried out using OLS regressions with principal components, and using the same data.

The first model to be selected was in regard to the innovative performance variables. This section documents the model selection process as well as the setting up of the data for analysis. Two of the response variables in this set of regressions, `InnoGoods` and `InnoServ`, are based on a question in the AEGIS questionnaire where respondents were provided with the option of indicating that they don't know what proportion of innovative goods or services are representative of the firm's total sales during the past 3 years. Additionally, the question that `RadInn` was based on also offered a *'don't know'* alternative. A substantive interpretation of these don't know answers is difficult, due to the fact that in this situation, *don't know* gives no indication of the true ratio of goods or services to sales. As *don't know* answers are difficult to interpret when the question asked requires

[12]In most cases however I include the more basic techniques in one column of the regression tables to follow for comparison

specific knowledge and/or not interpretable via any underlying continuum (here the latent variable *Innovative Performance*), these cases were subject to listwise deletion in accordance with the recommended missing value analysis literature (Acock, 2005; Little and Rubin, 2002). As a result the number of cases drops from 4004 to 3659 firms for models with response variables `InnoGoods` and `InnoServ`, and 3855 firms for `RadInn`[13].

While simply performing a logit transformation on the dependent variable may produce more desirable results when dealing with a proportion (Fox and Weisberg, 2011), this is not always practiced in innovation studies. Commonly, when two of the variables of interest are cornered, or censored, as a proportion is, Tobit's model of censored regression may be used in subsequent analyses (Long, 1997). However, since in this dataset, these depending variables which are fractions between 0 and 1 do not have a substantial pileup effect at both ends of the spectrum (Wooldridge, 2002/2012), as well as the conceptual mismatch between the variables being defined as, rather than limited to, lying between 0 and 1 (Cook, et al., 2008), the Tobit model was not an appropriate choice. In the modelling of the dependent variables `InnoGoods` and `InnoServ`, which have a large percentage of values clustered at 0, we have applied the Bernoulli quasi-maximum likelihood estimation (QMLE) Fractional Logit model as developed by Papke and Wooldridge (1996). The third variable, `RadInn` was originally intended for use in an ordered logit model. However, due to violations of the parallel regression assumptions of that particular model using a Brant (1990) test[14] , an alternating least squares optimal scaling (ALSOS) routine was conducted using the `optiscale` package in R, and the optimally scaled variable was then modeled using OLS regression (cf. Young, 1981; Jacoby, 1999; Jacoby, forthcoming). The resultant variable `RadInnOS`, is proposed to also represent Innovative Performance. The tables below document the results of the analysis. ALSOS is commonly used in order to test the measurement assumptions of a variable, and through it, empirical transformations of variable values may give insights about the appropriate level of measurement for that variable (Jacoby, 1999). The technique, commonly used in psychometric analysis, optimally scales the variable to find the maximized goodness of

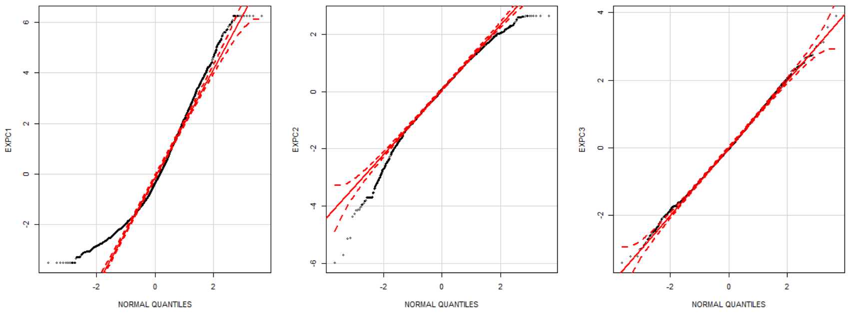
[13] Additionally, firms that introduced no innovations in the last 3 years were given the value of 0 for the dependent variables `InnoGoods`, `InnoServ`, and `RadInn`. I have provided descriptive statistics for the trimmed sample, i.e 3659 firms, above. The distributions for the `RadInn` response regression are not substantively different from those of the descriptive shown, so they have been omitted.

[14] In the ordered logit model, the cumulative probability curves, that is, the probability that the unit being analyzed falls into one of the ordered categories of the dependent variable, are assumed to be parallel. The Brant (1990) test is used to test if any variables violate this. In this case, several variables were in violation, thus an alternative model was needed. Robustness checks using an ordered logit model are nonetheless presented below.

fit between the analytical model and empirical observations (Young, 1981), by relaxing the assumption that the measurement scale of the variable is fixed. Using this, we have rescaled the dependent variable `RadInnOS`, which is assumed to be interval level data, into an ordinal variable which minimizes the sum of squared residuals (For a more comprehensive discussion of the ALSOS method, see Young, 1981 and Jacoby, 1999). Once this is done the optimally scaled variable lends itself very well to standard OLS regression (See appendix, Table 8.1 and Figure 8.1 for variable transformations). Though ALSOS has limited application in the field of innovation and entrepreneurship studies (one example however, can be found in Laursen and Foss (2003) in conjunction with principal components analysis), it offers an extremely flexible solution set to dealing with qualitative variables on nominal and ordinal levels. As Jacoby (1991: 76) describes it: "(T)he ALSOS approach gives a very reasonable solution to the problem of regression with qualitative variables. Nominal and ordinal variables would simply be assigned values that result in the highest possible R^2 and still maintain either the categories (for nominal variables) or the ordering (for ordinal variables) of the original observation categories."

As stated above, Model 2 was carried out using OLS regressions using the AEGIS survey data, with principal components derived from the survey's reliance on external knowledge sources for entrepreneurial and business opportunities filling the role of the response variable. The numerical values and span of these components makes OLS regression a suitable alternative. As with Model 1, listwise deletion for missing values was employed due to the nature of missingness not overtly interfering with the analysis, as well as a lack of interpretability of survey responses where the respondent denoted that he or she did not know the answer to the question(s). Concerning the choice of model, I have used ordinary least squares (OLS). Upon viewing these centered and standardized principal component variables' quantile normal distributions (Figure 4.15), this choice seems appropriate at least in the sense that the standardized principal component variables are approximately normally distributed. Although, some conceptual mismatch may be present regarding the interpretation of an effect on a response variable based on a principal component, which is itself merely a mathematical transformation of a summated rating scale. The interpretation therefore is based more upon the existence of statistically significant relationships and not concerned with magnitude of change in the dependent variables.

Figure 4.15: Quantile comparison plots of PCA variables based on external knowledge source reliance



4.6 Results

4.6.1 Model 1 - External knowledge intensity as affecting innovative performance

Tables 4.16 conveys the results of the first model specification using `InnoGoods` as the response variable. The regression of innovative goods/sales on `Breadth` and `Depth` yields some interesting results: `Breadth` alone produces a statistically significant association at the 10% level (spec. I) while Depth alone produces an association at the 1% level. In specification III, both `Breadth` and `Depth` in linear terms are included in the model. It can be seen that `Depth` takes a significant[15] coefficient while `Breadth` does not. Adding a quadratic term to `Breadth` (spec. IV) but not to `Depth` produces significant coefficients for `Breadth`, `Breadth`^{2}, as well as `Depth` , which remains linearly significant. The coefficient for `Breadth`^{2} is negative. Trying the the regression again with a quadratic term for `Depth` but not for `Breadth` produces a non-significant coefficient for `Breadth`, and significant coefficients for `Depth` and `Depth`^{2}. When both terms are allowed quadratic terms, the linear terms become significant for `Breadth` and `Depth`, but only `Breadth`^{2} is statistically significant in specification VI. Normally one might stop there, however, thus far the specifications have shown a potentially unstable relationship between `Breadth` and `Depth`. Indeed, they are quite correlated as we showed earlier (Table 4.15), with a correlation coefficient of 0.56.

[15]In the context of the models, I refer to *statistical* significance of coefficients, and not their *economic* significance. That is to say, I do not readily analyze the magnitude of the coefficients and interpret this at this time.

Table 4.16: Model 1.1 - `InnoGoods` regressed on breadth and depth of external knowledge search

	I	II	III	IV	V	VI	VII	VIII
	FLogit	FLogit	FLogit	FLogit	FLogit	FLogit	FLogit(OP)	OLS
Breadth	$0.0245^{\dagger}$		0.0046	0.2241^{**}	0.0010	0.1935^{*}	1.2052	$0.1242^{\dagger}$
	(0.0130)		(0.0150)	(0.0787)	(0.0151)	(0.0819)	(2.5311)	(0.0644)
Breadth^2				-0.0163^{**}		-0.0142^{*}	-5.4874^{*}	$-0.0091^{\dagger}$
				(0.0057)		(0.0059)	(2.2870)	(0.0048)
Depth		0.0458^{**}	0.0433^{**}	0.0469^{**}	0.1347^{**}	0.1073^{*}	6.8212^{**}	0.1082^{*}
		(0.0141)	(0.0164)	(0.0165)	(0.0473)	(0.0485)	(2.3107)	(0.0442)
Depth^2					-0.0099^{*}	-0.0066	-2.6997	-0.0066
					(0.0048)	(0.0050)	(2.0382)	(0.0047)
FirmAge	-0.0337^{*}	-0.0328^{*}	-0.0328^{*}	-0.0328^{*}	-0.0323^{*}	-0.0325^{*}	-0.0325^{*}	-0.0306^{*}
	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0129)
log(Emp)	0.0854^{**}	0.0925^{**}	0.0912^{**}	0.0902^{**}	0.0909^{**}	0.0902^{**}	0.0902^{**}	0.1108^{***}
	(0.0304)	(0.0301)	(0.0304)	(0.0304)	(0.0304)	(0.0304)	(0.0304)	(0.0282)
logit(IntlSales)	0.0720^{***}	0.0714^{***}	0.0714^{***}	0.0707^{***}	0.0713^{***}	0.0707^{***}	0.0707^{***}	0.0738^{***}
	(0.0156)	(0.0156)	(0.0156)	(0.0156)	(0.0156)	(0.0156)	(0.0156)	(0.0157)
logit(R&DInt)	0.2575^{***}	0.2527^{***}	0.2523^{***}	0.2529^{***}	0.2516^{***}	0.2523^{***}	31.6556^{***}	0.2907^{***}
	(0.0234)	(0.0234)	(0.0234)	(0.0234)	(0.0234)	(0.0234)	(1.9366)	(0.0279)
logit(R&DInt)^2	-0.0402^{***}	-0.0405^{***}	-0.0403^{***}	-0.0391^{***}	-0.0400^{***}	-0.0391^{***}	-8.8888^{***}	-0.0267^{***}
	(0.0074)	(0.0074)	(0.0074)	(0.0074)	(0.0074)	(0.0074)	(1.6884)	(0.0078)
(Intercept)	-0.2259	-0.2564	-0.2797	-0.9293^{**}	$-0.4162^{\dagger}$	-0.9375^{**}	-1.0020^{***}	-1.1536^{***}
	(0.2387)	(0.2263)	(0.2389)	(0.3328)	(0.2481)	(0.3328)	(0.2153)	(0.2811)
N	3673	3673	3673	3673	3673	3673	3673	3673
Sector Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3673	3673	3673	3673	3673	3673	3673	3673
Null Deviance	1713.2	1713.2	1713.2	1713.2	1713.2	1713.1	1713.1	
Residual Deviance	1443.8	1440.9	1440.9	1437.3	1439.1	1436.5	1436.5	
Dispersion (ϕ)	0.4281	0.4249	0.4252	0.4255	0.4251	0.4256	0.4256	
McFadden's R^2	0.1572	0.1589	0.1589	0.1610	0.1599	0.1615	0.1615	
Pearson's R^2								0.1758
adj. Pearson's R^2								0.1724
Resid. sd								1.6678

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Based on the relationships observed thus far in the model, I employ orthogonal polynomials in specification VII in order to remove the moderate to high correlation between these variables, and see if the significance levels and direction of effect differed. Now, one can see the significance of `Breadth` drops below the 10% level, while $\texttt{Breadth}^2$ retains its significance. `Depth` is positive and significant, but $\texttt{Depth}^2$ is insignificant in this specification. Column VIII in table 4.19 shows the same specification as IV, while instead relying on a log-odds (logit) transformation of the response variable and running the model using ordinary least squares estimation. Here we see 10% significance for positive `Breadth` and negative $\texttt{Breadth}^2$ coefficients, while `Depth` is positive and significant at the 5% level. As there does seem to be a bit of a multi-collinearity issue with `Breadth` and `Depth`, I will select model VII for further diagnostics and explanatory power/hypothesis testing.

4.6.1.1 Interpretation of Model 1.1

For this first case, a curvilinear relationship is indeed found between `Breadth` and the response, however, the linear component of the orthogonal polynomial is not found to be significant. One interpretation might be that while too much breadth is indeed producing a negative marginal relationship with the response variable, just having *some* breadth is not necessarily a good thing after all for our sampled firms (at least when one looks at the orthogonal polynomials, which is advisable in this case due to the instability of results across the first 4 specifications). Graphical interpretation of the effects plots lends some credence to this[16]. The first plot in Figure 4.15 shows that at the 0-values of the x-axis, that is, very small levels of breadth, the range of the 95% confidence bounds is extremely wide. This shows that just having a few sources is not conclusive in its effect on innovative performance. As the number of sources grow however, the bounds of the interval tighten, so that somewhere around the 4th source of knowledge we have a positive relationship between the explanatory variable `Breadth` and the response variable `InnoGoods`. However, at just a few additional sources of knowledge beyond this point, the declining marginality sets in, suggesting that these potentially KIE firms begin to draw innovative gains from breadth conclusively at on average 4 sources, but more than 6-7 proves to be too much for them to adequately manage for full benefit. While the

[16] In plotting my effects in all models, I make use of the effects package in R, authored by John Fox (2003: 4). It makes use of plotting fitted values for each interaction term of the explanatory variable(s): "The lower-order 'relatives' of a high-order term are absorbed into the term, allowing the predictors appearing in the high-order term to range over their values. The values of the other predictors are fixed at typical values."

`Breadth` variable indicates a potential inverse quadratic (though not necessarily linear) relationship, `Depth` produces some different effects. While the linear component of the polynomial is significant and positive, the quadratic term is not, suggesting a simpler linear relationship for these firms between depth of search and innovative performance of goods. The `Depth` plot in Figure 4.16 lends a bit more explanatory power. Up until about 6 sources of depth, the gains are net positive, beyond this though, the bounds widen and performance could move in either direction. The interplay between the `Depth` and `Breadth` variables throughout all 6 models makes it difficult to assess exactly what the relationship might be. The sectoral controls produce somewhat expected results, with clear differences emerging in a hierarchical sense: HTMS and LTMS being the most innovative in terms of goods, followed by KIBS and lastly by OBS. All in all, the control variables seem relatively stable throughout all 6 specifications.

Table 4.17 shows the next set of specifications for `InnoServ`, and it can be seen that regressing sale of innovative services divided by total sales yield quite different results. The first few specifications seem largely similar to `InnoGoods`' regressions. `Depth` is statistically significant but not `Breadth` in I, II, and III, while `Breadth`, `Breadth`^{2}, and `Depth` are all statistically significant in IV. In V none are significant, while in VI, only the linear and quadratic terms of `Breadth`. Orthogonal polynomials seem to stabilize the model somewhat, returning statistical significance to the quadratic of breath, and the linear term of `Depth` , as in the `InnoGoods` models. The logit-transformed response variable OLS specification, VIII, yields statistical significance for both the linear and quadratic components of `Breadth`, but none for `Depth` . Again I focus on specification VII for interpretation.

Figure 4.16: Effects plots for 1.1

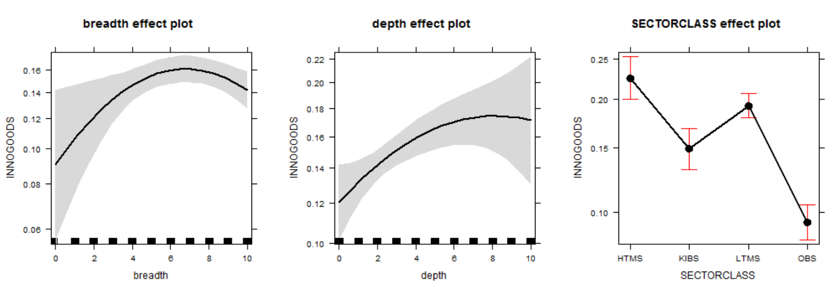
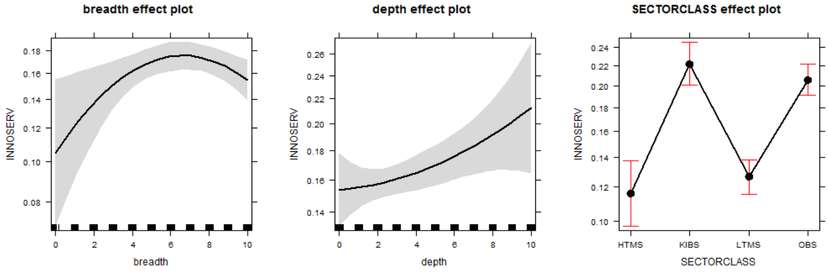


Table 4.17: Model 1.2 - `InnoServ` regressed on breadth and depth of external knowledge search

	I	II	III	IV	V	VI	VII	VIII
	FLogit	FLogit	FLogit	FLogit	FLogit	FLogit	FLogit (OP)	OLS
Breadth	0.0160		0.0010	0.1649*	0.0009	0.1788*	0.4844	0.1903**
	(0.0124)		(0.0145)	(0.0699)	(0.0146)	(0.0730)	(2.4097)	(0.0662)
Breadth^2				−0.0125*		−0.0134*	−5.1847*	−0.0134**
				(0.0052)		(0.0054)	(2.0763)	(0.0050)
Depth		0.0336*	0.0330*	0.0364*	0.0342	0.0078	4.9284*	0.0119
		(0.0139)	(0.0162)	(0.0163)	(0.0448)	(0.0459)	(2.2440)	(0.0455)
Depth^2					−0.0001	0.0032	1.3093	0.0026
					(0.0046)	(0.0048)	(1.9590)	(0.0049)
FirmAge	−0.0239†	−0.0231†	−0.0231†	−0.0234†	−0.0231†	−0.0236†	−0.0236†	−0.0135
	(0.0135)	(0.0135)	(0.0135)	(0.0135)	(0.0135)	(0.0135)	(0.0135)	(0.0133)
log(Emp)	0.0197	0.0244	0.0241	0.0230	0.0241	0.0230	0.0230	0.0514†
	(0.0291)	(0.0289)	(0.0291)	(0.0292)	(0.0291)	(0.0292)	(0.0292)	(0.0290)
logit(IntlSales)	−0.0063	−0.0070	−0.0070	−0.0074	−0.0070	−0.0075	−0.0075	−0.0149
	(0.0163)	(0.0163)	(0.0163)	(0.0163)	(0.0163)	(0.0163)	(0.0163)	(0.0161)
logit(R&DInt)	0.1302***	0.1260***	0.1259***	0.1270***	0.1259***	0.1270***	19.8915***	0.1272***
	(0.0256)	(0.0257)	(0.0257)	(0.0258)	(0.0257)	(0.0258)	(1.8927)	(0.0295)
logit(R&DInt)^2	−0.0380***	−0.0383***	−0.0382***	−0.0372***	−0.0382***	−0.0372***	−8.4638***	−0.0433***
	(0.0076)	(0.0076)	(0.0076)	(0.0077)	(0.0076)	(0.0077)	(1.7400)	(0.0082)
(Intercept)	−1.3752***	−1.4124***	−1.4175***	−1.8826***	−1.4191***	−1.8791***	−1.8702***	−2.6217***
	(0.2419)	(0.2308)	(0.2429)	(0.3128)	(0.2495)	(0.3129)	(0.2195)	(0.2896)
Sector Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Region Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	3671	3671	3671	3671	3671	3671	3671	3671
Null Deviance	1618.6	1618.6	1618.6	1618.6	1618.6	1618.6	1618.6	
Residual Deviance	1489.1	1487.4	1487.4	1484.9	1487.4	1484.7	1484.7	
Dispersion (ϕ)	0.4187	0.4186	0.4186	0.4203	0.4187	0.4203	0.4203	
McFadden's R^2	0.0800	0.0811	0.0811	0.08260	0.0811	0.0827	0.0827	
Pearson's R^2								
adj. R^2								0.0876
Resid. sd								1.7160

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Figure 4.17: Effects plots for 1.2



4.6.1.2 Interpretation of Model 1.2:

Looking at specification VII in table 4.17, innovation in services seems to be driven by slightly different levels of breadth and depth, though the trend from the goods response variable is largely found here as well: A non-significant linear component and significant quadratic component for `Breadth`, with the exact opposite for `Depth`. The effects plot in Figure 4.17 above shows a potential exponential trend in the non-significant curvilinear component of `Depth`, but cannot be evaluated due to the lack of precision in the confidence bounds. Here we find much less significance in our controls, firm size and internationalization have little to no effect (the latter is not entirely surprising due to the common conception of services being locational commodities that must be consumed where they are administered). Surprising is the still incredibly strong effect of R&D Intensity for innovativeness among service-centered innovators, suggesting that it is an important component for this type of firm as well as traditional goods innovators.

Table 4.18 shows the next set of specifications[17]: The variable `RadInnOS`, when used as the response, shows statistical significance for `Breadth` and `Depth` in specification I, in II, when the quadratic term of `Breadth` is added, all 3 terms are significant, with `Breadth` squared being a negative coefficient. III does not produce this effect when `Depth` is squared, rather only the linear associations are present. IV shows a inverse quadratic association for `Breadth`, with no significance for `Depth` at all. The orthogonal model, V, retains the same `Breadth` shape as IV, but adds a statistically significant coefficient for `Depth` . I will focus on this specification in my interpretation.

[17]In this model I do not regress `Breadth` and `Depth` separately before combining them. I performed this during robustness checks, and both variables retained the same approximate coefficients as seen here in specification I. Both were significant at the $p<.001$ level.

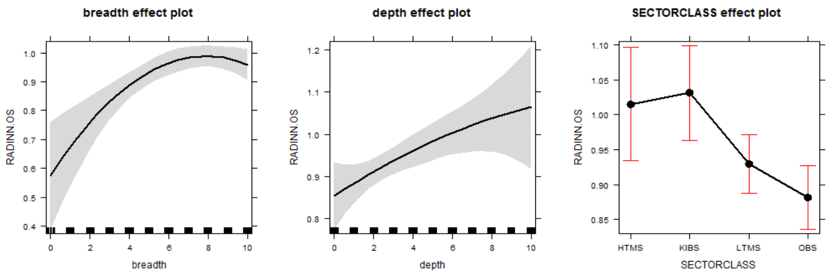
Table 4.18: Model 1.3 - `RadInn` regressed on breadth and depth of external knowledge search

	I ALSOS OLS	II ALSOS OLS	III ALSOS OLS	IV ALSOS OLS	V ALSOS OLS(OP)	VI O-LOGIT
Breadth	0.0192**	0.1091***	0.0182**	0.1049***	2.8962**	0.2562***
	(0.0065)	(0.0294)	(0.0066)	(0.0308)	(1.0635)	(0.0739)
Breadth^2		−0.0070**		−0.0067**	−2.5655**	−0.0161**
		(0.0022)		(0.0023)	(0.8907)	(0.0055)
Depth	0.0207**	0.0221**	0.0456*	0.0308	3.0698**	0.0978*
	(0.0075)	(0.0075)	(0.0202)	(0.0208)	(1.0310)	(0.0468)
Depth^2			−0.0028	−0.0010	−0.3978	−0.0064
			(0.0021)	(0.0022)	(0.8885)	(0.0048)
FirmAge	−0.0018	−0.0020	−0.0016	−0.0020	−0.0020	−0.0107
	(0.0063)	(0.0062)	(0.0063)	(0.0062)	(0.0062)	(0.0142)
log(Emp)	0.0819***	0.0814***	0.0819***	0.0815***	0.0815***	0.1686***
	(0.0136)	(0.0136)	(0.0136)	(0.0136)	(0.0136)	(0.0306)
logit(IntlSales)	0.0365***	0.0366***	0.0366***	0.0366***	0.0366***	0.0816***
	(0.0076)	(0.0076)	(0.0076)	(0.0076)	(0.0076)	(0.0173)
logit(R&DInt)	0.0725***	0.0732***	0.0723***	0.0730***	11.8961***	0.1272***
	(0.0135)	(0.0135)	(0.0135)	(0.0135)	(0.9112)	(0.0298)
logit(R&DInt)^2	−0.0240***	−0.0234***	−0.0239***	−0.0234***	−5.3276***	−0.0592***
	(0.0038)	(0.0038)	(0.0038)	(0.0038)	(0.8568)	(0.0085)
(Intercept)	1.0432***	0.7973***	1.0071***	0.7951***	0.8992***	
	(0.1103)	(0.1352)	(0.1136)	(0.1353)	(0.0992)	
Sector Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Region Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	3910	3910	3910	3910	3910	3910
R^2	0.1315	0.1337	0.1319	0.1338	0.1338	
adj. R^2	0.1286	0.1306	0.1288	0.1304	0.1304	
Resid. sd	0.8336	0.8327	0.8335	0.8328	0.8328	
McFadden's R^2						0.0539

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Though the parallel regression assumptions of the model were seen to be violated using the Brant test on the model, specification VI applies the ordinal logit method for comparison and robustness checking purposes. It yields similar directions and coefficients to that of the ALSOS model V with orthogonal polynomials included, and produces standard errors and coefficients roughly 2.5x the size of IV, though of course the log-odds interpretation renders a wholly different meaning to the numbers. The results of the Brant test may be found in the appendix (Table 8.2) and it shows that almost all the control variables used in the model violate the parallel regression assumptions of the ordinal logit model. Although the results of the ordinal model largely mirror the ALSOS models in terms of direction and significance of coefficients, failing this test implies that the model is not appropriate, and thus I will not rely on this specification for further analysis. Additionally, using OLS greatly facilitates diagnostics and effects plotting. I will rely on specification V, which makes use of orthogonal polynomials, for subsequent interpretation of results.

Figure 4.18: Effects plots for 1.3



4.6.1.3 Interpretation of 1.3

What then can be said of the relationship between radicalness of innovations and breadth and depth of external search? In this case, we find the hypothesized relationships hold for breadth, with statistically significant positive (negative) coefficients for the linear (quadratic) terms of the variable. Looking at the collection of effects plots in Figure 4.18, the confidence bounds are admittedly still rather large in the effect plot below for lower values of breadth. This suggests that having just a few sources does not necessarily affect a company's radical innovativeness. Also interesting is that the bend in the curve does not come until quite many sources, up around 8. This may indicate a certain resiliency on the part of KIE firms in terms of how drive to radically innovate might be associated with extensive search. For `Depth`, we again see the non-significant quadratic term while the linear term remains significant. The control variables take on similar shapes in the effects plots to those of the Model 1.1 plots on goods innovation, suggesting that these firms may steal the thunder of the service oriented firms in this case, since the dependent variable does not specify radicalness in goods or services, but rather combines the two. This is unfortunate, but such are the drawbacks of utilizing external survey data for analysis. The situation is not so dire though, as by glancing at the `SectorClass` effect plot, one may see that knowledge intensive business services are roughly similarly distributed on the response variable as high- and mid-tech manufacturers. So, we may not attribute all the explained variance to that of the goods-focused innovators.

4.6.1.4 Interpretation of Models 1.4 – 1.6

Table 4.19 - 4.21 convey the next series of regressions using principal components. Applying principal components regression to the innovative performance response variables is a bit more elucidating than looking solely at the levels of breadth and depth and their influence. As a reminder to the reader, the principal components used in the regression are as follows:

> `EXPC1`: External, non-industry sources of knowledge (or specialized knowledge providers (Tether and Tajar, 2008), mainly related to the collaboration with state, national, or regional research-based or academic entities.
>
> `EXPC2`: Business and operations-based relationships (made up of clients, customers, suppliers and competitors)
>
> `EXPC3`: Sources of codified knowledge stemming directly from academia and related communities.

Looking at Model 1.4, spec. I-VI (Table 4.19. Effects shown in Figure 4.19), where `InnoGoods` is the response variable, we have confirmed hypotheses. `EXPC1`, `EXPC2` and `EXPC3` are all positive and statistically significant in terms of their relationship with `InnoGoods`. Again, international sales and firm size seem to play a role in the regression, while again we see a curvilinear effect from R&D Intensity. No curvilinear relationships were detected between the components and the response variable.

Figure 4.19: Effects plots for 1.4

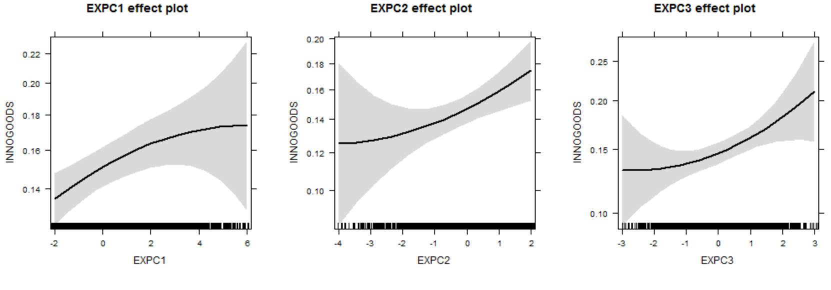


Table 4.19: Model 1.4: `InnoGoods` regressed on principal components regression of external knowledge

	I	II	III	IV	V	VI	VII
	FLogit	FLogit	FLogit	FLogit	FLogit	FLogit(OP)	OLS
EXPC1	0.0415**			0.0454**	0.0563**	6.2491**	0.0548**
	(0.0159)			(0.0161)	(0.0202)	(2.0791)	(0.0178)
EXPC1^2					−0.0047	−1.4267	−0.0034
					(0.0067)	(2.0130)	(0.0063)
EXPC2		0.0839**		0.0837**	0.0837**	5.7371**	0.0767**
		(0.0270)		(0.0270)	(0.0280)	(2.0189)	(0.0262)
EXPC2^2					0.0094	1.2764	0.0080
					(0.0147)	(2.0067)	(0.0132)
EXPC3			0.1037***	0.1033***	0.0945**	6.0914**	0.0931***
			(0.0300)	(0.0300)	(0.0305)	(1.9101)	(0.0285)
EXPC3^2					0.0181	1.5746	0.0076
					(0.0210)	(1.8284)	(0.0200)
Firm_age	−0.0335*	−0.0325*	−0.0331*	−0.0322*	−0.0322*	−0.0322*	−0.0298*
	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0129)
log(Emp)	0.0877**	0.0944**	0.0963**	0.0927**	0.0912**	0.0912**	0.1110***
	(0.0302)	(0.0301)	(0.0302)	(0.0302)	(0.0303)	(0.0303)	(0.0281)
logit(IntlSales)	0.0718***	0.0722***	0.0687***	0.0676***	0.0680***	0.0680***	0.0716***
	(0.0156)	(0.0156)	(0.0156)	(0.0157)	(0.0157)	(0.0157)	(0.0157)
logit(R&DInt)	0.2524***	0.2704***	0.2595***	0.2569***	0.2552***	32.0915***	0.2931***
	(0.0235)	(0.0234)	(0.0232)	(0.0237)	(0.0237)	(1.9552)	(0.0280)
logit(R&DInt)^2	−0.0403***	−0.0411***	−0.0424***	−0.0400***	−0.0399***	−9.0687***	−0.0272***
	(0.0074)	(0.0074)	(0.0074)	(0.0074)	(0.0074)	(1.6885)	0.0078
(Intercept)	−0.0837	−0.0442	−0.0833	−0.1439	−0.1591	−1.0342***	−0.5706**
	(0.2170)	(0.2162)	(0.2168)	(0.2173)	(0.2214)	(0.2154)	(0.2123)
Sector Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Region Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	3673	3673	3673	3673	3673	3673	3673
Null Deviance	1713.2	1713.2	1713.2	1713.2	1713.2	1713.2	
Residual Dev.	1442.5	1441.2	1440.3	1433.0	1432.3	1432.3	
Dispersion (ϕ)	0.4264	0.4266	0.4267	0.4241	0.4247	0.4247	
McFadden's R^2	0.1580	0.1588	0.0827	0.1636	0.1696	0.1696	
Pearson's R^2							0.1783
adj. R^2							0.1745
Resid. sd							1.666

Standard errors in parentheses
† significant at $p < .10$; $^*p < .05$; $^{**}p < .01$; $^{***}p < .001$

Figure 4.20: Effects plots for 1.5

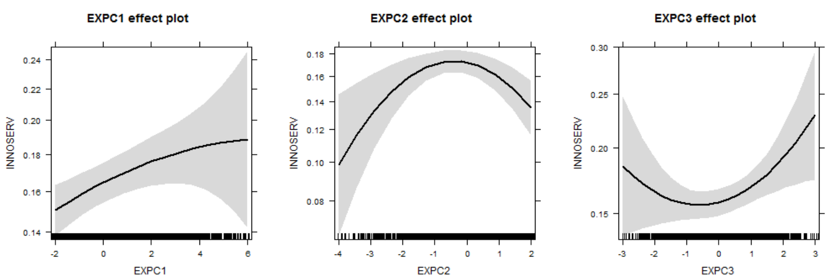


Table 4.20: Model 1.5: `InnoServ` regressed on principal components regression of external knowledge

	I FLogit	II FLogit	III FLogit	IV FLogit	V FLogit	VI FLogit(OP)	VII OLS
EXPC1	0.0427**			0.0433**	0.0465*	5.2971**	0.0601**
	(0.0155)			(0.0155)	(0.0190)	(1.9949)	(0.0183)
EXPC1^2					−0.0032	−0.9496	−0.0060
					(0.0064)	(1.9168)	(0.0065)
EXPC2		−0.0061		−0.0056	−0.0418	−0.6424	−0.0320
		(0.0256)		(0.0256)	(0.0288)	(2.0182)	(0.0270)
EXPC2^2					−0.0506**	−6.9066**	−0.0490***
					(0.0158)	(2.1504)	(0.0136)
EXPC3			0.0406	0.0427	0.0451	3.1226†	0.0542†
			(0.0287)	(0.0288)	(0.0290)	(1.8210)	(0.0292)
EXPC3^2					0.0366†	3.1886†	0.0353†
					(0.0198)	(1.7241)	(0.0204)
Firm_age	−0.0234†	−0.0240†	−0.0236†	−0.0232†	−0.0233†	−0.0233†	−0.0128
	(0.0135)	(0.0135)	(0.0135)	(0.0135)	(0.0135)	(0.0135)	(0.0132)
log(Emp)	0.0196	0.0239	0.0247	0.0204	0.0193	0.0193	0.0494†
	(0.0289)	(0.0289)	(0.0289)	(0.0289)	(0.0291)	(0.0291)	(0.0289)
logit(IntlSales)	−0.0070	−0.0057	−0.0067	−0.0080	−0.0075	−0.0075	−0.0146
	(0.0162)	(0.0162)	(0.0163)	(0.0163)	(0.0163)	(0.0163)	(0.0161)
logit(R&DInt)	0.1224***	0.1331***	0.1325***	0.1205***	0.1240***	19.2591***	0.1247***
	(0.0258)	(0.0256)	(0.0255)	(0.0259)	(0.0260)	(1.9072)	(0.0296)
logit(R&DInt)^2	−0.0376***	−0.0393***	−0.0394***	−0.0380***	−0.0356***	−8.1039***	−0.0417***
	(0.0076)	(0.0076)	(0.0076)	(0.0076)	(0.0077)	(1.7425)	(0.0082)
(Intercept)	−1.3047***	−1.2470***	−1.2665***	−1.3257***	−1.2926***	−1.8895***	−1.9437***
	(0.2211)	(0.2201)	(0.2206)	(0.2216)	(0.2253)	(0.2191)	(0.2188)
Sector Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Region Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	3671	3671	3671	3671	3671	3671	3671
Null Deviance	1618.6	1618.6	1618.6	1618.6	1618.6	1618.6	
Residual Dev.	1486.7	1489.8	1489.0	1485.8	1479.9	1479.9	
Dispersion (ϕ)	0.4179	0.4189	0.4189	0.4179	0.4180	0.4180	
McFadden's R^2	0.0815	0.0796	0.08	0.0820	0.0857	0.0857	
Pearson's R^2							0.0950
adj. R^2							0.0908
Resid. sd							1.7130

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

In Model 1.5, specification VI (Table 4.20. Effects shown in Figure 4.20), where the response variable is `InnoServ`, the pattern alters. We find 5% and 10% respective positive statistical significance levels for `EXPC1` and `EXPC3`, but not for `EXPC2`. Here the linear term is not significant. However, the quadratic term is, and is negative. This, when plotted, shows a relationship not unlike that of `Breadth` with `InnoServ`, showing no significant relationship at low levels of intra-industry knowledge reliance, and a slowly tightening confidence bound as levels increase. Eventually, the effect flips and marginal decline sets in. Hence, moderate levels of reliance on clients, customers, and suppliers seem to produce the highest levels of service innovations, while high levels of reliance seem to decrease in effectiveness. Worth noting is that surprisingly, the OLS specification (VII) seems to capture similar results to the orthogonal polynomial fractional logit (VI). To maintain consistence however, I will perform diagnostics on VI.

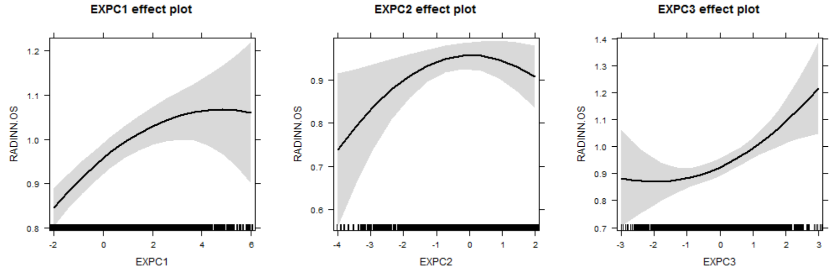
Table 4.21: Model 1.6: `RadInnOS` regressed on principal components regression of external knowledge

	I ALSOS OLS	II ALSOS OLS	III ALSOS OLS	IV ALSOS OLS	V ALSOS OLS	VI ALSOS OLS(OP)	VII O-LOGIT
EXPC1	0.0395***			0.0390***	0.0459***	4.9191***	0.1092***
	(0.0073)			(0.0073)	(0.0086)	(0.9257)	(0.0198)
EXPC1^2					−0.0048	−1.4447	−0.0169*
					(0.0029)	(0.8845)	(0.0066)
EXPC2		0.0164		0.0158	0.0019	0.7780	0.0229
		(0.0117)		(0.0116)	(0.0127)	(0.8819)	0.0289
EXPC2^2					−0.0131*	−1.7920*	−0.0231
					(0.0063)	(0.8589)	(0.0144)
EXPC3			0.0593***	0.0580***	0.0555***	3.6071***	0.0957**
			(0.0135)	(0.0135)	(0.0138)	(0.8690)	(0.0313)
EXPC3^2					0.0140	1.2202	0.0382†
					(0.0097)	(0.8454)	(0.0219)
Firm_age	−0.0020	−0.0018	−0.0018	−0.0015	−0.0015	−0.0015	−0.0093
	(0.0063)	(0.0063)	(0.0063)	(0.0062)	(0.0062)	(0.0062)	0.0142
log(Emp)	0.0829***	0.0869***	0.0889***	0.0849***	0.0832***	0.0832***	0.1723***
	(0.0135)	(0.0135)	(0.0135)	(0.0135)	(0.0135)	(0.0135)	0.0306
logit(IntlSales)	0.0367***	0.0377***	0.0362***	0.0353***	0.0356***	0.0356***	0.0800***
	(0.0076)	(0.0076)	(0.0076)	(0.0076)	(0.0076)	(0.0076)	(0.0173)
logit(R&DInt)	0.0711***	0.0821***	0.0793***	0.0712***	0.0712***	11.7771***	0.1254***
	(0.0135)	(0.0135)	(0.0134)	(0.0135)	(0.0135)	(0.9164)	(0.0300)
logit(R&DInt)^2	−0.0246***	−0.0260***	−0.0263***	−0.0246***	−0.0237***	−5.3805***	−0.0598***
	(0.0038)	(0.0038)	(0.0037)	(0.0037)	(0.0038)	(0.8549)	(0.0085)
(Intercept)	1.2460***	1.2879***	1.2628***	0.8705***	1.2386***	0.8751***	
	(0.1008)	(0.1009)	(0.1008)	(0.0990)	(0.1025)	(0.0990)	
Sector Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Region Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	3910	3910	3910	3910	3910	3910	3910
R^2	0.1308	0.1248	0.1286	0.1354	0.1374	0.1374	
adj. R^2	0.1282	0.1221	0.1259	0.1323	0.1336	0.1336	
Resid. sd	0.8338	0.8367	0.8349	0.8319	0.8312	0.8312	
McFadden's R^2							0.0437

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

For Model 1.6 (Table 4.21), where `RadInnOS` is regressed on the `EXPC` variables, the first 4 specifications point towards positive and significant coefficients for `EXPC1` and `EXPC3`, while polynomial models (V and VI) point towards the same, with the addition of a significant and negative quadratic term for `EXPC2`. These effects are clearly discernable in the effects plot below for specification VI. The orthogonal specification (VI) does not appear to change much, which is not surprising given that the components are per definition not correlated. In column VII one can see that when the regression is run as an ordered logit model, a slight inverse curvilinear relationships is produced for `EXPC2`, while an even weaker exponential curvilinear relationship for `EXPC3` is found, where both coefficients take positive significant values. Again, this model is only here for comparison, and as the ALSOS models seem to be more in line with what was found in Models 1.4 and 1.5 effect-wise, it reinforces the view that ordinal logit may not be appropriate.

Figure 4.21: Effects plots for 1.6



A Brant test for this regression specification may be found in the appendix (Table 8.2)

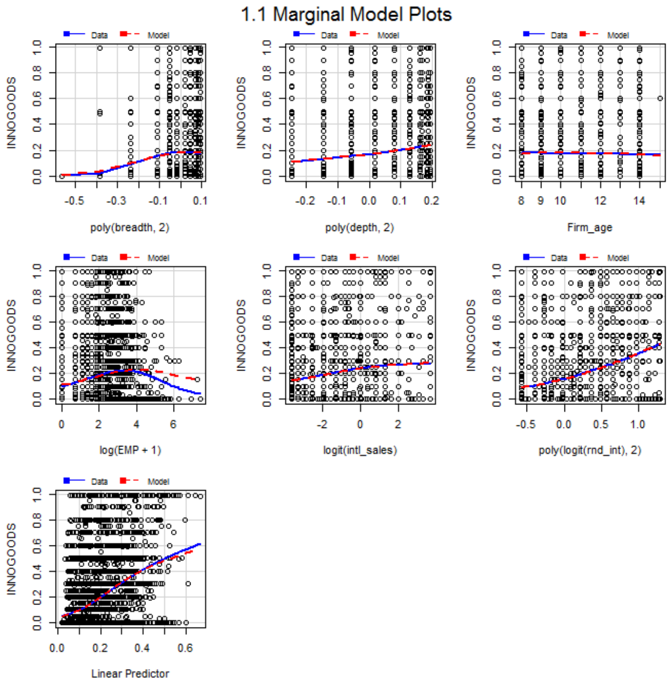
4.6.1.5 Summary of effects and hypothesis confirmation - Model 1

While search breadth does indeed appear to be curvilinear in all cases, it is not always apparent that it is linearly related to innovative performance. Evidence from the first specifications of Models 1.1 - 1.3 reveal that it does appear to be linearly associated with `RadInnOS`, and, conditional on the model also including a quadratic term, it is linearly related to both `InnoGoods` and `InnoServ`. So **H-1.1a is left partially confirmed**. Since the regression do yield quadratic associations in all 3 fully specified models, **I lean towards confirmation for H-1.1b.** Search depth appears to be linearly related to all constructs of innovative performance, so H-1.2a is confirmed, but H-1.2b is not confirmed since it is not curvilinear in any of the models, save for `InnoGoods` when the quadratic term for `Breadth` is not included. Looking at the effects of the coefficients across the models, `Depth` more often than not has the larger coefficient, while quadratic effects favor `Breadth` as the stronger predictor. Nonetheless, **H-1.3 is not confirmed**. In terms of certain types of external knowledge and how they affect innovative performance, **we find mixed confirmation in both WH-1.4 and WH-1.5**. Greater reliance on knowledge by specialist knowledge provider-like actors in the extra-industry environment, as well as codified scientific and technological knowledge, both positively relate to innovative performance in all 3 principal component-based regressions, though neither were quadratically associated. Intra-industry knowledge sources were positive-linearly related to goods and radicalness. However, we cannot detect a linear relationship between this variable and services, though we can detect an inverse curvilinear effect ($p<.01$).

4.6.1.6 Diagnostics Model 1

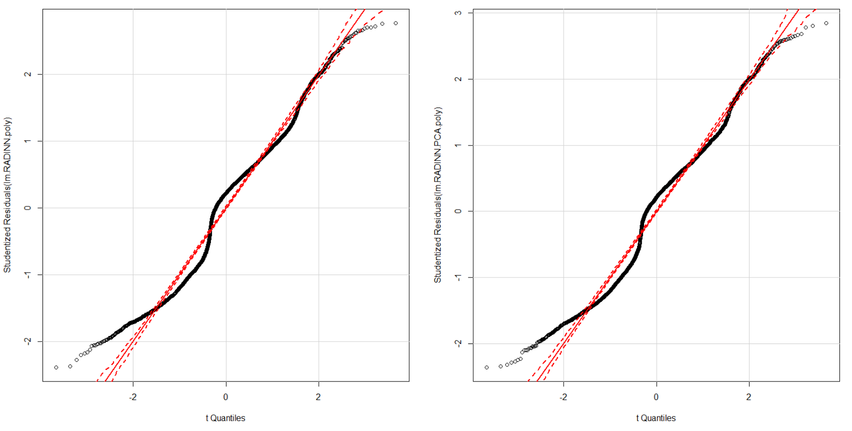
To ensure that the fit of the model is adequate, for the fractional logit models on `InnoGoods` and `InnoServ`, I have employed marginal model plotting (Cook and Weisberg, 1997; Fox and Weisberg, 2011). These plots use a locally weighted scatterplot smoother (loess) curve to plot the values of the response variable (the solid blue line) as well as to plot the fitted values generated by the regression (the dashed red line) according to the conditional marginal effect provided by each of the individual predictor variables. Generally, an adequate fit between both loess lines indicates that the model does indeed fit the data well. The only discernable non-fit is present in the logarithmic Emp variable, denoting firm size, which deviates at higher values. One can also observe some deviation from the fitted model in the tail of some of the principal components. This is however very likely due to the distribution of the variable values, and not enough of a real deviation to worry about model fit. I include only the plots for Model 1.1 here (Figure 4.22), the rest may be found in the appendix in Figures 8.2 - 8.4.

Figure 4.22: Marginal model plots of 1.1



The reader will recall that the `RadInnOS` variable, denoting radicalness of innovation, was originally an ordinal variable, but was rescaled to be numeric using the alternating least squares optimal scaling method. The question is whether this transformation indeed makes the variable a viable candidate for an OLS regression model. Producing a quantile comparison plot of the studentized residuals of the models *vis á vis* the quantiles of the normal distribution model itself can alleviate this worry: The model's fitted value distribution follows quite closely that of the normal quantiles, denoting a good fit to an OLS regressive model:

Figure 4.23: Normal quantiles vs. studentized residuals of `RadInnOS` models 1.3 and 1.6



Additionally, I performed variance inflation factor tests on all models, returning no significant problems in any of the 6 tests[18], leading to a moderate degree of certainty that multi-collinearity is not a problem in any of the above regressions. The VIF tests are available in the appendices, Tables 8.4 and 8.5. As stated earlier, orthogonal polynomials are used to avoid collinearity issues in quadratic terms. For now, I leave this set of regressions to focus on the next set of models.

4.6.2 Model 2 - Internal knowledge intensity as affecting external knowledge intensity

In this set of regressions, I map the predictive power and association of variables signifying the human capital of founders, employees, and the firm

[18] All values are quite close to 1, indicating that the confidence interval of the variables would not drastically be affected by conditions of orthogonality among the predictor variables

at its inception, on that of external knowledge source reliance. Namely, I model the effect of different predictors on the principal components of external knowledge source reliance derived previously in the chapter. The dependent variables in these models are the following:

> `EXPC1`: External, non-industry sources of knowledge (or specialized knowledge providers (Tether and Tajar, 2008), mainly related to the collaboration with state, national, or regional research-based or academic entities.
>
> `EXPC2`: Business and operations-based relationships (made up of clients, customers, suppliers and competitors)
>
> `EXPC3`: Sources of codified knowledge stemming directly from academia and related communities.

The first step was to select an adequate set of regressions by way of their focal predictors. I regressed all three dependent variables on the founder variables, varying only in the variables representing functional heterogeneity: Blau's index, Attneave's entropy, and Cantner et al.'s knowledge scope and knowledge disparity. While results were quite similar across the board, I relied on the R squared of the model to select which variables to retain. The knowledge scope and disparity models retained the highest explanatory power in 2 of 3 instances, and the other one was a marginal difference (Δ=.002). Thus I opted for the conceptualization of functional background heterogeneity applied by Cantner et al.(2005) in these specifications. I modify the model specification gradually by introducing clusters of variables based on the operational categories of the variables: Education levels within the venture (founders, employees); Experience of founders (Entrepreneurial, industry, academic, other) as well as functional heterogeneity in background of founders; Formal organizational origin (Spinoff or not); and factors expressed as important for firm formation (`FF1-FF3`).

Even though these latter variables represent different concepts, they are per-definition orthogonal with one another so there is no harm done by including them together in a specification. The final column (VI) utilizes orthogonal polynomials for the R&D Intensity controls, but is otherwise identical to model V. Tables 4.22 – 4.24 show the results.

4.6.2.1 Interpretation of Model 2

It is worth restating the Model 2 hypotheses before continuing to aid interpretation and to help the reader follow the results:

> WH-2.1: Higher levels of stocks of human capital in the form of education will positively influence the reliance on external knowledge of the firm
>
> WH-2.2a: Founders having higher previous work experience at a university or research institute will be positively associated with the reliance on knowledge stemming from non-industry sources of knowledge such as specialist knowledge providers.
>
> WH-2.2b: Founders having higher previous work experience in the same industry or having entrepreneurial experience will be negatively associated with the reliance on knowledge stemming from non-industry sources of knowledge such as specialist knowledge providers.
>
> WH-2.3a: Founders' previous work experience at a university/research institute will be negatively associated with the reliance on intra-industry sources of knowledge
>
> WH-2.3b: Founders' previous work experience in the same industry, or previous entrepreneurial experience, will positively associated with the reliance on intra-industry sources of knowledge
>
> WH-2.4a: In terms of functional heterogeneity of the founding team, the level of knowledge scope should positively associated with the level of reliance on external knowledge from all categories.
>
> WH-2.4b: The level of knowledge disparity should negatively associate with the level of reliance on external knowledge from all categories.
>
> WH-2.5: For entrepreneurial firms, stemming from a previous organization (spinoff) will positively impact the reliance on knowledge stemming from all types of external knowledge sources.
>
> WH-2.6: The extent to which a firm's formation was based on novel opportunities should positively influence the reliance on knowledge stemming from all types of external knowledge sources.

Specification I in Table 4.22 shows a mildly statistically significant relationship between education level of employees (being a tertiary degree or higher) and reliance on the *specialist knowledge providers* represented by `EXPC1`. The second specification shows that entrepreneurial experience negatively associates, while university/academic experience positively associates with `EXPC1`. Also, older founding teams associate with higher

Table 4.22: Model 2.1: `EXPC1` regressed on variables of internal knowledge intensity

	I OLS	II OLS	III OLS	IV OLS	V OLS	VI OLS(OP)
EmpEdu	0.3495†				0.4170*	0.4170*
	(0.1880)				(0.1782)	(0.1782)
EmpHiEdu	−0.1942				−0.2916	−0.2916
	(0.2214)				(0.2083)	(0.2083)
FoundEdu	0.0067				0.0263	0.0263
	(0.0348)				(0.0335)	(0.0335)
FoundEnt		−0.2131**			−0.1244†	−0.1244†
		(0.0651)			(0.0648)	(0.0648)
FoundInd		0.0004			−0.0039	−0.0039
		(0.0032)			(0.0033)	(0.0033)
FoundUni		0.9345***			0.9362***	0.9362***
		(0.1849)			(0.1838)	(0.1838)
AgeMax		0.0967*			0.0651	0.0651
		(0.0422)			(0.0416)	(0.0416)
KScope		0.1444			−0.1069	−0.1069
		(0.1490)			(0.1492)	(0.1492)
KDisp		0.0053			0.0820	0.0820
		(0.1851)			(0.1831)	(0.1831)
Spinoff			0.0528		0.0324	0.0324
			(0.0859)		(0.0867)	(0.0867)
FF1 (Opps)				0.4368***	0.4287***	0.4287***
				(0.0175)	(0.0189)	(0.0189)
FF2 (Exp/Net)				−0.1338***	−0.1188***	−0.1188***
				(0.0233)	(0.0257)	(0.0257)
FF3 (Spec)				−0.0107	−0.0175	−0.0175
				(0.0256)	(0.0271)	(0.0271)
Firm_age	−0.0128	−0.0154	−0.0011	−0.0129	−0.0304*	−0.0304*
	(0.0144)	(0.0140)	(0.0137)	(0.0129)	(0.0138)	(0.0138)
log(Emp)	0.0846**	0.0925**	0.1004***	0.0310	0.0113	0.0113
	(0.0321)	(0.0302)	(0.0297)	(0.0282)	(0.0309)	(0.0309)
logit(IntlSales)	0.0266	0.0199	0.0246	0.0300†	0.0283†	0.0283†
	(0.0175)	(0.0168)	(0.0166)	(0.0156)	(0.0165)	(0.0165)
logit(R&DInt)	0.2397***	0.2159***	0.2411***	0.1600***	0.1448***	21.4324***
	(0.0306)	(0.0300)	(0.0294)	(0.0279)	(0.0293)	(2.0214)
logit(R&DInt)^2	−0.0409***	−0.0422***	−0.0391***	−0.0333***	−0.0369***	−8.3961***
	(0.0085)	(0.0083)	(0.0082)	(0.0077)	(0.0081)	(1.8410)
(Intercept)	1.1821***	0.9580***	1.1157***	1.1289***	1.0520***	0.4428†
	(0.2543)	(0.2594)	(0.2205)	(0.2084)	(0.2718)	(0.2671)
Sector Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Region Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	3507	3782	3910	3749	3331	3331
R^2	0.1692	0.1755	0.1663	0.2925	0.2988	0.2988
adj. R^2	0.1659	0.1718	0.1638	0.2899	0.2937	0.2937
Resid. sd	1.8148	1.8096	1.8289	1.6848	1.6661	1.6661

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

reliance on `EXPC1` knowledge sources. In specification III, spinoffs do not seem to differ from non-spinoffs. In IV, we can see that both opportunity based formation factors' importance (`FF1`) and experiential and network factors (`FF2`) are associated with `EXPC1`, the former positively and the latter negatively at the p<.001 level. Surprisingly , in this specification, the importance of technical and design knowledge for firm formation (`FF3`) has no statistically significant association with reliance on knowledge stemming from specialist knowledge providers. Specification V and VI show that when all variables are present in the model, the positive association of employee education (`EmpEdu`) becomes more defined at the p<.05 level, while the negative association of founders having entrepreneurial experience weakens from the 1% to the 10% level. The latter may be due to the fact that the `FF2` variable in part denotes the extent to which firm founding was based on experiential knowledge, and including these two in the same model results in the `FF2` variable taking some of the significance away from `FoundEnt`. The strongest positive predictor seems to be whether or not the founding team has some type of academic experience.

Proceeding to the second response variable, Table 4.23 models the importance of knowledge stemming from business and operations based relationships (`EXPC2`) regressed on the Internal Knowledge Intensity variables. In this set of models, a high degree of disturbance was detected during trial regressions between `EmpEdu`, `EmpHiEdu`; and `FoundEdu`, with the presence of all 3 variables resulting in obscured effects on the part of `EmpEdu` especially. This is likely due to the potentially problematic high correlation between the 3 variables. This makes intuitive sense, since the proportion of Master and PhD degrees is a subset of the proportion of Tertiary degrees (one can almost never acquire the former without obtaining the latter), and the fact that the small average size of the firm makes Founder and employee education likely somewhat convergent, as well as the high potential of founders counting their own education among that of the employee average.

For this reason, in Model 2.2 I omit `FoundEdu`, and avoid having `EmpEdu` and `EmpHiEdu` in the same specification[19]. In the two columns, one can see that there is a negative association between the proportion of tertiary educated employees, and of Master and PhD educated employees, with the response variable. The third column yields negative associations with founders having university experience (p<.001) and age of the founding team (p<.01), and a positive association with knowledge scope of the founding team (p<.10). The fourth column yields a negative association between the venture coming from a pre-existing organization (Spinoff) and the response variable at the 1% level. Column V shows that opportunity-based and experience and network-based formation factors are positively associated with the response variable, while technical and design knowledge are negatively associated (all p<.001).

[19] After extensive testing, this model, and the survival model for 3.3, were the only two in which this relationship proved problematic, therefore all other models retain all 3 education variables

Table 4.23: Model 2.2: `EXPC2` regressed on variables of internal knowledge intensity

	I OLS	II OLS	III OLS	IV OLS	V OLS	VI OLS	VII OLS	VIII OLS (OP)
EmpEdu	-0.3762^{***}					-0.3464^{***}		-0.3464^{***}
	(0.0794)					(0.0802)		(0.0802)
EmpHiEdu		-0.5754^{***}					-0.5322^{***}	
		(0.0929)					(0.0942)	
FoundEnt			0.0432			0.0522	0.0447	0.0522
			(0.0405)			(0.0412)	(0.0412)	(0.0412)
FoundUni			-0.4982^{***}			-0.3657^{**}	-0.3430^{**}	-0.3657^{**}
			(0.1154)			(0.1193)	(0.1209)	(0.1193)
AgeMax			-0.0697^{**}			-0.0946^{***}	-0.0952^{***}	-0.0946^{***}
			(0.0230)			(0.0234)	(0.0234)	(0.0234)
KScope			$0.1657^{\dagger}$			0.1076	0.1121	0.1076
			(0.0928)			(0.0942)	(0.0946)	(0.0942)
KDisp			0.1073			0.1561	0.1362	0.1561
			(0.1146)			(0.1160)	(0.1163)	(0.1160)
Spinoff				-0.1422^{**}		-0.1706^{**}	-0.1493^{**}	-0.1706^{**}
				(0.0538)		(0.0552)	(0.0552)	(0.0552)
FF1					0.0874^{***}	0.0877^{***}	0.0892^{***}	0.0877^{***}
					(0.0117)	(0.0119)	(0.0119)	(0.0119)
FF2					0.0649^{***}	0.0749^{***}	0.0703^{***}	0.0749^{***}
					(0.0156)	(0.0157)	(0.0158)	(0.0157)
FF3					-0.1033^{***}	-0.1114^{***}	-0.1096^{***}	-0.1114^{***}
					(0.0171)	(0.0173)	(0.0174)	(0.0173)
Firm_age	-0.0116	$-0.0146^{\dagger}$	-0.0108	$-0.0159^{\dagger}$	-0.0138	-0.0040	-0.0070	-0.0040
	(0.0087)	(0.0087)	(0.0087)	(0.0086)	(0.0086)	(0.0088)	(0.0088)	(0.0088)
log(Emp + 1)	0.0125	0.0001	0.0023	0.0040	$-0.0325^{\dagger}$	-0.0084	-0.0195	-0.0084
	(0.0191)	(0.0191)	(0.0187)	(0.0186)	(0.0188)	(0.0195)	(0.0196)	(0.0195)
logit(IntlSales)	-0.0023	0.0030	-0.0005	-0.0024	-0.0040	-0.0010	0.0028	-0.0010
	(0.0106)	(0.0106)	(0.0104)	(0.0104)	(0.0104)	(0.0106)	(0.0106)	(0.0106)
logit(R&DInt)	-0.0708^{***}	-0.0668^{***}	-0.0681^{***}	-0.0847^{***}	-0.0813^{***}	-0.0577^{**}	-0.0563^{**}	-3.7077^{**}
	(0.0189)	(0.0188)	(0.0188)	(0.0184)	(0.0187)	(0.0192)	(0.0192)	(1.2982)
logit(R&DInt)^2	-0.0110^{*}	$-0.0097^{\dagger}$	-0.0069	-0.0123^{*}	-0.0103^{*}	-0.0067	-0.0060	-1.5235
	(0.0052)	(0.0052)	(0.0052)	(0.0051)	(0.0052)	(0.0053)	(0.0053)	(1.1958)
Sectoral Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Regional Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
(Intercept)	0.1642	0.2237	0.1930	0.1298	0.2090	0.3454^{*}	0.4125^{*}	0.4293^{**}
	(0.1406)	(0.1410)	(0.1614)	(0.1379)	(0.1393)	(0.1640)	(0.1643)	(0.1616)
N	3795	3758	3857	3910	3749	3646	3607	3646
R^2	0.0420	0.0459	0.0458	0.0397	0.0657	0.0820	0.0843	0.0820
adj. R^2	0.0390	0.0428	0.0419	0.0367	0.0622	0.0767	0.0790	0.0767
Resid. sd	1.1404	1.1367	1.1385	1.1442	1.1262	1.1147	1.1129	1.1147

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 4.24: Model 2.3: `EXPC3` regressed on variables of internal knowledge intensity

	I	II	III	IV	V	VI
	OLS	OLS	OLS	OLS	OLS	OLS(OP)
EmpEdu	0.2120*				0.2527*	0.2527*
	(0.1027)				(0.1056)	(0.1056)
EmpHiEdu	−0.2520*				−0.2916*	−0.2916*
	(0.1209)				(0.1234)	(0.1234)
FoundEdu	0.0629***				0.0620**	0.0620**
	(0.0190)				(0.0198)	(0.0198)
FoundEnt		0.0761*			0.0947*	0.0947*
		(0.0355)			(0.0384)	(0.0384)
FoundInd		0.0003			0.0002	0.0002
		(0.0017)			(0.0020)	(0.0020)
FoundUni		−0.0363			−0.1190	−0.1190
		(0.1009)			(0.1089)	(0.1089)
AgeMax		−0.0403†			−0.0412†	−0.0412†
		(0.0230)			(0.0246)	(0.0246)
KScope		0.0456			0.0347	0.0347
		(0.0813)			(0.0884)	(0.0884)
KDisp		0.1196			0.1581	0.1581
		(0.1010)			(0.1084)	(0.1084)
Spinoff			−0.0241		−0.0224	−0.0224
			(0.0464)		(0.0513)	(0.0513)
FF1 (Opps)				−0.0039	0.0040	0.0040
				(0.0103)	(0.0112)	(0.0112)
FF2 (Exp/Net)				0.0201	0.0194	0.0194
				(0.0136)	(0.0152)	(0.0152)
FF3 (Spec)				0.0169	0.0192	0.0192
				(0.0150)	(0.0160)	(0.0160)
Firm_age	−0.0065	0.0001	−0.0038	−0.0036	−0.0006	−0.0006
	(0.0079)	(0.0076)	(0.0074)	(0.0076)	(0.0082)	(0.0082)
log(Emp)	−0.0252	−0.0295†	−0.0335*	−0.0334*	−0.0184	−0.0184
	(0.0175)	(0.0165)	(0.0160)	(0.0165)	(0.0183)	(0.0183)
logit(IntlSales)	0.0207*	0.0263**	0.0243**	0.0249**	0.0212*	0.0212*
	(0.0096)	(0.0092)	(0.0090)	(0.0091)	(0.0098)	(0.0098)
logit(R&DInt)	0.0171	0.0228	0.0236	0.0207	0.0155	0.7541
	(0.0167)	(0.0164)	(0.0159)	(0.0164)	(0.0174)	(1.1971)
logit(R&DInt)^2	0.0018	0.0040	0.0024	0.0017	0.0029	0.6581
	(0.0047)	(0.0045)	(0.0044)	(0.0045)	(0.0048)	(1.0903)
(Intercept)	0.2432†	0.4347**	0.4589***	0.4491***	0.1536	0.1390
	(0.1389)	(0.1416)	(0.1191)	(0.1221)	(0.1610)	(0.1582)
Sector Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Region Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	3507	3782	3910	3749	3331	3331
R^2	0.0281	0.0259	0.0241	0.0254	0.0330	0.0330
adj. R^2	0.0242	0.0215	0.0210	0.0217	0.0259	0.0259
Resid. sd	0.9912	0.9875	0.9877	0.9869	0.9867	0.9867

Standard errors in parentheses
† significant at $p < .10$; $^*p < .05$; $^{**}p < .01$; $^{***}p < .001$

Combining all specifications yields largely the same levels, with the exception of founder age and knowledge scope losing significance entirely, and the max age of the founder becomes statistically significant and negative (p<.001). Additional testing yields that adding any of the three formation factors (`FF1-FF3`) to specification II contribute to this variable becoming more significant. Specification VIII shows that orthogonal polynomials for R&D Intensity do not really affect the coefficients' significance, though the intercept becomes weaker when added.

The third response variable of Model 2, EXPC3, or, the importance of codified academic and trade knowledge, is regressed on the internal knowledge intensity variables in Table 4.24. The first specification shows that while the proportion of tertiary educated employees is positively related to EXPC3, the proportion of Master and PhD level educated employees is negatively associated. Also, higher founder education levels positively associated with reliance on EXPC3. Specification II yields weaker results, with only founder's entrepreneurial experience and age of founding team have a statistical association. The next two specifications yield no associations at any level. Combining all variables in the final 2 specifications does not yield any major deviations from what we have observed thus far. Nonetheless it is noteworthy that the intercept value does not retain significance in statistical terms when all variables are added to the model. Also of note in all six specifications is the relatively sparse statistical significance of the control variables. Only international sales seems stable across all specifications.

4.6.2.2 Summary of effects and hypothesis confirmation - Model 2

By focusing on the sixth column (the fully specified model) of the regression results in Tables 4.22 –4.24, one can see the effects of the different internal knowledge intensity variables on that of the EXPC1, EXPC2, and EXPC3 components representing specialist knowledge provider-class actors, intra-industry actors, and codified academic and trade knowledge, respectively.

The relationship between education and knowledge reliance is elucidating. First in terms of education of employees, EmpEdu (a percentage of employees holding at least a tertiary degree) has a statistically significant and positive effect on the reliance on knowledge stemming from EXPC1 as well as EXPC3, suggesting that higher educated workforces make new ventures more prone to increased reliance on these types of knowledge, both formal and informal science and technology actors and relationships. Conversely, EmpEdu is negatively associated with EXPC2, intra-industry knowledge source reliance. EmpHiEdu, or, the fraction of employees holding a PhD, is statistically significant and negatively related with both EXPC2 and EXPC3 (at $p<.01$ and $p<.05$, respectively). This suggests that highly Master- and PhD-level educated workforces rely less on intra-industry sources, which is not unexpected really, as they may have established more relational networks within academia and related research fields and may opt for those instead. However, one must be careful in this interpretation since there is no significant coefficient for

EXPC1, and moreover the sign is negative. Looking at the effects plots on the following page however, this could be due to the low variance in that particular variable at higher levels. At the far right of the plot mapping the effect of EmpHiEdu onto EXPC1, one can see that the confidence bounds increase dramatically as the variable grows past 0.2, before which the relationship looks more statistically significant and negative. A similar trend is noted when comparing EmpEdu and EmpHiEdu's relation to the other 2 dependent variables, though not as dramatic. In terms of the founder education patterns, they have a similar directional effect to that of PhD frequency in employees. So, **WH-2.1 is partially confirmed but on a whole must be rejected**, since firms with a higher percentage of well-educated employees tend to value higher specialist knowledge providers and codified knowledge, but firms with a large number of tertiary and PhDs devaluate both intra-industry sources; tertiary education values codified knowledge while PhD devalues it; and highly educated founders tend to devaluate intra-industry sources and to increasingly valuate codified knowledge. This is reinforced when one looks at the principal component based variable FF3, depicting the importance of specialized knowledge like technical and design skills in forming the firm, which could also be used to proxy educational attainment as well as its importance for the firm. This variable negatively associates with EXPC2 ($p<.001$) but has no association with EXPC1 or EXPC3. It becomes clear that these effects point to more complexity than anticipated in these interrelations between employee and founder education with that of external knowledge source reliance.

Moving on to the previous experience of the founders in a more specific capacity, we find that the presence of former university or research institute employees increases the reliance on knowledge stemming from specialist knowledge providers (EXPC1), while decreasing the reliance on intra-industry sources (EXPC2), while an effect on EXPC3 was not discernable from the data. Industry and entrepreneurial experience on the founder level provided less conclusive results: Entrepreneurial experience seems potentially negatively related to reliance on specialist knowledge provider (EXPC1) sources ($p<.10$), especially when the experience constructs are isolated as in specification II of Model 2.1 (Table 4.22). The same variable was actually positively associated with codified sources (EXPC3), and not associated with intra-industry sources. Years of experience on the other hand, is only statistically significant ($p<.001$) and negatively related to intra-industry knowledge sources, though this only occurs in the fully specified model, potentially because one of the formation factors (FF2) stipulates how much experiential and network factors had to do with founding the firm. Possibly, the lower the average

age of the founders, the more experiential knowledge drives how much firms value intra-industry sources. Age of the founder only associated with EXPC1 (positive, $p<.05$) and EXPC2 (negative, $p<.10$) in specification II, with only the experiential variables included in the model, and with EXPC3 (negative, $p<.10$) in all specifications. FF2, which captures the importance of experience and networks in founding the firm, negatively associates with EXPC1 ($p<.001$) and negatively associates with EXPC2 ($p<.001$), which is in line with the other industry experience proxies in the model. All of this leads to a bit of a jumbled hypothesis confirmation for WH-2.2a, 2.2b, 2.3a, and 2.3b. While each gives results that lends credence to the hypotheses' confirmation, all variables included in the operationalizations do not yield similar results.

Concerning Cantner et al.'s functional diversity indicators, knowledge scope and knowledge disparity, there are little to no significant effects. In fact across all models they are insignificant[20]. **Thus WH-2.4a and WH2-4b are rejected.**

`Spinoff` activity produces no statistically significant effects for either extra-industry knowledge source category (EXPC1 and EXPC3), while a negative relationship is significant for EXPC2 at the $p<.05$ percent level. **WH-2.5 is then rejected according to the data.**

What then do the results say about novel opportunities as formation factors of the firm? Here we find some quite strong associations. Formation factors derived from novel opportunities (stemming from technical change, regulatory/institutional change, newly arisen market needs, or public procurement) showed strongly significant and positive effects on both reliance on SKP knowledge (EXPC1) and intra-industry knowledge (EXPC2), though not codified academic knowledge (EXPC3). **WH-2.6 then is partially confirmed.**

Again using the `car` and `effects` packages in R, I have created effects plots for each of the 3 fully specified models (column VI in each table). The shapes of the effects depict the variation in statistical significance throughout the data space, and can serve as a backdrop to discuss differing values in explanatory variables as having different marginal effects. These may be found in the appendices, Figures 8.5 – 8.7.

[20] When however, the founder background variables are isolated in Table 4.23, specification III, then EXPC2 is weakly associated with Knowledge Scope, but that is the only tangible change in significant effects for these particular variables.

4.6.2.3 Diagnostics - Model 2.1

I begin diagnosing Model 2 by constructing some basic residual plots: These map the Pearson residuals onto the fitted values of the response variable as well as the predictor variables: The residual plots for all three models can be seen below (Figures 8.8 – 8.10 in the appendices). The basic takeaway from these plots is in that the Pearson residuals ought to be independent of the predictor values and the fitted values of the regression model. Systematic deviations from the 'null' plot can indicate violations of the assumptions. It must be noted that these plots are more useful in identifying problems than in solving them however. A lack-of-fit test is also computed for each set of regressions (the table accompanying each plot). Clearly non-linear patterns may be identified by consulting both the graphs, as well as the p-values for the lack-of-fit test. In this case, some of the specifications in the models seem to be potentially problematic. Significant p-values are returned for `FF1` and `FF2` in 2.1 (`EXPC1` as response), `FF2` and `FF3` in 2.2 (`EXPC2` as response), and `EmpEdu` and `EmpHiEdu` in 2.3 (`EXPC3` as response. Tukey's (1949) test of non-additivity tests the squares of the fitted values against the standard normal distribution. Significant values indicate inadequate model fit. We can see from below that while there are some problematic variables, the condition is most serious in Model 2.1, where the Tukey test has actually been failed. Fox and Weisberg (2011) recommend that an often ideal starting place to correcting for non-constant error variance and similar diagnosed problems in a linear model is with the transformation of variables in the regression to normalize the fit. Since the Tukey test was failed in 2.1, this indicates a failure of the dependent variable to adhere to the normal distributional assumptions. I use the Yeo-Johnson (Yeo and Johnson, 2000) family of power transformations to approximate what type of transformation would best maximize the so-called *profile likelihood function*, the reason for which being that approximating the power transformation that maximizes this function behaves in a way like a maximum likelihood estimation function, and helps to specify what type of transformation of a variable will approximate normality in the variable to the highest degree possible (Box and Cox, 1964). The transformation approximates that around λ=0.7 we get the ideal likelihood. I approximate this by transforming to 0.5, or, the square root of the response variable; this is often a good alternative to more complex transformations (Fox, 2016).

Figure 4.24: Yeo-Johnson recommended power transformation of `EXPC1` in 2.1

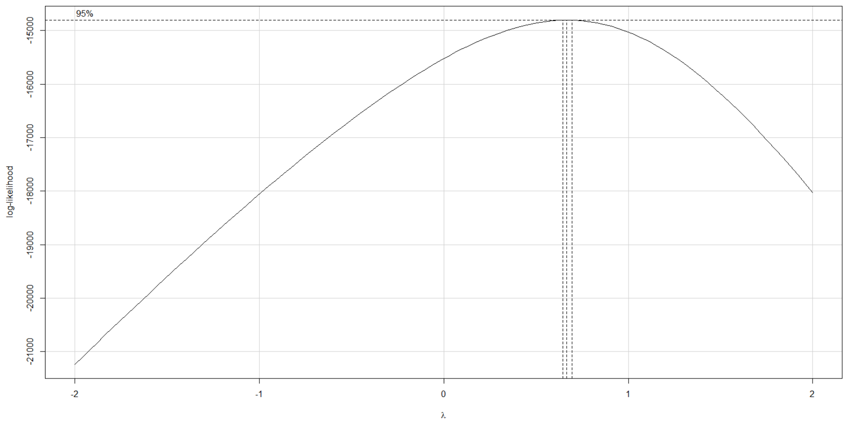


Table 4.25: Lack-of-fit test following Yeo-Johnson power transformation of EXPC1

	Test stat	Pr(>|t|)
EmpEdu	-0.42	0.67
EmpHiEdu	0.44	0.66
FoundEdu	2.09	0.04
FoundEnt	0.66	0.51
FoundUni	0.92	0.36
AgeMax	-0.58	0.56
KScope	0.39	0.70
KDisp	-0.88	0.38
FF1	1.20	0.23
FF2	-2.97	0.00
FF3	-0.63	0.53
Firm_age	0.88	0.38
log(Emp + 1)	0.91	0.36
logit(IntlSales)	-1.22	0.22
Tukey test	1.03	0.30

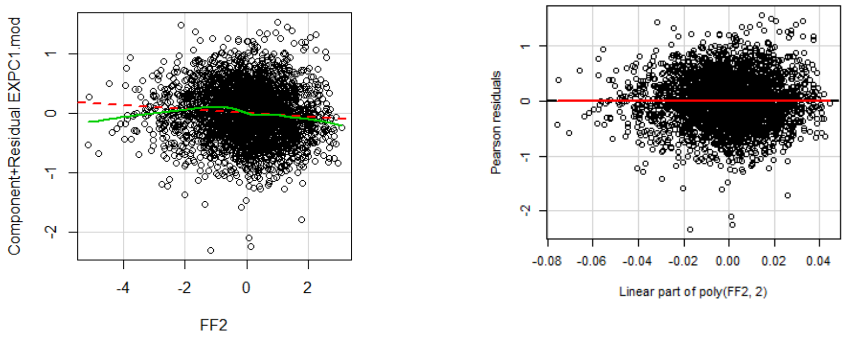
The regression is rerun after adding a constant to the response variable (3.5) to ensure that all values are positive, and taking the square root: I will not show the residual plots again for reasons of space and clutter, but the non-additivity test in Table 4.25 is retrieved from the modified regression. We now no longer have a failed Tukey test, indicating an adequate fit of the model with the assumptions of the normal distribution. However, we still have an explanatory variable that has failed the lack-of-fit test, `FF2`. While the variable `FoundEdu` also is in violation, it is not as close to 0 and not of particular concern[21]. I construct component residual plots, similar to the marginal model plots I used in Model 1 to assess the marginal fit

[21] Additionally, upon viewing the updated marginal model plot of this variable, the fitted and observed values coalesce quite nicely, suggesting that a power transformation will not be

Figure 4.25: Residual plots of FF2 in 2.1

Component residual plot of FF2 in 2.1

Residual plot of FF2 following square transformation



of this particular explanatory variable with the response variable. It can be observed in the plot that there is a slight bend to the relationship, indicating that a polynomial of degree 2 might fit the equation a bit better (Shown in Figure 4.25, left pane).

I rerun the regression for 2.1 a third time, this time with the transformed response (`EXPC1`) and explanatory (`FF2`) variables. The resulting model retains the same direction of effects with improved significance levels in several coefficients (see Table 4.26). Additionally, the model's residual plots of the `FF2` variable now show no deviation from the null line (see Figure 4.25, right pane):

helpful. In this researcher's opinion, the disturbance is likely due to the ordinal construction of the variable.

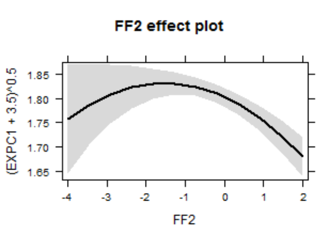
Table 4.26: Model 2.1 with transformed response and explanatory variables

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.9029	0.0725	26.25	0.0000
EmpEdu	0.0821	0.0369	2.22	0.0262
FoundEdu	0.0035	0.0089	0.40	0.6900
FoundEnt	-0.0249	0.0176	-1.41	0.1578
FoundUni	0.2267	0.0499	4.55	0.0000
FoundInd	-0.0013	0.0009	-1.43	0.1537
AgeMax	0.0140	0.0113	1.24	0.2161
KScope	-0.0167	0.0406	-0.41	0.6811
KDisp	0.0408	0.0498	0.82	0.4129
Spinoff	0.0116	0.0236	0.49	0.6221
FF1	0.1129	0.0053	21.39	0.0000
FF2	-2.1500	0.4765	-4.51	0.0000
$FF2^2$	-1.4140	0.4714	-3.00	0.0027
FF3	-0.0073	0.0074	-0.99	0.3225
Firm_age	-0.0075	0.0037	-2.00	0.0461
log(Emp + 1)	0.0104	0.0084	1.24	0.2145
logit(IntlSales)	0.0080	0.0045	1.77	0.0772
logit(R&DInt)	5.3905	0.5005	10.77	0.0000
$\text{logit(R\&DInt)}^2$	-2.4260	0.4624	-5.25	0.0000
SectorCLASS[T.KIBS]	-0.1023	0.0322	-3.17	0.0015
SectorCLASS[T.LTMS]	0.0003	0.0279	0.01	0.9922
SectorCLASS[T.OBS]	-0.0137	0.0290	-0.47	0.6367
Sector[T.Mid EU]	-0.1814	0.0301	-6.02	0.0000
Sector[T.North EU]	-0.1604	0.0298	-5.37	0.0000
Sector[T.South EU]	0.0661	0.0294	2.25	0.0245

Using Akaike's information criteria (AIC), it becomes apparent that while the first transformation vastly improved the likelihood fit of the model; the second power transformation, while also an improvement, is more marginal (See Table 4.27). Nonetheless, the finalized model has increased the R squared by about 0.0015, which while minuscule is still an improvement, and the exercise has proved quite enlightening. We now have a significant quadratic negative compound effect from the `FF2` variable, or, the importance of experiential and network knowledge in the formation of the company, onto the reliance placed upon knowledge stemming from specialist knowledge providers.

Table 4.27: AIC for all versions of 2.1, and effect plot for FF2 post-transformation in 2.1

	df	AIC
3.1 original fit	25.00	13074.52
3.1 transformed EXPC1	25.00	4210.84
3.1 transformed EXPC1 & FF2	26.00	4204.61



4.6.2.4 Diagnostics - Model 2.2

I continue the diagnostics by analyzing the results of the lack-of-fit tests of model 2.2, in which `FF2` and `FF3` were both found to be in violation. The Tukey test was however passed, so I will not labor with transforming the dependent variable this time. Suspecting similar curvilinearity in these variables I skip some steps and include a quadratic polynomial of `FF2` and `FF3` in the regression for model 2.2. The results are somewhat surprising (note that I remove all non-significant and control variables from the output but they were included in the regression with non-substantive differences:

Table 4.28: Model 2.2 with quadratic polynomials included

	$EXPC2$
EmpEdu	-0.3434^{***}
	(0.0830)
FoundEnt	0.0144
	(0.0430)
FoundUni	-0.4144^{***}
	(0.1204)
AgeMax	-0.0897^{***}
	(0.0241)
KScope	0.0765
	(0.0985)
KDisp	$0.2104^{\dagger}$
	(0.1213)
Spinoff	-0.1255^{*}
	(0.0575)
FF1	0.0995^{***}
	(0.0130)
FF2	4.3294^{***}
	(1.1295)
FF2^2	4.3580^{***}
	(1.1560)
FF3	-6.7088^{***}
	(1.1738)
FF3^2	3.2573^{**}
	(1.1536)
Firm_age	0.0006
	(0.0091)
log(Emp + 1)	-0.0005
	(0.0205)
logit(IntlSales)	-0.0052
	(0.0110)
logit(R&DInt)	-3.6204^{**}
	(1.2188)
logit(R&DInt)^2	-1.2927
	(1.1273)
N	3331
R^2	0.0817
adj. R^2	0.0753
Resid. sd	1.1052

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

It would appear that there are some more hidden quadratic relationships in Model 2.2 . Exponential marginal benefits appear to occur from formation factors like industry experience and networks (`FF2`) onto reliance on intra-industry knowledge (`EXPC2`), while a weak U-shaped

Figure 4.26: FF2 and FF3 quadratic effects on EXPC2

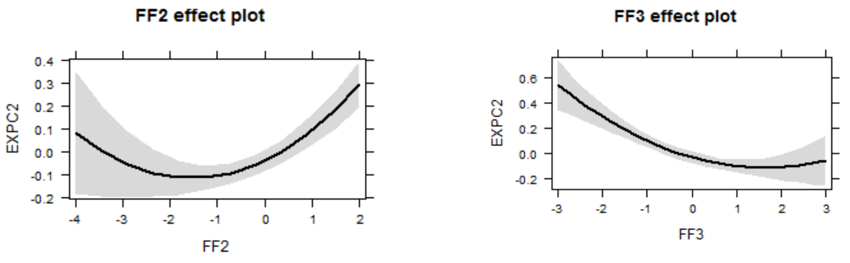
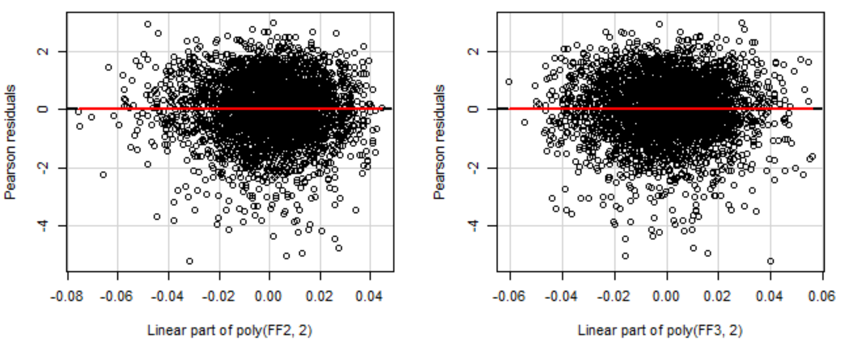


Figure 4.27: 2.2 modified model predictor residual plots



relationship is observed between formation factors of specialized knowledge like technical and design expertise and that of `EXPC2` (see Figure 4.26). Interpretation of this second effect is that potentially, *lower importance of specialized knowledge in the formation of the company reduces the reliance on industry knowledge, but as this knowledge improves or is more crucial to the foundation of the venture in question, the effect is that a higher reliance is place on intra-industry partners than before, coupled with a heightened sophistication of technique and design in the delivery of the product or service.* All in all, some interesting results that could not be obtained at first glance of the model. The adjusted R-squared has now fractionally improved, and checking the residual plots again yield much improved 'null' relationships between the variables and their residuals (see Figure 4.27).

Finally, we can address the third model 2.3, with informal, codified scientific and technological knowledge as the response. Problematic variables during the lack-of-fit test in this model were `EmpEdu` ($p<.01$) and `EmpHiEdu` ($p<.02$). I try first to address the more serious fit issue of the `EmpEdu` variable, before pursuing an eventual modification or transformation of the `EmpHiEdu` variable. It can be especially sensitive since the two are related to one another in that they are both derived similarly from the same survey variable set.

The component residual plot of that variable (just below) shows that the actual smoothed, non-parametric loess curve denotes a somewhat curvilinear relationship, as opposed to the dotted red line of the fitted regression. So, I add a squared term to the model using an orthogonal polynomial fit for `EmpEdu`. The results of this can be seen below (again I remove non-significant and control variables for clarity of interpretation, no major changes of coefficients or significance have occurred concerning these).

Figure 4.28: Component residual plot for 2.3 predictor EmpEdu

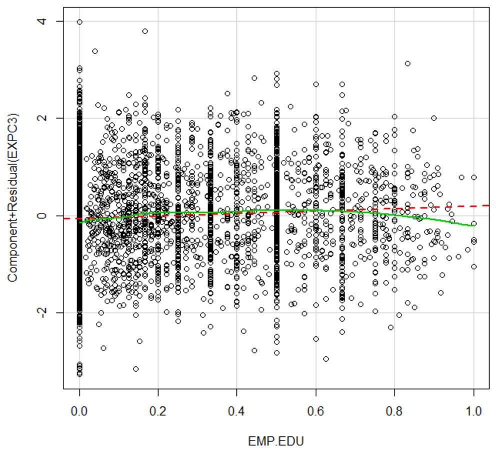


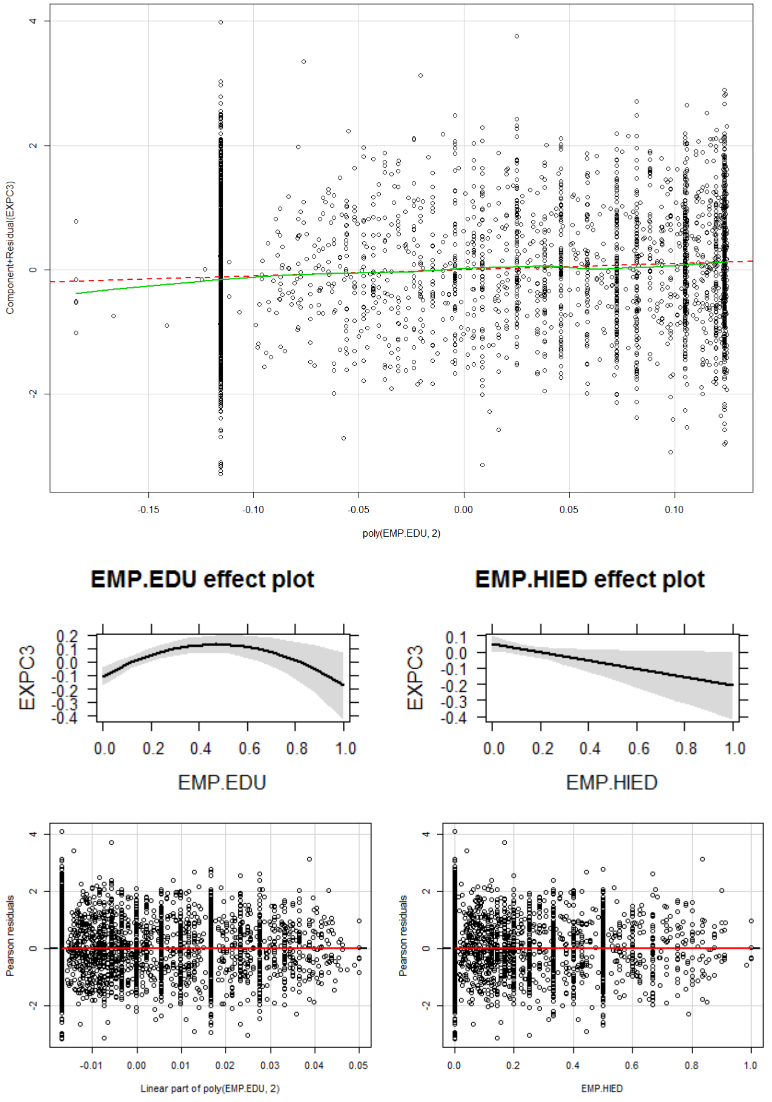
Table 4.29: Modified 2.3 with 2nd degree polynomial for EmpEdu included

	EXPC3
(Intercept)	0.2507
	(16.3958)
poly(EmpEdu, 2)1	3.9876^{*}
	(1.5860)
poly(EmpEdu, 2)2	-3.9591^{***}
	(1.0298)
EmpHiEdu	-0.2584^{*}
	(0.1242)
FoundEdu	0.0455^{*}
	(0.0204)
FoundEnt	0.0964^{*}
	(0.0386)
AgeMax	$-0.0435^{\dagger}$
	(0.0248)
logit(IntlSales)	0.0228^{*}
	(0.0099)
N	3296
R^2	0.0378
adj. R^2	0.0305
Resid. sd	0.9858

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

I can now see that there exists a statistically significant curvilinear relationship between `EmpEdu` and `EXPC3` that the first version of the model did not capture. The final step is to rerun the lack-of-fit test to see if the problem has been resolved. Although now that the polynomial term is in the regression the index cannot be calculated so we must rely on graphical interpretation to confirm we have made the right choice for the `EmpEdu` polynomial addition. A follow up p-value of `EmpHiEdu` with the new polynomial for `EmpEdu` included is given as 0.923, denoting that the real culprit was the `EmpEdu` variable, and transforming that one has fixed the problem for both variables. It is also worth noting that the AIC improved by about an 8 point decrease between these two models. Figure 4.28 shows the original variable residual plot, while Figure 4.28 shows plots for the revised model with the included quadratic term. Graphically, in the lower panes of Figure 4.29, we can see that in both the residual and component residual plots that the relationship is predominantly a 'null' plot, as we had hoped it would be. While the bottom tail of the component residual plot is still slightly sagging down, I am not concerned, since I know that the variable itself has a considerable pileup around 0, conveying that many of the ventures do not have tertiary-educated employees. Clearly this is partly due to the sampling of the survey, which included diverse sectors, some of which (namely Wood and furniture and Textiles) have means of below 10% for this particular variable, While more high tech industries are often between 30% and 50%.

Figure 4.29: Component residual, effects, and residual plots for EmpEdu after diagnostics



Some effects plots of these variables from the modified regression model show the relationship more clearly. As the percentage of tertiary-educated employees increases, there is a higher likelihood of higher reliance on codified science and technological knowledge, but as this increases even more, the likelihood of high reliance declines marginally (and though the confidence bound widens, but the decrease is still apparent). The `EmpHiEdu` variable maintains its former shape, still signifying that with greater PhD proportions in the employee body, the reliance on codified knowledge of this type may decline (slightly or greatly according to the confidence bounds). The width of the bounds should be taken with a grain of salt though, as the data space on the right tail of the distribution of the variable is somewhat sparse, with few firms having more than 40% PhD educated employees.

Finally, the variance inflation factor tests for each of these revised models show that collinearity is not a serious problem for any of the modified regressions (Table 8.6 in appendix). This concludes Chapter 4, which gave a detailed account of the data, variables, models, and diagnostics of the regression sets in Model 1, which detailed the relationship between external knowledge intensity and innovative performance, and Model 2; which detailed the relationship between internal knowledge intensity and external knowledge intensity. The next chapter, Chapter 5, will overview the regression models used to fit the combined AEGIS and Orbis dataset. The regression sets for Model 3 and Model 4 will be covered; Model 3 shows how internal knowledge intensity may potentially affect business performance of the KIE ventures, while Model 4 looks at the relationship between innovative performance, our dependent variable from Chapter 4, and business performance, and how the former may affect the latter.

Chapter 5

Research method, data, descriptives, and analysis: Part 2 - Models 3 and 4

I now turn to the examination of the remaining two empirical sets, those that aim to achieve the following research objectives:

RO3: *Explore the association between internal knowledge intensity and business performance of the entrepreneurial firm.*

RO4: *Explore the association between innovative performance and business performance of entrepreneurial firms.*

In order to do this, we constructed several hypotheses for each research objectives through Chapter 3's review of the relevant literature related to these phenomena:

> H-3.1: Higher levels of human capital in the form of education of founders and employees are positively associated with business performance in entrepreneurial firms.
>
> H-3.2: Higher levels of human capital in the form of entrepreneurial, industrial, and academic experience of the founding team are positively associated with business performance in entrepreneurial firms

These hypotheses stems from the human capital literature accounting for the effect of human capital investments and human capital outcomes (Becker, 1964). I expect that higher levels of education and relevant experience of founders and employees will positively affect business performance

> H-3.3a: In terms of founding team functional heterogeneity: Knowledge scope is positively associated with business performance.
>
> H-3.3b: At higher levels of knowledge scope, it will have a diminished marginal association with business performance for entrepreneurial firms.
>
> H-3.4: In terms of founding team functional heterogeneity: Knowledge disparity is negatively associated with business performance for entrepreneurial firms.

These hypotheses stem from the notion that a successful founding team composition in a new venture both maximizes the scope of knowledge encompassed by the team and minimizes the disparity, or, non-beneficial differences in way of working, social categorization, and dissociative team properties. This should hold true for knowledge intensive entrepreneurial ventures due to the fact that there is a certain emphasis on adaptability, and a maximum need for diverse resources to succeed, and a diverse but not redundant or divergent management team best suites maximized performance on the business level. Additionally, I have argued previously that overly broad team knowledge scope will have diminished marginal benefits to performance, when a KIE founding team becomes too wide in their functional heterogeneity and cannot focus on performance outcomes. These hypotheses are grounded in the work of Cantner et al. (2010) and others who commonly relate team functional heterogeneity to performance outcomes in small firms.

> H-3.5: Having the organizational origin of being a corporate spinoff entrant is positively associated with business performance for entrepreneurial ventures

This hypothesis rises from the work done on organizational origins, and how different types of spinoffs perform compared with de novo entrants (Klepper, 2001; Dencker et al., 2009).

> H-3.6: The extent to which a firm's formation was based on novel opportunities should be positively associated with business performance in entrepreneurial firms.

This hypothesis stems largely from the literature on entrepreneurial opportunities and how they are recognized, created, and drive venture performance (Shane, 2000; Shane and Venkatraman, 2001; Brush et al., 2001).

Model 4 tests several working hypotheses relating innovative performance of KIE firms to their business performance, with higher innovative returns in different respects leading to heightened business performance:

> WH-4.1 Higher innovative performance in goods sales will positively associate with business performance for entrepreneurial firms.
>
> WH-4.2 Higher innovative performance in service sales will positively associate with business performance for entrepreneurial firms.
>
> WH-4.3 A higher degree of radicalness in products and services will positively associate with business performance for entrepreneurial firms.

Much of the innovation studies literature puts innovativeness at the forefront of competitiveness, but not so much has been done to measure its effect for new and potentially knowledge intensive ventures. The broader connotations of KIE firms' impact on economic growth have been well documented early on in this dissertation (Chapter 2 and latter parts of Chapter 3), but to what extent will we find a heightened economic or business performance on the firm level. Before a societal gain is to be reaped from the KIE activities, it is likely that the firm must find itself in a stable upward trend of its own, in order for the knowledge to have the opportunity to 'spill over', transfer, or otherwise reach society. Studies have shown recently that there is a dynamic relationship between innovation and growth for new firms, so how will this relationship materialize for these KIE firms? These questions are the focus of Model 4.

5.1 Data

While comprehensive in its scope and ambitious in its depth of analysis of core constructs to isolate the knowledge intensive components of the KIE venture, the AEGIS survey does not delve deeply into more tangible accounts of performance, nor does it offer any sort of longitudinal or hierarchical modelling potential, due to its cross-sectional nature. Hopefully, more data will be collective systematically from the firms sampled here, and many re-sampled sets of KIE firms from the same population will yield confirmatory and exploratory studies of how knowledge intensive firms compete, grow, and contribute to societal well-being. But for now, this has not been carried out. So, in order to deepen the analysis of the KIE concept and how it relates to these specifically sampled firms, I have returned to the source of much of the sampling frame in the AEGIS survey project to gather follow up data. Namely, the Amadeus/Orbis database. To construct a suitable dataset with which to answer the above-mentioned objectives and test the hypotheses, complementary data on the firms captured in the AEGIS survey originally stemming from the Amadeus database was obtained by utilizing their unique Bureau van Dijk (the owner of Amadeus) identification numbers, or BVD ID numbers. This was done for 2978 of the 4004 firms in the AEGIS sample.[1] 31 firms had been assigned new BVD ID numbers, so the necessary changes were made to match this

[1] Unfortunately, though many firms in the AEGIS sample come from other databases, no identification numbers were provided in the survey results, making tracking them down in their native databases, or in other databases, a difficult and unwieldy task to complete in a short period of time. Therefore, I limit this portion of the analysis to those firms collected from Amadeus.

dataset with the AEGIS sample. Hear I used the Orbis database of global firms, of which all Amadeus firms have been integrated by Bureau van Dijk. The following variables were obtained for the remaining 2978 firms.

Table 5.1: Variables extracted from Orbis database to match with the AEGIS survey

Company Name	Number of Patents	BVD ID Number
Number of available years	Market status (Listed/Unlisted/Delisted	NACE Rev 2 code
Status date	Date of incorporation	Delisted date (if delisted)
Profit/loss ration before tax $\{Y_0 - Y_{-9}$; 2006–2015$\}$	Net income $\{Y_0 - Y_{-9}$; 2006–2015$\}$	Cash Flow $\{Y_0 - Y_{-9}$; 2006–2015$\}$
Shareholder Funds $\{Y_0 - Y_{-9}$; 2006–2015$\}$	Current Ratio $\{Y_0 - Y_{-9}$; 2006–2015$\}$	Profit Margin $\{Y_0 - Y_{-9}$; 2006–2015$\}$
Return on Current Equity $\{Y_0 - Y_{-9}$; 2006–2015$\}$	Solvency Ratio $\{Y_0 - Y_{-9}$; 2006–2015$\}$	Number of Employees $\{Y_0 - Y_{-9}$; 2006–2015$\}$
Number of Trademarks	Last available year	Status /(active, inactive, deactivated)
Operating Revenue $\{Y_0 - Y_{-9}$; 2006–2015$\}$	Total Assets $\{Y_0 - Y_{-9}$; 2006–2015$\}$	Return on Equity $\{Y_0 - Y_{-9}$; 2006–2015$\}$

The variables collected directly from Orbis were then merged with the AEGIS survey data in order to provide a semi-longitudinal data set, with the Orbis variables being variant over time, and those from the AEGIS survey being in stasis in 2010/2011. The variables from AEGIS are associated with a 'snapshot' of where the firm was in its development and strategic environment in that specific time period, and will be interpreted as such throughout the modelling, analytical, and summative sections.

5.2 Operationalization and descriptive statistics

5.2.1 Internal knowledge intensity and innovative performance

The derivation of these sets of variables is actually identical to how I have carried out the process in Models 1 and 2 in Chapter 4, albeit with a slightly reduced sample size. I have matched the two datasets together by each firm's unique Bureau van Dijk identifier. By comparing density plots between variables in the two different data sets, it can verified we have not drastically altered any data characteristics of the variables despite reduction in sample size. The following page shows the rest of the variables as density plot overlays for each sample by variable, the dashed red line representing the variable taken from the full AEGIS survey, the solid black line taken from the reduced size sample of 2978 firms (those with valid BVD id numbers). Fortunately, after viewing the above plots, I can confidently say that there are no major deviations from the original sample. The sub-sampled descriptive statistics of each of the models appear in Table 5.2.

Figure 5.1: Density comparisons of the variable samples used in the original (AEGIS) and combined (AEGIS + Orbis) samples

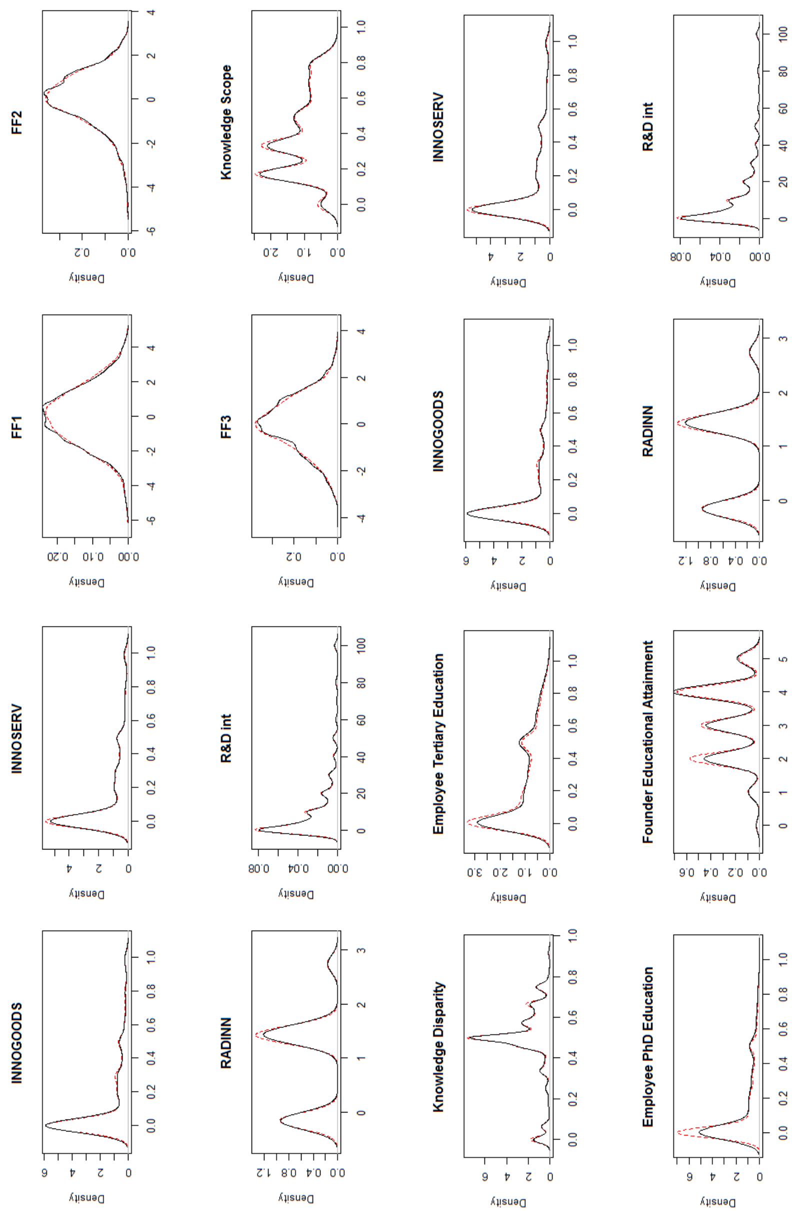


Table 5.2: Descriptive statistics for explanatory variables: Model 3

	n	mean	sd	med	trim	mad	min	max	skew	kurt
R&D Int	2501	12.83	19.79	5.00	8.26	7.41	0.00	100.00	2.55	7.07
FoundEdu	2501	3.24	1.07	3.00	3.25	1.48	1.00	5.00	-0.28	-0.81
FoundEnt	2501	0.31	0.46	0.00	0.27	0.00	0.00	1.00	0.81	-1.35
FoundUni	2501	0.03	0.17	0.00	0.00	0.00	0.00	1.00	5.55	28.80
FoundInd	2501	15.29	10.64	15.00	14.67	10.38	0.00	60.00	0.62	0.16
AgeMax	2501	3.21	0.82	3.00	3.29	1.48	1.00	4.00	-0.65	-0.56
EmpEdu	2501	0.28	0.27	0.22	0.25	0.33	0.00	1.00	0.58	-0.83
EmpHiEdu	2501	0.18	0.23	0.03	0.14	0.04	0.00	1.00	1.18	0.36
Spinoff	2501	1.14	0.35	1.00	1.05	0.00	1.00	2.00	2.07	2.28
ATTNEAVE	2501	-0.11	0.12	-0.08	-0.10	0.09	-0.82	0.18	-0.89	2.15
KScope	2501	0.39	0.22	0.34	0.38	0.25	0.00	0.93	0.33	-0.79
KDisp	2501	0.49	0.17	0.50	0.51	0.09	0.00	0.92	-1.23	1.84
Blau	2501	0.71	0.19	0.75	0.72	0.18	0.00	1.00	-1.16	2.39
FF1	2501	-0.01	1.61	0.00	-0.01	1.62	-5.49	4.59	-0.06	0.02
FF2	2501	0.02	1.20	0.10	0.08	1.12	-5.01	3.10	-0.52	0.56
FF3	2501	0.01	1.16	0.06	0.03	1.15	-3.94	3.48	-0.16	-0.09

Industry means	n	Spinoff	R&DInt	FoundEdu	FoundEnt
ICT	98	10.87	22.59	3.49	0.26
Machinery and equipment	110	11.83	10.24	2.78	0.27
Chemical industry	24	4.76	22.24	3.90	0.24
Paper and printing	258	12.61	9.70	3.00	0.35
Textile and Clothing	114	19.23	11.66	2.57	0.36
Food, beverages, and tobacco	160	19.29	6.75	2.71	0.39
Wood and furniture	150	24.62	8.40	2.38	0.33
Telecommunications	23	10.00	14.00	3.10	0.35
Computer and related	448	11.90	18.89	3.43	0.36
R&D	65	14.29	40.86	4.41	0.32
Other business services	1200	13.14	10.75	3.46	0.29

Industry means	FoundUni	FoundInd	AgeMax	EmpEdu	EmpHiEdu
ICT	0.08	17.12	3.38	0.27	0.18
Machinery and equipment	0.01	17.44	3.31	0.19	0.12
Chemical industry	0.05	13.14	3.76	0.27	0.23
Paper and printing	0.02	15.39	3.18	0.22	0.14
Textile and Clothing	0.01	16.49	3.18	0.09	0.06
Food, beverages, and tobacco	0.01	14.86	3.13	0.14	0.08
Wood and furniture	0.00	15.55	3.10	0.07	0.04
Telecommunications	0.05	11.90	2.90	0.22	0.11
Computer and related	0.04	12.81	3.01	0.35	0.21
R&D	0.27	16.66	3.41	0.52	0.38
Other business services	0.02	15.66	3.29	0.33	0.22

Industry means	KScope	KDisp	Attneave	Blau
ICT	0.42	0.50	-0.13	0.70
Machinery and equipment	0.42	0.50	-0.10	0.72
Chemical industry	0.42	0.47	-0.13	0.70
Paper and printing	0.40	0.50	-0.11	0.74
Textile and Clothing	0.39	0.52	-0.09	0.74
Food, beverages, and tobacco	0.41	0.51	-0.10	0.74
Wood and furniture	0.45	0.52	-0.12	0.75
Telecommunications	0.39	0.45	-0.07	0.70
Computer and related	0.40	0.49	-0.11	0.72
R&D	0.38	0.48	-0.08	0.70
Other business services	0.37	0.47	-0.11	0.68

Industry means	N	% No Inno.	% Firm	% Mkt	% World	% R&D Int	% Goods	% Serv
ICT	98	26.53	44.90	16.33	12.24	20.14	0.33	0.16
Machinery etc.	110.00	38.18	36.36	17.27	8.18	10.46	0.24	0.09
Chemical industry	24	29.17	33.33	20.83	16.67	23.08	0.20	0.15
Paper and printing	258	44.57	32.95	17.05	5.43	9.88	0.17	0.14
Textile and Clothing	114	42.98	37.72	13.16	6.14	9.91	0.26	0.10
Food, beverages, etc	160	41.25	35.62	15.00	8.12	7.13	0.17	0.07
Wood and furniture	150	46.00	32.67	12.67	8.67	7.88	0.21	0.09
Telecom	23	26.09	56.52	8.70	8.70	13.48	0.15	0.34
Computers	448	30.13	37.72	24.33	7.81	18.43	0.19	0.25
R&D	65.00	32.31	27.69	21.54	18.46	42.74	0.23	0.24
Other business serv.	1200	46.42	33.08	16.17	4.33	10.81	0.10	0.19
Mfg. of metals	102	50.98	33.33	13.73	1.96	8.78	0.15	0.10

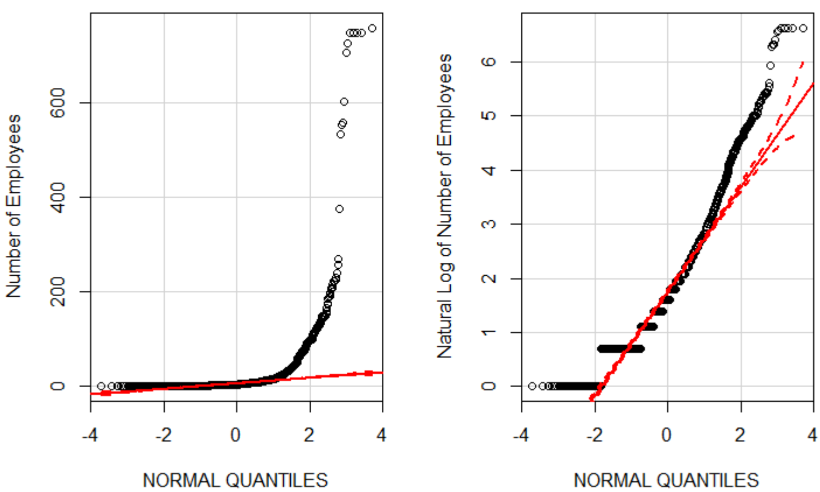
	n	mean	sd	med	trim	mad	min	max	rng	skew	kurt	se
InnoGoods	2752	0.16	0.25	0.00	0.10	0.00	0.00	0.99	0.99	1.67	1.98	0.00
InnoServ	2752	0.17	0.25	0.00	0.12	0.00	0.00	0.99	0.99	1.54	1.70	0.00
RadInnOS	2752	0.86	0.91	1.39	0.80	0.23	-0.16	2.75	2.90	0.13	-1.11	0.02

5.2.2 Business performance variables

As previously recounted, the variables used to measure the firms' business performance were derived from the complementary dataset obtained from Orbis, and hence, only encompass the 2978 firms that were drawn from Amadeus during the sampling procedure for the AEGIS survey. These variables are:

1. The natural logarithm of the *number of employees* each year ranging from 2010 (the year the survey was administered) to 2015, the present year of this study, denoted as `logEmp`. This is used as an approximate operationalization of firm growth, following Colombo and Grilli (2005). Quantile comparison plots again provide good justification of this transformation with the left figure untransformed, and the right figure transformed by the natural logarithm. While obviously imperfect, the values are brought in by the log transformation to lie closer to normality. It still may not be the best representation of the variable, but it is acceptably close for my purposes.

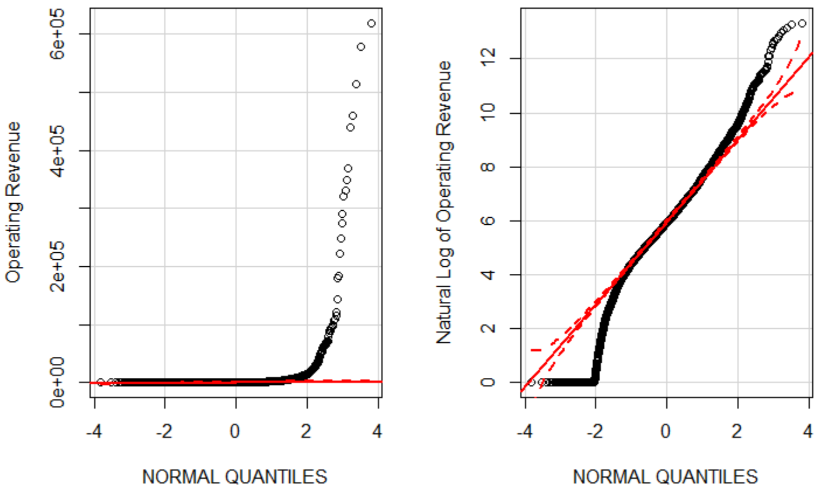
Figure 5.2: Transformation of dependent variable Number of Employees before regression



2. The natural logarithm of the operating revenue of the firm for the

same period, denoted as `logOpRev`. I justify again using quantile plotting (there are still a deal of outliers to the normal distribution following the transformation but again the deviation is markedly less).[2]

Figure 5.3: Transformation of dependent variable Operating Revenue before regression



3. Additionally, firm survival, or, the Cox (1972) proportional hazard function of whether or not the firm was marked as inactive during the period between founding and the present date was derived. The Cox (1972) survival indicator, `Surv`, combined with a baseline hazard indicator, `Lifespan`, and an indicator for the beginning of the observation period, `Time1` were constructed.

A variable was extracted from the Orbis database of firms with unique Bureau van Dijk identifiers, giving the following categorizations for the status of the firms in question:

Additionally, two additional variables were used in calculating the survival variable: `Status Date`, or, the date that Orbis assigned this

[2]Additionally, some firms were marked in Orbis with a negative operating revenue just below zero. In order to ease computation, these were left censored at zero prior to transformation.

Table 5.3: Status of firm coding from Orbis

Status	n
Active	2361
Active (default of payment)	7
Active (dormant)	25
Active (insolvency proc.)	72
Bankruptcy	26
Dissolved	217
Dissolved (bankruptcy)	37
Dissolved (liquidation)	17
Dissolved (merger or take-over)	6
Inactive (no precision)	4
In liquidation	53
Unknown situation	117
Missing values	36

externally retrieved status to the firm, and `StatusUpdatedInOrbis`, or the date when Orbis actually changed the data on this particular firm in their own database. In order to establish a valid survival indicator, the following steps were performed on these variables:

Initial transformations

1. I created a new variable: `MergedStatus`, which combines the two above-named variables, giving precedence to `StatusDate` and using `StatusUpdateInOrbis` if `StatusDate` is missing.

2. I recoded the variable `Status` as follows and renamed it to `Surv`, to denote whether the firm survived or not to the present day.[3] The recoding was due to the fact that I cannot interpret the status of the firm if the situation is unknown, already a missing value, or the firm has been merged or acquired. So, I decided to remove these values from the analysis and just coded them as missing, rather than risk clouding the interpretation of the regression to follow.

3. I created a dummy variable `DateInactive` that takes the value of 1 if the firm is inactive according to the transformed `Status/Surv` variable and 0 otherwise.

4. I created a variable `PossibleExit` that takes a missing value if the firm is active, and the year beginning the proposed inactivity (exit) otherwise.

5. I created a variable to measure the `Span` of the feasible exit period of the inactive firms where: `Span` = `PossibleExit` − `LastAvailableYear`.

[3]This recoding embodies the assumption that any of the above categorizations taken by the original `Status` variable convey inactivity, or at least, the beginning of the process of dematerialization. I assume here that if the process had been hindered, and the health of the firm had been restored, that it would have been updated as such in the database.

6. I created a variable `Feasibility` that takes a 1 if the span variable is between -1 and 2, 0 otherwise. This is by me assumed to be a reasonable bound within which to assume that the Possible exit variable represents a plausible 'year of exit' estimate. That is, if difference between the possible year of exit according to the status variable and the last year of available data in the Orbis database is larger than -1 year and smaller than 2 years, that it can be seen as an acceptable approximation of the year the firm became inactive or died.

7. Thus, I then created the variable `AssumedExit` that takes the value of the *PossibleExit* variable if the variable `Feasible` is 1, and 0 otherwise.

8. Finally, I created the variable `Lifespan` that takes the value of the `Start year` of the firm subtracted from `AssumedExit` if the firm is inactive, and the value of `StartYear` subtracted from the number 2015 if the firm is active.

At the end of this lengthy variable transformation and creation process, the result is two variables suitable for use in a Cox proportional hazards model. That is, one can then perform a Cox proportional hazard model on the `Surv` variable (the so-called event dummy variable denoting firm survival or exit as of 2015) as the response variable, with the baseline hazard function supplied by the `Lifespan` variable (the length of time the firm was or has been in existence in years). Additionally, left truncation to the model was supplied by specifying the time in the firms' lifespan which the surveying took place using the `Time1` variable. This is to combat potential survival bias in the data, since none of the firms could have possibly failed in the time leading up to the point when they were actually sampled and surveyed during the AEGIS survey.

5.3 Estimation Method

Following this, an analysis of business performance of select firms from the sample based on the indicators developed in the previous chapters was conducted, and are represented by datasets and Models 3 and 4. As previously mentioned. I have gathered additional data using the Amadeus/Orbis database, retaining the following indicators for n = 2978 of the sampled AEGIS survey firms. Thus the sample is slightly different for sets 3 and 4, but as we have shown, no distributional anomalies arose upon closer inspection of the data.

This data complements the AEGIS survey and helps to build a more structured mapping of these companies in a panel sense, combining time invariant constructs reported in 2010, time variant but snapshot-based constructs reported in 2010, and time invariant constructs reported in the Orbis database and retrieved in 2015. This makes the data used here quite blended, and consequently, somewhat difficult to structure and analyze. Various techniques have been applied throughout the study to ease interpretation and facilitate statistical modeling.

In order to derive appropriate survival analysis functions from the data, the AEGIS survey data was combined with the firm status variables used in Orbis, and matched by ID number, thus the survival analysis dataset is only those firms in the AEGIS survey that were taken from Amadeus/Orbis, so, 2978 firms. Since the survival indicator was collected via Orbis at a later date (October, 2015) than the administration of the survey (ranging from the fall of 2010 to the spring of 2011), it serves as a valid indicator. The Cox proportional hazard model for cross-sectional data (Cox, 1972; Cox and Oakes, 1984; Fox and Weisberg, 2011, online appendix) was used to assess the hazard function for these firms.

For the remaining business performance models, the data used was a combined set of variables from the AEGIS survey and the Orbis database. Those from the AEGIS survey represent the values of the data points taken in late 2010/early 2011 during the time of the survey. Those from the Orbis database range from 2010 to 2015. To combine these in a meaningful way, the data needed to first be merged, and then transformed so that the modelling procedure could be carried out. After first merging the two sets together using the BVD ID numbers of the 2978 firms, closer inspection of the data revealed some potentially serious problems concerning missing data values among the Orbis sub-sample. The following table outlines the percentages of missing data upon inspection using the R statistical computing environment and the package `Amelia II` (Honaker et al., 2011). A problem was thus encountered in that a sizable portion (up to 70% for Number of Employees, a key growth indicator for entrepreneurial firms) of some of the key response variables were missing in the data. This meant that a more advanced strategy of missing data treatment needed to be considered.[4]

Using `R` statistical computing environment, and the `Amelia II` package [5],

[4]Note: Even though I will now proceed to document the missing data imputation process, and analyze the subsequently generated empirical models. I have in fact tested the regressions without imputing any data and using listwise deletion instead, and have full result tables available for this in the appendices. This will also receive more extensive treatment in the form of robustness checks later on.

[5]I will not review the computational procedure performed by the package here, for a

Table 5.4: Missingness proportions for imputed variables

Variable of Interest	% missing
Number of patents	1.2
Number of trademarks	1.2
Operating Revenue	56.8
Profit/Loss Ratio	54.1
Net Income	54.0
Cash Flow	60.4
Total Assets	39.9
Share Funds	39.8
Current Ratio	43.1
Profit Margin	63.2
Return of Equity	59.6
Return on Current Equity	77.0
Solvency Ratio	42.5
Number of Employees	70.3
Export Revenue	84.8

a multiple imputation procedure upon the dataset was evaluated. Multiple imputation is an approach to data with missing values that has been shown to reduce bias and to increase computational and modeling efficiency compared with that of the more commonly practice listwise deletion (Honaker et al., 2011). Multiple imputation creates m values for each of the missing observations in ones data set, doing so across m different "imputed" datasets. The observed values do not change in these data sets, only the missing values are imputed (ibid.). The imputation model of Amelia II requires two criteria: first, the complete data (both present and missing values) must be following a multivariate normal distribution. This might sound like a grand assumption, but in most cases can be quite plausible concerning large datasets even with a large variation of different variable types (Schafer and Olsen, 1998). Second; the researcher must be able to make the assumption that the data are missing at random (MAR), meaning that the pattern of 'missing-ness' does not depend on unobserved data, but only on that data which is present in the sample.[6] Another contemporary interpretation is that this missingness is "related to the observed data but - conditioning on the observed data - not to the missing data" (Fox, 2016: 606).

Though the MAR assumption is only truly formally testable upon gathering evidence through an extensive follow up study or an alternative gathering of the missing values (Schafer, 1997), the scale and scope of this study did not allow such an undertaking. However, checks found no evidence that the distribution of missingness among the variables was systematically related to any particular sector, country, region, time-related patterns, or

detailed account of the expectation maximization bootstrapped (EMB) algorithm employed by the package, consult Honaker and King (2010), Honaker et al. (2011), and Fox (2016), Chapter 20

[6] One may consult the work of Rubin (1976) for a more detailed account of the generally established missing data typology in statistical research.

any other key indicator, so it was decided that the assumption that the data was missing at random (MAR) was reasonable. Thus, I did not find any reason to reject the MAR assumption or that of multivariate normality as described above, so the use of multiple imputation was deemed to be a reasonable solution to the missing data problem present here.

Since the time series functionality of Amelia II is quite accommodating, with the ability to specify constant or time-invariant variables, variables denoting the passage of time (in this case, `Year`), and the cross–sectional variable across which the variation over time is relevant (here, the specific firm), the combined dataset made from the Orbis and AEGIS samples was put into long, or panel format, with the result being a set of variables that are time-invariant across all years (those from the AEGIS survey) and a set of variables that are time-variant from year to year, both for each of the 2978 firms that were originally drawn from Amadeus by the AEGIS project and survey. This was done using the `dplyr` and `string`' packages in `R`.

Through correspondence with select authors of the `Amelia II` package, and consulting reliable missing data literature (compiled in Fox, 2016), it was suggested that an imputation model should have at least as many (if not more) parameters as are to be used in subsequent regression. Thus I have included in the imputation model all the variables obtained via the Orbis database regardless of their inclusion in the final regressions. Also, all explanatory variables in both models and their controls were included in the imputation process, but often not directly imputed themselves. Based on Graham et al.'s (2007) recommendation of decreasing levels of power falloff by increasing the number of imputations dramatically as opposed to what is commonplace (more than 10 is considered unusual by some), I have instead opted for maximum statistical power retention given the high number of missing values in the Orbis data (see the preceding table) and imputed 100 different datasets. While analysis post-imputation is rendered more cumbersome and computationally demanding when this many imputations are used, clarity in the results benefits greatly from it. The subsequent regression analysis is attained by combining the regression estimates and averaging across them, so that the parameter $\bar{q}$ is the average of the m separate estimations so that, given that:

$$q_j(j = 1, ..., m) : \bar{q} = \frac{1}{m} \sum_{j=1}^{m} = q_j \tag{5.1}$$

That is, the regression coefficients and standard errors are averaged across all 100 imputations for each predictor variable and fitted value of the response variable. This option is specified and encouraged by Honaker et al. (2011). The regressions and averaging procedures were carried out using the `Zelig` package in `R` (Imai, King, and Lau, 2007, 2008; Lam, 2007), as well as the `mi` function in Stata.

Figure 5.4: Missingness map of variables used in imputation process: Y-axis by observation across variables

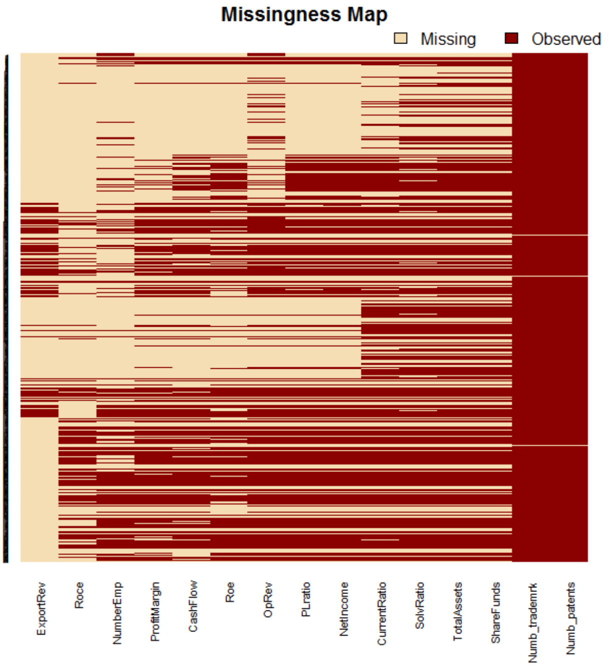
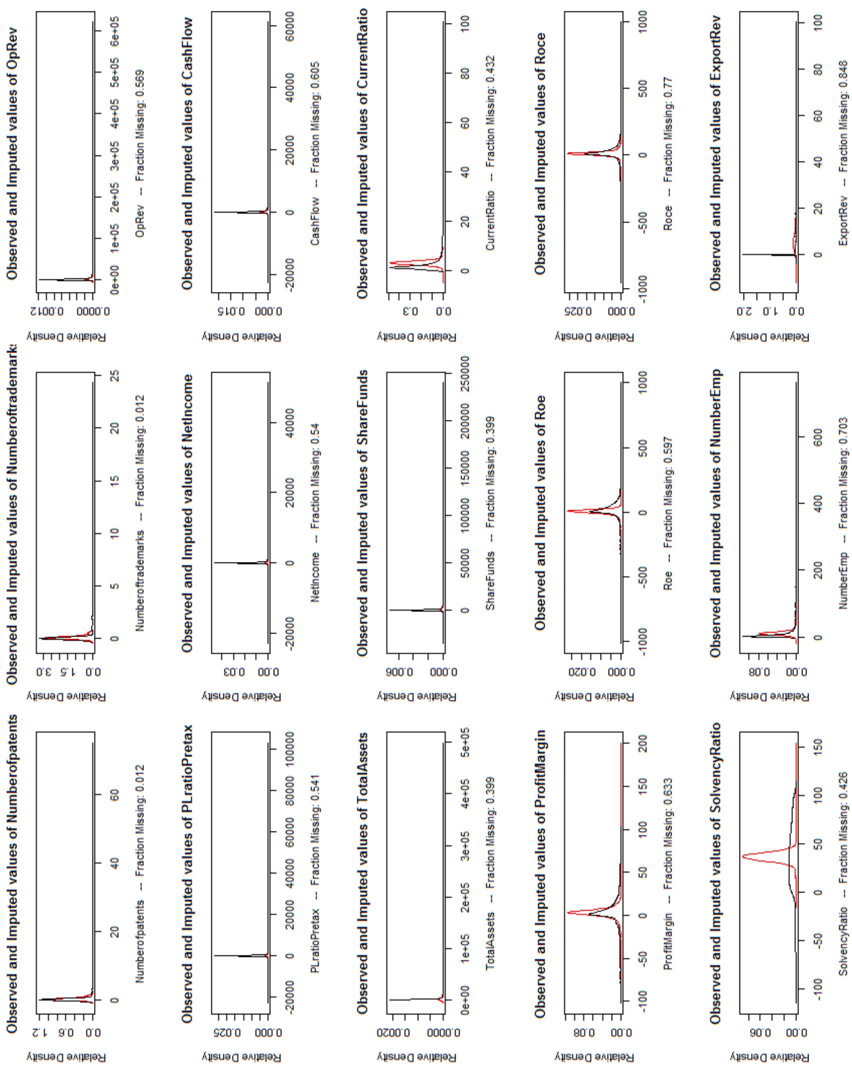


Figure 5.4 maps the missingness of the variables included in the imputation model. The red area signifies non-missingness (note that almost all the variables shown in the red "block" were excluded from the outputted imputation process, but that they were used in computing the imputation regression algorithm). As can be seen, the missingness was most extreme in many of the key variables obtained from Orbis. The plausibility of the imputation model employed is readily assessable by surveying the distribution of the new imputed data values and comparing with the distribution of the data observed before imputation, which is subsequently accomplished in Figure 5.5. Figure 5.5 plots these two distributions for each of the imputed financial variables taken from the Orbis database. The red line is the imputed data value distribution, and it can be seen that the imputed values largely follow closely to the

original observed data values' distributions. The map includes more variables than were actually regressed to increase precision, as well as some change-over-time variables that were not ultimately used in any regression. We can also plot the confidence intervals of the imputations vis á vis the original data values using the overimpute function in Amelia (see graphs 'Observed versus Imputed values', Figure 8.11, in the appendix). Color coding represents the fraction of missing observations present in the specific pattern of missingness, of which this imputation procedure contains around 1300 unique patterns, assigned to a particular observation. While there are some outliers present, it is apparent from viewing both these figures that there is a high percentage of imputed observations lying on the line, meaning that the true value with 90% confidence falls within this given range.

In order to estimate the regression appropriately given the unique structure of the combined data sets with a combination of time invariant explanatory variables and time-variant response variables, the use of generalized estimating equations (GEE) was chosen (Liang and Zeger, 1986). The technique was developed to provide a more resilient approach to non-independent residuals and clustered data commonly found in hierarchical and longitudinal datasets. The technique has seen some application in the field of innovation studies, often instead of a mixed effects model (Katila and Ahuja, 2002; Di Gregorio and Shane, 2003). More specifically, the MI-GEE approach where generalized estimating equations are used as a tool of inference in the presence of multiple imputed longitudinal data, given the historical effectiveness of GEE models with multiple imputed datasets (Shen and Chen, 2013). Additionally I find the technique is quite resilient when dealing with unorthodox data structures, such as this combination of panel and cross sectional data used here. The key difference between GEE models and more common mixed effects models (commonly called panel data models), is that GEE models fit the marginal mean models, while mixed effects and generalized mixed models fit the conditional means. This means that GEE models are interested in estimating the average effects over the whole population the sample represents, and not in subject-specific changes. When modeling linear relationships (Gaussian response variables) the results are nearly identical (Hubbard et al., 2010). Due to the fact that some firms have exited during the period from 2010-2015, it was necessary to deal with these firms in the imputation procedure.

Figure 5.5: Density comparisons prior (black)- and post- (red) imputation



Prior to imputation, all firms that were marked as having potentially exited during the creation of the survival variables were excluded from the analysis. *This is a precautionary measure that prevents multiple imputation of yearly values during years where the firm in all likelihood no longer existed.* Therefore, the results of the GEE models portray the response variable regressed on the explanatory variables *conditional on firm survival until at least 2015.*

The nature of time series cross sectional data makes it necessary to forgo certain assumptions of more commonly employed generalized linear models, including OLS. Namely, that the residuals are not serially correlated, and that the covariance of the errors is a stochastic process. Since the same units are observed in subsequent time periods, it is natural to assume that there will be some correlation among the errors. This is a type of autocorrelation, or, a non-random degree of association between the current and past values of a series. It is often advisable try and find a certain type of specified autoregressive or moving average function (ARMA) that can at least discernibly be related to the residuals of the data. I assess this using Durbin-Watson tests and by mapping the autocorrelation and partial autocorrelation functions of the time series data (ACF and PCF) (Fox and Weisberg, 2010). The Durbin-Watson statistics are computed on an OLS regression using the same data to be used in time series regression, then the autocorrelation points may be better identified.[7] Upon performing this for Models 3 and 4, it was found that autocorrelation was statistically significant in each of the 5 time periods. The pattern is quite similar in all 4 models, and while it is generally not possible to with certainty identify which type of ARMA process the data actually has, it can be fruitful to look for certain patterns that increase the likelihood of certain processes as opposed to others. In terms of specifying the autocorrelation or moving average process, it is useful to recognize the different patterns taken on by different processes. The AR(1) process has a PCF with a decaying exponential function of autocorrelation and a singular spike at the first lag (Fox and Weisberg, 2010). Moving average (MA) processes often connote a shifting positive-negative trend in the ACF and PCF quite early in the lag process. With the visualizations below, it is at quite probable that an AR(1) process will fit the estimation specification. Thus, AR(1) is specified in all generalized estimating equations to follow for H2 and H3. Figures 8.12 and 8.13 in the appendix show the ACF and PCF functions of each of the dependent variables `logEmp` and `logOpRev`.

[7] In this case, for the logged firm size variable the test statistics was 1.5961, and for logged operating revenue, 1.5108, attesting that there is autocorrelation at the $p<0.0001$ level.

Table 5.5: Model 3.1 - `logEmp` conditional on survival (intercept) regressed on internal knowledge intensity

	I MI-GEE	II MI-GEE	III MI-GEE	IV MI-GEE	V MI-GEE	VI MI-GEE	VII MI-GEE
EmpEdu	0.417^{***} (0.119)					0.470^{***} (0.117)	0.452^{***} (0.117)
EmpHiEdu	-0.271^{*} (0.132)					-0.287^{*} (0.129)	-0.269^{*} (0.128)
FoundEdu	0.0103 (0.0220)					−0.00659 (0.0217)	−0.0110 (0.0216)
FoundEnt		0.119^{**} (0.0419)				0.0995^{*} (0.0412)	0.101^{*} (0.0411)
FoundInd		0.000148 (0.00239)				−0.00154 (0.00242)	−0.00180 (0.00240)
FoundUni		0.129 (0.113)				0.110 (0.111)	0.113 (0.110)
AgeMax		0.0415 (0.0288)				0.0300 (0.0278)	0.0283 (0.0276)
KDisp			-0.404^{***} (0.118)			-0.313^{**} (0.115)	-0.592^{***} (0.130)
KScope			0.379^{***} (0.0914)			0.298^{**} (0.0909)	1.921^{***} (0.367)
KScope^2							-1.793^{***} (0.389)
Spinoff				0.275^{***} (0.0614)		0.247^{***} (0.0604)	0.234^{***} (0.0599)
FF1					0.0626^{***} (0.0125)	0.0600^{***} (0.0125)	0.0610^{***} (0.0125)
FF2					$0.0284^{\dagger}$ (0.0158)	0.0248 (0.0162)	0.0259 (0.0161)
FF3					-0.0584^{***} (0.0175)	-0.0524^{**} (0.0173)	-0.0517^{**} (0.0172)
FirmAge	0.0119 (0.00901)	0.0118 (0.00911)	0.0103 (0.00895)	0.0117 (0.00895)	0.0141 (0.00896)	0.00665 (0.00887)	0.00535 (0.00881)
logit(R&DInt)	0.0128 (0.0145)	0.0145 (0.0145)	0.0170 (0.0142)	0.0193 (0.0142)	0.0151 (0.0146)	0.00546 (0.0147)	0.00475 (0.0147)
(Intercept)	2.241^{***} (0.177)	2.178^{***} (0.170)	2.430^{***} (0.168)	2.350^{***} (0.160)	2.329^{***} (0.160)	2.226^{***} (0.187)	2.129^{***} (0.186)
Sectoral controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Regional controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	9955	9955	9955	9955	9955	9955	9955
Groups	1991	1991	1991	1991	1991	1991	1991
est. QIC	12805	12860	12837	12821	12722	12550	12508

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

5.4 Results

5.4.1 Model 3 - Internal knowledge intensity as affecting business performance

5.4.1.1 Model 3.1

Table 5.5 displays the results of the regressions of Model 3.1. Again, as the coefficients can sometimes be difficult to interpret, I rely mainly on effects plots to aid in this. These plots are available in the appendices, Figure 8.14.

Moving through the stepwise specifications of the model, one can see that direction of effects and overall significance of the variables remains relatively stable throughout. According to the models, the more employees that have completed a tertiary education is clearly positively associated to the size of the firm, while having a PhD among employees is negatively associated with it. The level of founder educational attainment actually negatively affects the size of the firm in the single imputation chosen for analysis, but in the average statistics reported in the regression table, there is no significant effect. Entrepreneurial experience has a positive association, but university experience, years of industry experience and the maximum age of the founder, while promising in the graphical display of the single imputation, do not have significant effects in the model. Knowledge disparity has a negative and statistically significant effect on firm growth, while knowledge scope indeed positively affects it, before a marginally declining relationship sets in. Spinoffs also seem to be more likely to experience higher levels of firm size in terms of employees over time. In terms of the 3 formation factor variables: `FF1` (opportunities derived from technical change, etc.) is positively associated and highly statistically significant in this model, while `FF2` is only partly significant, and `FF3` is negatively associated with firm size over time.

Before proceeding with additional diagnostics, I will briefly comment on the robustness check performed pre-imputation in order to validate the interpretation and the use of the MI-GEE method I have chosen. The first step was to simply regress each yearly value of the `logEmp` variable on the full specification of the independent variables. This regression can be found in Figure 8.12 in the appendix in the first 6 columns. By doing this, I find that for any given year, the same direction of effects and statistical significance levels largely persist for most of the explanatory variables as were found in the MI-GEE model above. The number of observations has been moreover reduced by about two-thirds. Not surprisingly, some variables experience a waning of statistical significance and coefficient worth as years progress, suggesting that over time, the conditions present at the start of the firm (as captured by the survey in 2010-2011) affect the size of the firm each given year less and less in some cases. The important thing to convey is that pre-imputation, the OLS models work reasonably well any given year of the response variable. Next I compose a GEE model based on the pre-imputed data in the same was as was done post imputation. This table can be seen in the appendix, Table 8.8, column I, and Table 8.9. Here again, if the reader reviews, it is apparent that the models are very similar, though the statistical significance of the coefficients of the MI-GEE model above is stronger than in the pre-imputed model. This is with 1306 observations instead of the 1991 above. To me this serves as compelling evidence that the imputed data model above not only captures the essence of the statistical models had they not been imputed, but improves upon them.

For diagnosis of model 3.1 (and those to follow also employing multiple imputed generalized estimating equations) I address the fully specified model including polynomials, specification VII in Table 5.5. I rely on residual plots of an extracted imputation of each one, randomly selected, for which I perform a generalized linear model GEE (a "GEEGLM") using the R package `geepack` (Højsgaard et al., 2006) with the same correlation structure specifications. For these plots I rely on the "working" residuals, the last ones computed in the iteration process that estimated the GEE model. The plots (Figure 8.15 in appendix) indicate minor subtle deviations from the null plot line on several variables. Though the output does not automatically generate the lack-of-fit tests, I am able to construct them manually. Tukey's test of non-additivity on the residuals revealed a significant p-value, indicating a possibly inadequate fit.[8] Performing a lack-of-fit test on the model yields that several models test positive for the addition of a quadratic term: namely, the 1st and 3rd formation factor (`FF1` and `FF3`), percentage of employees holding a tertiary degree (`EmpEdu`), and the level of the founder's educational attainment (`FoundEdu`).

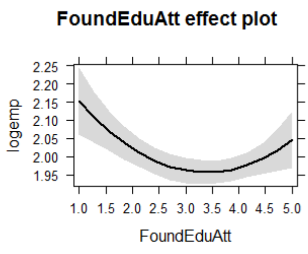
Although this does not give us information about the whole range of imputed datasets, it serves to aid in diagnosing the presence of major problems with residual correlation with the variables or the fitted values (beyond the autocorrelation in the clusters which we have already corrected for with the GEE model). The corresponding residual plots can be seen on the next page below (the red line is a loess fit).

[8] For 3.1, Tukey's test statistic =3.106 at $p<0.0019$.

Table 5.6: Lack-of-fit tests for 3.1

	lack of fit	P-value
FF1	3.28	0.00
FF2	0.30	0.76
FF3	2.44	0.01
Knowledge Disparity	0.79	0.43
Employee (tertiary) ed.	4.20	0.00
Employee (PhD) ed.	1.21	0.23
Founder ed. att.	7.12	0.00
Founder ind. exp.	0.06	0.95
Founder max age	0.05	0.96
R&D Int.	1.58	0.12

Figure 5.6: Quadratic effect plot of founder educational attainment for 3.1



Updating the regression with the added squared terms for the 4 variables pointed out above which had significant p-values in the lack of fit test yields the following output for the averaged full imputation model. The single imputation GEEGLM model of the same regression has a quasi-likelihood under the independence model criterion[9] (QIC) with a slightly lower value than that of the original 3.1 model (12508 > 12486 QIC), indicating that the modified model fits the data better for that single imputation, though perhaps marginally so. Indeed, when adding the squared terms to the MI-GEE model, which averages across all 100 imputations, the only added squared terms that are significant that were not in the original model is that of the Founder's educational attainment, which shows a p<.01 significance when averaged across all 100 imputations, and `EmpEdu`, which had a positive (p<.10) squared term. For reasons of space I will not show the full results of the modified model, as this addition does not markedly change anything about the results reported above. This shows us that while 1 particular imputation might find more nuanced results, the average effect across all imputations might be marginally significant. Nonetheless, there is a positive and statistically significant relationship between the quadratic term of the Founder's educational attainment and the size of the firm in terms of number of

[9]Analogous to the AIC and BIC for use of testing the fit of quasi-maximum likelihood estimated models (See Pan 2001 for theoretical details).

employees. Suggesting that while the linear term is not significant, as the variable increases in size it becomes statistically significant and positive. The effect plot to the right shows the shape of the effect from the single imputed dataset mixed model equation with the squared term for `FoundEdu` included. This suggests that founder educational does not matter so much until the level is moderate to high, which does positively affect firm size over time.

5.4.1.2 Model 3.2

I turn now to Model 3.2 in Table 5.7: For the natural log of operating revenue over the 5 year period, the effects are similar in some respects. Employee education at the tertiary level has an effect at the 5% level, while entrepreneurial experience of the founder and the firm being a formal spinoff also have an effect at this level of significance. The effect of knowledge scope takes the same shape and direction of effects as the previous model, with a positive yet negatively curvilinear effect (at least at the 10% level for the latter component). Knowledge disparity again is negatively associated with the response variable, though only at the 5% level. The roles of `FF1` and `FF2` are nearly reversed, with `FF2` (work experience and networks influential in forming the company) being the most statistically significant formation factor component. `FF1` is only positive and significant at the 10% level, suggesting that growth and profitability may be driven by different entrepreneurial motives as materialized by these factors. `FF3` has no discernable effect on operating revenue over the period. R&D Intensity seems to be negatively related to operating revenue over time at the 1% level. The lagged value of logged firm size/growth was controlled for in the regression, producing an extremely significant and sizable coefficient compared with the rest of the model's coefficients. This suggests that much of the profitability, unsurprisingly, depends on the firm's size and growth trajectory from the previous year. If one observes the effects plots, one can see that the effects of the variables vary a bit over the data range of the explanatory/response variables. For the most part, they seem well suited to represent the regressions above visually (See Figure 8.16 in the appendix).

Table 5.7: Model 3.2 - `logOpRev` conditional on survival (intercept) regressed on internal knowledge intensity

	I MI-GEE	II MI-GEE	III MI-GEE	IV MI-GEE	V MI-GEE	VI MI-GEE	VII MI-GEE
EmpEdu	0.502*					0.572**	0.551**
	(0.208)					(0.209)	(0.209)
EmpHiEdu	−0.111					−0.122	−0.0995
	(0.236)					(0.236)	(0.236)
FoundEdu	0.0222					−0.000870	−0.00645
	(0.0389)					(0.0394)	(0.0394)
FoundEnt		0.221**				0.200**	0.203**
		(0.0725)				(0.0729)	(0.0728)
FoundInd		0.00353				0.000626	0.000302
		(0.00423)				(0.00427)	(0.00426)
FoundUni		0.00974				−0.0532	−0.0491
		(0.237)				(0.235)	(0.234)
AgeMax		0.000323				−0.0110	−0.0131
		(0.0495)				(0.0487)	(0.0487)
KDisp			−0.639**			−0.538**	−0.894***
			(0.204)			(0.202)	(0.228)
KScope			0.629***			0.534***	2.597***
			(0.161)			(0.162)	(0.647)
$KScope^2$							−2.278***
							(0.684)
Spinoff				0.364***		0.322**	0.307**
				(0.103)		(0.105)	(0.104)
FF1					0.0526*	0.0480*	0.0495*
					(0.0227)	(0.0228)	(0.0227)
FF2					0.0920**	0.0791**	0.0806**
					(0.0288)	(0.0292)	(0.0292)
FF3					−0.0702*	−0.0608†	−0.0601†
					(0.0322)	(0.0323)	(0.0322)
FirmAge	0.0200	0.0224	0.0165	0.0194	0.0237	0.0160	0.0144
	(0.0154)	(0.0158)	(0.0154)	(0.0154)	(0.0154)	(0.0157)	(0.0156)
logit(R&DInt)	−0.0488†	−0.0393	−0.0378	−0.0339	−0.0324	−0.0506†	−0.0515†
	(0.0269)	(0.0271)	(0.0261)	(0.0263)	(0.0271)	(0.0281)	(0.0282)
L.log(Emp)	0.276***	0.276***	0.274***	0.272***	0.272***	0.262***	0.258***
	(0.0456)	(0.0461)	(0.0456)	(0.0457)	(0.0456)	(0.0447)	(0.0445)
(Intercept)	5.344***	5.417***	5.678***	5.568***	5.568***	5.490***	5.376***
	(0.297)	(0.281)	(0.288)	(0.272)	(0.272)	(0.318)	(0.319)
Sectoral controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Regional controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	7964	7964	7964	7964	7964	7964	7964
Groups	1991	1991	1991	1991	1991	1991	1991
est. QIC	25121	25134	25091	25151	25087	24993	24956

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Similar to Model 3.1, here I have also performed pre-imputation robustness checks on the data and models. In the appendix the reader may find the OLS regressions on the logged values of Operating Revenue for 2010 - 2014 in table 8.13 columns I-V, and the GEE regressions in Tables 8.8 and 8.10. The results of the OLS trials give conflicting results for the `EmpEdu` variable, which appears negative but non-significant. Other than this, there are no major differences in direction of effects between the OLS trials and the MI-GEE model 3.2. A few variables appear significant in certain iterations, such as university experience of the founder in 2013 and 2014, but since they are not significant in the imputed model no comparison can be made. The GEE model pre-imputation shows significant coefficients for knowledge disparity and scope, as well as founder entrepreneurial experience. In both checks, it can be seen that variables are on average less significant pre-imputation, but since most of the directions of coefficients are the same, it seems to have improved upon the model. Though obviously this variable behaves less well than the number of employees natural logarithm when comparing the imputed and non-imputed models.

Again, I plot the working residuals against the explanatory variables and the fitted values of the model in Figure 8.17 in the appendix, and construct lack-of-fit and Tukey tests for the components of the model in Table 5.7.

The graphs indicate a fairly stable trend of null relationships, though there are some minor deviations, though R&D Intensity as well as the founder's educational attainment, year of industry experience, `FF1` and `FF3` are significant in the lack of fit tests. I construct Tukey tests for non-additivity for the fitted values manually, and it did not reveal a lack of fit.[10] I rerun the multiple imputation GEE model for 3.1 with the additional polynomials, but none are significant at the 5% level or below. Thus I leave the model 3.2 unaltered after diagnostics.

Table 5.8: Lack-of-fit tests for 3.2

	l.o.f. statistic	P-value
FF1	3.46	0.00
FF2	0.72	0.47
FF3	0.00	1.00
Knowledge Disparity	1.05	0.29
Employee (tertiary) ed.	0.00	1.00
Employee (PhD) ed.	0.05	0.96
Founder ed. att.	2.94	0.00
Founder ind. exp.	3.33	0.00
Founder max age	0.25	0.80
R&D Int.	14.65	0.00

[10] For Model 3.2.2, Tukey's test statistic = 0.0868 at p<0.9308.

5.4.1.3 Model 3.3 - Survival analysis

I now proceed with the Cox proportional hazard model on Internal Knowledge Intensity. As expressed in Model 2.2, I omit the variables `EmpHiEdu` and `FoundEdu` due to their high correlation with `EmpEdu` resulting in obscured effects. I obtain first a fit of the survival frequency of the firms, and its pattern over the lifespan of each firm:

Table 5.9: Survival fit descriptives: 3.3 (& 4.3)

	Time	Survival	Deaths	Survival SE	Cumulative Events	Number remaining
1	5.00	0.96	0.04	0.01	13.00	366.00
2	6.00	0.93	0.07	0.01	45.00	1018.00
3	7.00	0.92	0.08	0.01	70.00	1294.00
4	8.00	0.90	0.10	0.01	106.00	1595.00
5	9.00	0.86	0.14	0.01	166.00	1563.00
6	10.00	0.84	0.16	0.01	188.00	907.00
7	11.00	0.82	0.18	0.01	215.00	1274.00
8	12.00	0.79	0.21	0.01	247.00	955.00
9	13.00	0.78	0.22	0.01	258.00	645.00
10	14.00	0.74	0.26	0.01	289.00	634.00

Figure 5.7: Left truncated survival fit 2.3/4.3

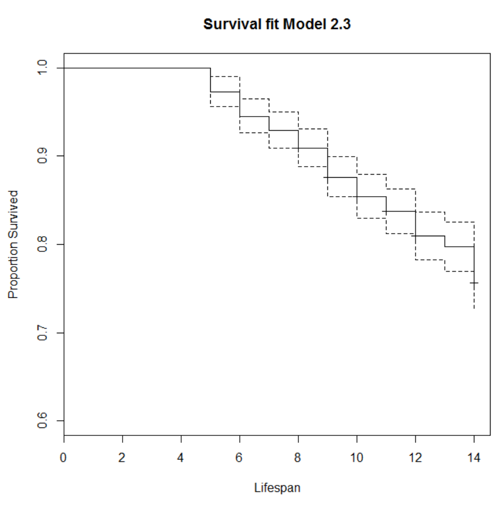


Table 5.10: Model 2.3 - Cox survival analysis based on internal knowledge intensity (conditional on survival up to survey date i.e. 2010)

	I Cox PH	II Cox PH	III Cox PH	IV Cox PH	V Cox PH	VI Cox PH	VII Cox PH	VIII Bin. Logit
EmpEdu	-0.5312^{*}					$-0.5228^{\dagger}$	$-0.5195^{\dagger}$	-0.6379^{**}
	(0.2704)					(0.2716)	(0.2712)	(0.2353)
FoundEnt		−0.2396				$-0.2767^{\dagger}$	$-0.2816^{\dagger}$	−0.1422
		(0.1493)				(0.1506)	(0.1508)	(0.1250)
FoundUni		−0.5044				−0.4616	−0.4705	−0.1610
		(0.5107)				(0.5143)	(0.5144)	(0.3753)
AgeMax		−0.0976				−0.0962	−0.0889	−0.0923
		(0.0782)				(0.0792)	(0.0793)	(0.0783)
KDisp			0.4948			0.5318	1.0690^{*}	0.9685^{*}
			(0.4041)			(0.4050)	(0.4717)	(0.4058)
KScope			−0.2693			−0.2608	-3.1404^{*}	$-1.9619^{\dagger}$
			(0.3214)			(0.3230)	(1.3396)	(1.1522)
KScope^2							3.1546^{*}	$2.1441^{\dagger}$
							(1.4154)	(1.2147)
Spinoff				0.1703		0.2404	0.2578	0.3037^{*}
				(0.1787)		(0.1814)	(0.1819)	(0.1539)
FF1					−0.0053	−0.0063	−0.0100	0.0194
					(0.0401)	(0.0406)	(0.0408)	(0.0359)
FF2					−0.0695	−0.0611	−0.0622	−0.0256
					(0.0504)	(0.0510)	(0.0510)	(0.0471)
FF3					0.0131	−0.0053	−0.0068	−0.0160
					(0.0570)	(0.0580)	(0.0580)	(0.0504)
FoundInd								−0.0013
								(0.0062)
logit(R&DInt)	−0.0327	−0.0380	−0.0484	−0.0477	−0.0508	−0.0261	−0.0239	−0.0058
	(0.0477)	(0.0479)	(0.0476)	(0.0473)	(0.0486)	(0.0498)	(0.0498)	(0.0414)
log(Emp + 1)	−0.1021	$-0.1190^{\dagger}$	$-0.1204^{\dagger}$	-0.1358^{*}	$-0.1249^{\dagger}$	−0.0861	−0.0724	−0.0528
	(0.0646)	(0.0646)	(0.0646)	(0.0643)	(0.0648)	(0.0663)	(0.0666)	(0.0565)
time1								-0.0574^{*}
								(0.0265)
(Intercept)								-1.4336^{**}
								(0.5059)
Sectoral controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Regional controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	2289	2289	2289	2289	2289	2289	2289	2458
Wald p	0.0096	0.0128	0.0307	0.0244	0.0411	0.0184	0.0077	
R^2	0.0097	0.0106	0.0087	0.0084	0.0089	0.0145	0.0166	
AIC								2246.3586
BIC								2757.3837
$\log L$								−1035.1793

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Looking at Table 5.10, we can see that the initial fit of the model shows that a stepwise approach is somewhat unhelpful. Only `EmpEdu` appears significant prior to running the more complex model specifications VI and VII. These reveal a weak negative association ($p<.10$) between founder entrepreneurial experience and firm exit. When quadratic polynomial of knowledge scope is added, some weak results begin to appear for knowledge scope, its quadratic term, being a corporate spinoff, and knowledge disparity. Knowledge disparity seems positively associated with firm exit, while knowledge scope is negatively associated with firm exit ($p<.05$), and a positive marginal relationship exists ($p<.05$) in specification VII. Overall, it shows quite weak association between the explanatory and response variables on a whole. Column VIII depicts the results of the binary logit regression, where 1 is the outcome where the firm has exited, and 0 otherwise, as the response variable. Results are similar, with varying power for most variables. Effects plots confirm these relationships graphically (See figure 8.18 in Appendix)

It is beneficial to check if the assumptions of the model are being violated by any particular variables. The proportional hazards assumption is crucial to the specification of the Cox regression model, and variables not behaving in accordance with this can give radically skewed results. For this, I use the proportional hazards test. This test is based on the correlations between the scaled Schoenfeld (1982) residuals and a transformation of the time variables (which defaults to the Kaplan-Meier (1958) estimate). The procedure is quickly and easily done using the `survival` package in `R`. Table 5.11 below conveys the result of this test for the model, while Figure 8.19 visualizes the test. Ideally the lines in the figure should be null plots with no vertical deviation, substantial deviation from 0 indicates potential violation of model assumptions. As can be seen by the lack of significant p-values of the test (and in the graph on the following page), no systematic violations of the assumptions are apparent. I also conducted tests of linearity of the covariates using Martingale residuals and partial residual plots (not shown), but no non-linear relationships could be detected.

5.4.1.4 Summary of effects and hypothesis confirmation - Model 3.

Hypothesis 3.1 suggested that higher levels of stocks of human capital in the form of education of founders and employees would positively associated with business performance . We found that higher levels of tertiary educated employees in the venture did positively impact both the number of employees and in financial turnover during the 5 year period,

Table 5.11: Proportional hazards test of 3.3

	rho	chisq	p
EmpEdu	0.0591	0.7908	0.3739
FoundEnt	0.0442	0.5083	0.4759
FoundUni	0.0700	1.2252	0.2683
AgeMax	-0.0192	0.1024	0.7489
KDisp	-0.0292	0.2135	0.6440
KScope	-0.0063	0.0099	0.9207
KScope^2	-0.0034	0.0029	0.9567
Spinoff	0.0264	0.1834	0.6685
FF1	-0.0816	1.5254	0.2168
FF2	-0.0022	0.0011	0.9730
FF3	-0.0564	0.7789	0.3775
logit(R&DInt)	0.0109	0.0347	0.8521
log(Emp + 1)	0.0198	0.1111	0.7389
GLOBAL		14.5706	0.7495

while survival/ exit were weakly associated with this variable. Higher levels of MSc and PhD educated employees were negatively associated with at least firm size during the 5 year period. Educational attainment of the founder did not seem to have any significant effect on business performance.

It seems then that employee education and founder experience are among the most important predictors of business performance among these variables. I do not find any effect supporting claims that formal founder education plays a significant role in shaping business performance, or that older founding teams perform better or worse than younger ones. University experience surprisingly has no discernable effects either from this sample. The third formation factor `FF3`, the importance of design and technical knowledge of the founding team (`FF3`), was negatively related to firm size, but not to any other response variables in the models. We must then settle for only partial confirmation of H-3.1.

Concerning Hypothesis 3.2, it is noteworthy that the presence of entrepreneurial experience in the founding team positively associated with all 3 models (at the 5% level for growth of employees and turnover, and at the 10% level for survival). University experience, years of industry experience, and increased age of the founding team did not affect business performance in any significant way in the models.This is also the case for importance of experience of networks and the industry (`FF2`), the effects are discernibly positive for firm size and revenue over time, but not for survival.

Hypothesis 3.3a and 3.3b suggested that knowledge scope of the founding team would be linearly, as well as inverse curvilinearly, related to business performance, while 3.4 predicted that knowledge disparity would be negatively related to business performance. These effects are indeed present in the multiple imputed generalized estimating equation-based models for firm size and firm financial performance in terms of operating revenue. Additionally, we found that firm's exit (that is, the 1 value of the `Surv` variable used in the Cox models) was positively influenced by the founding team's knowledge disparity, as well as by knowledge scope and its squared term (at least at the p<.05 level for the latter). These results seem to, for the most part, confirm our hypotheses 3.3a, 3.3b, and 3.4.

Hypothesis 3.5 proposed that being a formal spinoff should positively impact the performance of the new venture. This seems to be true for number of employees and operating revenue, while the variable actually did not seem to systematically associate with survival. So there is only partial confirmation for H-3.5

Hypothesis 3.6: Reacting to opportunities brought about by novelty in the system, that is, through technical change, market needs, or regulatory or institutional changes (`FF1`) the effects are discernibly positive for firm size and revenue over time, but not for survival.

5.4.1.5 Firm levels vs. firm growth relating to internal knowledge intensity

While the response variables above measure the amount of employees and operating revenue each year, they do not have any connotation to the actual level of growth from year to year, or over the period. Indeed, while I have used an operationalization of firm growth that is established in the literature, using the values of the number of employees while controlling for firm age (Colombo and Grilli, 2005) it is not the more common one. In latent growth curve modeling language, my measure might be referred to as the "intercept value" of firm growth, while the traditional measure, denoting a change from time *t-1* to time *t*, would be referred to as the "slope value" of firm growth (Bollen and Curran, 2006). With this in mind, I have also attempted to regress the natural log difference between 2014 values and 2010 values of the response variables, in order to see if the explanatory variables might affect the growth of the firm in terms of size and revenue. These regression specifications may be found in the appendix for the data pre-imputation. Post-imputation calculation growth rates is unreliable due to the fact that the imputation process of the EMB algorithm used focuses not on obtaining the individual missing values themselves, but by plugging in values multiple times to "preserve important characteristics of the dataset .... [m]ultiple imputation preserves means, variances, covariances, correlations, and linear regression coefficients", but it cannot necessarily adequately be used to, post-imputation, predict changes from year to year in a panel set

(Graham, 2009: 559). The pre-imputation growth rates are constructed by taking the difference of logged 2014 and 2010 values, the results yielded no significant coefficients for Number of Employees, and only a negative association with founders having university experience for Operating Revenue (along with a significant control variable for size of the firm). These models can be found in the appendix, Tables 8.14 and 8.15, for reference.

5.4.2 Model 4 - Innovative performance as affecting business performance

5.4.2.1 Models 4.1 and 4.2

We now move on to the results and interpretation of Model 4, starting with the MI-GEE regressions on the response variables `logEmp` and `logOpRev`. Table 5.12 shows the regression results of 4.1 in all specifications: Here there are some obvious issues with the specification, as it appears that including the independent variables in the same regression equation produces varying results in terms of significance and direction of coefficients. Specification I finds a highly statistically significant ($p<.01$) coefficient for `InnoGoods`, while II finds the same ($p<.05$ admittedly) for `InnoServ`. Combining the two in specification III gives around the same results, with the `InnoServ` coefficient degrading to the $p<.10$ level of statistical significance. `RadInnOS` on its own (specification VI) is highly significant as well ($p<.001$). However, when all are added to the model, only `RadInnOS` retains significance ($p<.05$). Controlling for corporate spinoff status yields stable results, and the age of the firm is weakly significant. The `RadInnOS` variable was earlier shown to correlate quite substantially with the other two variables (0.54 for goods, and 0.51 for services), so it seems that using these variables in the same specification results in potentially multiple collinearity problems. All variables seem to have an association with the response variable in some form however, suggesting that innovative performance may associate strongly (in a statistical sense) with firm size over time.

Moving on to Model 4.2 in Table 5.13, we see a near reflection of relative signs and statistical significance levels among the 3 primary explanatory variables. Again, the `RadInnOS` variable has the strongest effect both economically and statistically, and when it is included with the other two variables it usurps any significance they had. Here the R&D Intensity control negatively associated with firm revenue over time.

Once again I rely on plotting the effects of these regressions for

Table 5.12: Model 4.1 - `logEmp` conditional on survival (intercept) regressed on innovative performance

	I MI-GEE	II MI-GEE	III MI-GEE	IV MI-GEE	V MI-GEE
logit(InnoGoods)	0.0342^{**}		0.0321^{**}		0.0108
	(0.0113)		(0.0113)		(0.0144)
logit(InnoServ)		0.0213^{*}	$0.0179^{\dagger}$		−0.00362
		(0.0104)	(0.0103)		(0.0138)
RadInnOS				0.0962^{***}	0.0888^{*}
				(0.0228)	(0.0360)
Spinoff	0.267^{***}	0.265^{***}	0.269^{***}	0.261^{***}	0.262^{***}
	(0.0583)	(0.0583)	(0.0583)	(0.0578)	(0.0579)
FirmAge	$0.0142^{\dagger}$	$0.0136^{\dagger}$	$0.0145^{\dagger}$	0.0126	0.0129
	(0.00818)	(0.00821)	(0.00818)	(0.00812)	(0.00806)
logit(R&DInt)	−0.00685	−0.000173	−0.0106	−0.00996	−0.0117
	(0.0148)	(0.0148)	(0.0153)	(0.0151)	(0.0153)
(Intercept)	2.376^{***}	2.402^{***}	2.401^{***}	2.227^{***}	2.235^{***}
	(0.151)	(0.152)	(0.152)	(0.156)	(0.168)
Sectoral Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Regional Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	10870	10870	10870	10870	10870
Groups	2174	2174	2174	2174	2174
QIC	14528	14567	14525	14497	14502

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 5.13: Model 4.2 - `logOpRev` conditional on survival (intercept) regressed on innovative performance

	I MI-GEE	II MI-GEE	III MI-GEE	IV MI-GEE	V MI-GEE
logit(InnoGoods)	0.0560^{**}		0.0515^{**}		0.0238
	(0.0198)		(0.0199)		(0.0248)
logit(InnoServ)		0.0439^{*}	0.0385^{*}		0.0106
		(0.0171)	(0.0172)		(0.0215)
RadInnOS				0.154^{***}	0.115^{*}
				(0.0385)	(0.0580)
L. log(Emp)	0.267^{***}	0.268^{***}	0.267^{***}	0.266^{***}	0.265^{***}
	(0.0430)	(0.0429)	(0.0428)	(0.0428)	(0.0428)
FirmAge	$0.0256^{\dagger}$	$0.0247^{\dagger}$	$0.0262^{\dagger}$	0.0229	$0.0241^{\dagger}$
	(0.0145)	(0.0145)	(0.0145)	(0.0144)	(0.0144)
logit(R&DInt)	-0.0653^{*}	-0.0567^{*}	-0.0734^{**}	-0.0698^{**}	-0.0748^{**}
	(0.0257)	(0.0256)	(0.0263)	(0.0258)	(0.0262)
Spinoff	0.364^{***}	0.360^{***}	0.368^{***}	0.354^{***}	0.359^{***}
	(0.0976)	(0.0979)	(0.0977)	(0.0974)	(0.0974)
(Intercept)	5.631^{***}	5.681^{***}	5.685^{***}	5.395^{***}	5.471^{***}
	(0.261)	(0.262)	(0.261)	(0.265)	(0.277)
Sectoral Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Regional Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	8696	8696	8696	8696	8696
Groups	2174	2174	2174	2174	2174
QIC	27756	27745	27732	27712	27719

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

interpretation. The plots for 4.1 and 4.2 can be seen in Figures 8.20 and 8.22 in the appendix. Again, I diagnose the models by carefully looking at the residual plots taken from the GEEGLM model on a single imputation (also Figures 8.21 and 8.23 in appendix). All are null plots more or less. Tukey's test for non-additivity comes up also significant for 4.1, but not for 4.2.[11] Finally, I perform the lack-of-fit tests for each of the models' variables:

Table 5.14: Lack of fit tests for 4.1 and 4.2

	lof	Pvalue
4.1.innogoods	0.33	0.74
4.1.innoserv	1.70	0.09
4.1.radinn	0.56	0.58
4.1.R&DInt	0.84	0.40
4.2.innoggoods	2.05	0.04
4.2.innoserv	6.26	0.00
4.2.radinn	26.32	0.00
4.2.R&DInt	8.84	0.00

4.1 has no violations at the $p<.05$ level, while in 4.2, all of the four variables seem to be potentially miss-specified. Remodeling the regression 4.2 to include quadratic terms for `InnoGoods`, `RadInnOS`, and `R&D Int` yields statistically significant relationships for the latter two but not the first one. I then attempt to remodel the multiple imputation GEE model for 4.2 with the latter polynomial terms included. Since I already found some variation in the specifications depending on which explanatory variables are included, I run them stepwise. However, no curvilinearity is apparent in any specification for `InnoGoods` or `InnoServ`. `RadInnOS` does appear curvilinear in all specifications that include it. The results are shown in Table 5.15. By viewing the plot of `RadInnOS` effects (Figure 8.22 in appendix), we can see that it is beneficial for firm revenue growth to at least reach new to the firm innovations, but implementing new to the market or new to the world innovations has declining marginal benefits. R&D Intensity has a negative association with revenue growth.

As in Model 3, I have performed pre-imputation robustness checks on Model 4. These regressions can be found in the appendix: The fully specified GEEGLM regressions can be found in Table 8.8 specifications III and IV, and all specifications of this modeling can be found in Table 8.11; also, the pre-imputation OLS regressions of response variables `logEmp` and `logOpRev` for each year from 2010-2014 can be found in Table 8.12. The conclusions drawn from these models are similar to that of Model 3's robustness checks.

[11] 4.1: Tukey test statistic = 1.9813, $p<0.0476$; 4.2: Tukey test statistic 0.783, $p<0.434$.

Table 5.15: Model 4.2 modified

	(I) logOpRev	(II) logOpRev	(III) logOpRev	(IV) logOpRev	(V) logOpRev
logit(InnoGoods)	0.0300		0.0324		-0.00100
	(0.0277)		(0.0278)		(0.0298)
logit(InnoGoods)^2	-0.0127		-0.0106		-0.00590
	(0.00811)		(0.00862)		(0.00871)
logit(InnoServ)		0.0320	0.0343†		-0.00986
		(0.0197)	(0.0196)		(0.0231)
logit(InnoServ)^2		-0.00558	0.00219		0.0102
		(0.00609)	(0.00662)		(0.00695)
RadInnOS				0.375***	0.444***
				(0.0824)	(0.117)
RadInnOS^2				-0.122**	-0.141**
				(0.0420)	(0.0465)
L. log(Emp)	0.268***	0.270***	0.268***	0.270***	0.269***
	(0.0429)	(0.0430)	(0.0428)	(0.0430)	(0.0429)
FirmAge	0.0236	0.0237	0.0247†	0.0217	0.0214
	(0.0144)	(0.0145)	(0.0144)	(0.0144)	(0.0144)
logit(R&DInt)	-0.0825*	-0.0801*	-0.0885*	-0.0869*	-0.0873*
	(0.0392)	(0.0388)	(0.0398)	(0.0382)	(0.0394)
logit(R&DInt)^2	-0.00773	-0.00900	-0.00719	-0.00684	-0.00688
	(0.00978)	(0.00971)	(0.00980)	(0.00964)	(0.00974)
Spinoff	0.360***	0.360***	0.364***	0.359***	0.357***
	(0.0973)	(0.0977)	(0.0975)	(0.0971)	(0.0970)
(Intercept)	5.708***	5.715***	5.729***	5.385***	5.278***
	(0.260)	(0.264)	(0.262)	(0.266)	(0.292)
N	8696	8696	8696	8696	8696

Standard errors in parentheses
$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

5.4.2.2 Model 4.3 - Survival

The survival model for 4.3 is carried out similarly to Model 3.3, and are left-truncated models due to the firms having a given survival rate until at least 2010 when they were surveyed. I use the same fit commands to produce the same types of tables and graphs. Again I return an initial regression model after first fitting the survival function. Both can be seen below in Table 5.16. The survival curve is identical to that found in Model 3.3, and can be seen in Figure 5.7 in the previous section. This time, before interpretation I try to diagnose the proportional hazards assumption's intactness for the variables of the model, but find no variables in violation. So, I leave the variables untransformed. Table 5.16 shows the test of hazards, and the Schoenfeld residual plots for the violating variables in are in the appendix, Figure 8.24. Here, no specifications yield significant coefficients until all 3 explanatory variables of interest are added to the model. Once all 3 are present, we see a positive association with `InnoGoods` and a negative association with `RadInnOS` to the dependent variable, firm exit. This may be attributable to the idea that while radicalness may positively associate with survival, investing heavily in 'radical' goods for sale might be overall risky and result in diminished likelihood of survival. Effects plots of the Cox PH model 4.3 is also available in the appendix, Figure 8.25.

Table 5.16: Model 4.3 - Cox survival analysis based on innovative performance (conditional on survival up to survey date i.e. 2010)

	I	II	III	IV	V	VI
	Cox PH	Cox PH	Cox PH	Cox PH	Cox PH	Bin.Logit
logit(InnoGoods)	0.0345		0.0380		0.1176^{*}	0.0827^{*}
	(0.0363)		(0.0369)		(0.0487)	(0.0400)
logit(InnoServ)		−0.0122	−0.0194		0.0617	0.0557
		(0.0369)	(0.0379)		(0.0498)	(0.0402)
RadInnOS				−0.0884	-0.3114^{*}	$-0.1754^{\dagger}$
				(0.0719)	(0.1275)	(0.1007)
logit(R&DInt)	−0.0597	−0.0435	−0.0561	−0.0311	−0.0533	−0.0318
	(0.0491)	(0.0478)	(0.0496)	(0.0482)	(0.0495)	(0.0405)
Spinoff	0.2681	0.2622	0.2658	0.2626	$0.2810^{\dagger}$	0.3422^{*}
	(0.1662)	(0.1663)	(0.1663)	(0.1662)	(0.1665)	(0.1427)
log(Emp + 1)	-0.1511^{*}	-0.1451^{*}	-0.1498^{*}	-0.1360^{*}	-0.1339^{*}	$-0.0946^{\dagger}$
	(0.0636)	(0.0633)	(0.0636)	(0.0636)	(0.0639)	(0.0531)
(Intercept)						-1.5667^{***}
						(0.4595)
time1						$-0.0421^{\dagger}$
						(0.0248)
Sectoral Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Regional Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	2379	2379	2379	2379	2379	2565
Wald p-value	0.0183	0.0231	0.0270	0.0142	0.0067	
R^2	0.0078	0.0088	0.0093	0.0094	0.0119	
AIC						2383.7743
BIC						2711.3583
$\log L$						−1135.8872

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 5.17: Proportional hazards test for 4.3

	rho	chisq	p
logit(InnoGoods)	0.0455	0.5476	0.4593
logit(InnoServ)	-0.0465	0.5352	0.4644
RadInnOS	-0.0655	1.0152	0.3137
logit(R&DInt)	0.0014	0.0006	0.9807
Spinoff	0.0151	0.0617	0.8038
SectorCLASS[T.KIBS]	-0.0058	0.0088	0.9251
SectorCLASS[T.LTMS]	-0.0172	0.0776	0.7806
SectorCLASS[T.OBS]	0.0008	0.0002	0.9899
Sector[T.Mid EU]	-0.0051	0.0068	0.9341
Sector[T.North EU]	-0.0785	1.5967	0.2064
Sector[T.South EU]	-0.0146	0.0566	0.8120
log(Emp + 1)	0.0346	0.3959	0.5292
GLOBAL		11.6951	0.4705

5.4.2.3 Firm levels vs. firm growth and how they relate to innovative performance

As in Model 3, I have mainly looked at the levels, or intercepts, of firm size and revenue each year from 2010 to 2014. In addition to the slope regressions that accompany Model 3, I have run OLS pre-imputation regressions on the difference in natural logs between the 2014 and 2010 values of both above-named response variables. These results are available in the appendix, Tables 8.16 and 8.17. Unlike Model 3, which yielded little to no results in a growth model, Model 4's pre-imputation slope regressions do yield some results that are somewhat consistent with the intercept regressions. When growth in employees is regressed on all 3 dependent variables representing innovative performance, there is a major diminishing of effects of the variables in terms of statistical significance, in fact none are significant in a fully specified model. However, when each is isolated, and especially when `RadInnOS` is not included with the others, we find some positively significant coefficients for all 3 variables: `InnoGoods` at $p<.10$; `InnoServ` at $p<.05$, and `RadInnOS`at $p<.01$. Again though it is worth noting that this sample size due to the large amount of missing values in the response variable has gone down to between 580 and 630 observations. For growth in operating revenue during the period, we find positive and statistically significant coefficients for `RadInnOS` ($p<.10$ alone, $p<.05$ in full model) as well as `InnoServ` ($p<.10$ alone, $p<.05$ with `InnoGoods`), but a negative coefficient ($p<.05$) for `InnoGoods` in the full specification. Interpreting these findings is tricky due to the changes throughout the specifications, but on average, there seems to be some support for the notion that in this sub-sample, firms that produce more novel innovations tend to grow more than those that don't, while those that invest in innovative services may grow more than those investing in goods, which actually may not grow as much in terms of sales.

5.4.2.4 Interpretation of results: Model 4

Concerning the relationship between innovative and business performance for KIE firms, we had similar hypotheses: That higher innovative performance in terms of goods (WH-4.1), sales (WH-4.2), and radicalness of innovations (WH-4.3) would positively associate with business performance for KIE firms. Of the 3, the only one that can be given full confirmation is WH-4.3, since the variable `RadInnOS` positively associates with firm size, sales, and likelihood of survival. WH-4.1 is close to full confirmation since it seems to positively impact size, sales, and survival as well, though not as convincingly when combined with other explanatory variables. WH-4.2 does not appear significantly related to the likelihood of survival, but tends to show a positive association with size and sales over time.

Chapter 6

Summary of results and analysis

This chapter summarizes and contextualizes the empirical findings from the previous two chapters. Before delving into this though, I will readdress the concepts, constructs, and operationalizations that were used. It is difficult to pinpoint an adequate reference of origin for these types of classifications and taxonomies of measurement objects, as the dialog has been ongoing almost since the dawn of statistics. However, I have been alerted to key publications such as Blalock (1982); Cronbach and Meehl (1955); and Kerlinger (1986) for those readers wishing for a comprehensive account of what meaning is and has been attributed to these terms. See Table 6.1 below for a visualization of the categories used.

Throughout this text, I have come back again and again to the idea that knowledge intensity of the firm is intricately related its performance. For the purpose of carrying out the research objectives, I have derived 2 second order concepts from both; Internal- and External Knowledge Intensity, and Business- and Innovative Performance. Constructs used to approximate these concepts are given in the table below, along with their consequent variable operationalizations from the AEGIS survey and from Orbis. I have broken internal knowledge intensity into education, experience, organizational origins, and expressed formation factors. I have also broken down education into founder and employee-specific education, and broken down experience into entrepreneurial, academic, industrial, functional, and pre-determined experience (see table below for details). Organizational origins are expressed as the event of the company being a spinoff or not, but admittedly university experience could potentially have been included here as many view this as an often appropriate proxy for being an academic spinoff (Perkmann et al., 2013), as well as the founder having previous industry experience being synonymous with the term spinoff in the sense used by Klepper and Sleeper (2005); a firm founded by an previous employee of a firm within the same industry. Finally, expressed factors important for firm formation is a blend of both education and experience, as well as some hard to define/measure entrepreneurship concepts like opportunity recognition or opportunity creation. The benefit with these latter variables is that they are derived principle components that are orthogonal with one another, so they may be effectively used in regression analysis without multicollinearity occurring.[1]. The *intensity* component of knowledge intensity is here best

[1] In the hypotheses, first and second order constructs are prioritized over the

captured by the formation factors, and the percentile variables. Some ordinal and binary variables are also used due to lack of detailed information for certain concepts . External knowledge intensity is approximated on the construct level by the established notion of openness, which I have explained briefly in Chapter 3. Its operationalization comes initially from approximating previously used Breadth and Depth indicators (Laursen and Salter, 2006; Katila and Ahuja, 2002), but receives a more thorough treatment later by using principal components analysis to summarize the most important source components, or categories, of external knowledge for these firms. These components were: Specialist knowledge providers (Universities, research institutes and other non-industry sources); Intra-industry knowledge providers (Clients, customers, suppliers, and competitors); and Informal and codified academic knowledge providers (Conferences and publications). Although other factors may potentially play into what makes up a firms' external knowledge intensity, I have chosen to focus on these sources of knowledge to reduce the amount of inherent complexity in the model, as well as limitations in the dataset concerning external-to-the-firm variables. I am aware that I do not directly address the many types of formal external configurations that firms may have with one another, including alliances, contracting deals and other common modes of formal collaboration, and these could certainly be considered part of a firm's external knowledge intensity. However, I feel that using a measure for how much a firm relies on external knowledge for new business opportunities is a good operationalization of the concept for the purpose of this study.

In terms of Business Performance, my constructs have been firm size and firm revenue conditional on the survival of the firm, as well as firm survival itself. This I measure mainly through yearly levels (intercepts) as opposed to growth rates (slopes), though growth rate regressions were tested with largely marginal results. Innovative performance (which is often viewed as a form of operational performance), is, on the construct level, represented by the sales ratio of innovative goods and services, as well as an approximation of overall innovativeness in products and services. These innovative performance constructs seem to have an overlapping correlation problem (especially around the bottom end of the distribution), so often the first two needed to be separated from the latter in order to obtain meaningful results.

operationalizations in terms of theoretical focus, since knowledge intensity in internal (and external) terms is of chief interest here.

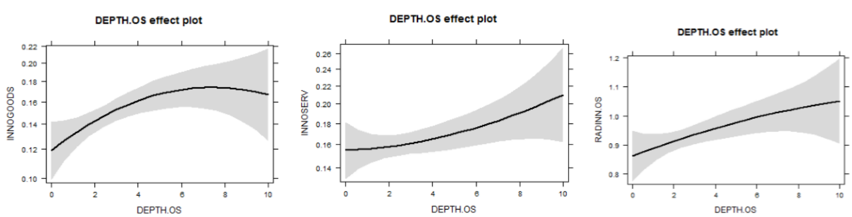
Table 6.1: Measurement of concepts throughout the study

Broad Concept	Narrow Sub-concept(s)	Construct	Operationalization on the construct (Variables) Category	Variable
Knowledge Intensity	Internal Knowledge Intensity	Education levels within the venture	Founder(s)	Max level of educational attainment within founding team
			Employees	1. Percentage of employees with a tertiary degree 2. Percentage of employees with a post-graduate degree
		Experience of the founders	Entrepreneurial Experience	Presence of prior entrepreneurial experiences in the founding team
			Industry Experience	Years of industry experience of the founding team
			University/Academic experience	Presence of prior academic work experience in the founding team
			Functional Experience	Knowledge scope of the founding team Knowledge disparity of the founding team
			Other	Age of the oldest founder
		Organizational origins	Stemming from a previous organization i.e. spinoff status	
		Expressed formation factors	Based on specific perceived external opportunities Experiential and network based factors Factors of technical, engineering or design knowledge	
	External Knowledge Intensity	Openness to external knowledge sources (in terms of reliance)	Breadth Depth Specific categories of sources/components of knowledge reliance EXPC1-3)	
Firm Performance	Business/Economic Performance	1. Firm size 2. Firm sales 3. Survival	1. No. Employees over time 2. Firm operating revenue over time 3. Firm survival up to 2015	
	Innovative/Operational Performance	Innovativeness in sales	Proportion of innovative goods and services/sales	
		Innovativeness in products/services	Highest level of radicalness in firm innovation	

I would now like to discuss the empirical findings once more, but this time in succession. A previous discussion of each models' findings can be found in the relevant empirical chapters, but the wealth of material analyzed warrants a holistic review.

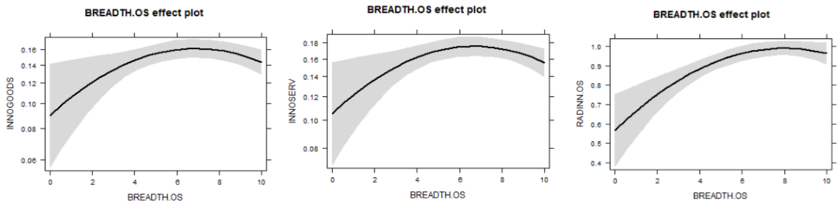
6.1 Model 1

For this subset of entrepreneurial firms, we hypothesized that external knowledge intensity by way of breadth and depth would exhibit positive linear, and negative quadratic, relationships with all 3 measures of innovative performance. Also, reliance on each distinct type of knowledge source category was hypothesized to have a positive assocation with innovative performance. However, we found differences within the relationships between each operationalization.



`Depth` is linearly related to both goods and service innovation-to-sales ratios as well as the level of radicalness of the firm's innovations overall, suggesting that, all else equal, an increase in depth, that is, adding an additional highly valued external source of knowledge, will overall have a positive effect on the innovative performance of the firm. It is also clear that after 6 or so sources, the benefits are less pronounced, probably due in part to the fact that fewer firms draw on more than this. It is noteworthy that the variable controlling for firm size was significant for innovative goods ($p<.01$) as well as radicalness ($p<.001$), suggesting that within the sample, which is 70% micro-firms, size on even this small a scale can make a difference in the association between external knowledge intensity and innovative performance.

`Breadth`, on the other hand, while only being significant in linear terms for 1 of the 3 indicators of innovative performance, had a quadratic (orthogonal) polynomial term that was statistically significant with a negative coefficient for all 3 responses (ranging from $p<.05$ to $p<.01$). This leads to the assertion that just having *some* breadth of search for



entrepreneurial firms will not necessarily result in innovative gains in relation to sales. Indeed, the effects plots of `Breadth` in 1.1; 1.2; and 1.3 (Figures 4.11, 4.12, and 4.13, reproduced above) show a more nuanced story: The effect of breadth of search on goods and service innovation does begin to produce noticeable gains at around 4 or more sources, before marginally declining at just about 7 sources. So while the resource constraints arguments that have permeated the literature (see Malerba et al., 2015) these firms do seem to experience a negative quadratic effect on their sales of innovative products and services by higher levels of breadth of knowledge sources, but the same effect does not here result from higher levels of depth.

In the case of radicalness, the third variable representing innovative performance, `Breadth` produces statistically significant and positive effects with additional sources of knowledge, and only at around 8 sources does its effect diminish. This leads to the following conclusion:

While depth of search is, all else held at its mean, beneficial for the innovative performance of these entrepreneurial firms, breadth is not necessarily beneficial at lower levels, but seems to be at higher levels, before diminishing marginal returns set in at 'too many' sources, ranging from 6-8 for all indicators. Radicalness of innovations is, all else at its mean, positively impacted by breadth as well as depth, but too much breadth may result in a lack of focus, dragging down the ability to successfully implement and introduce more radical products and services to the market.

Well these results are elucidating, the measurement was deemed as potentially problematic. Due to the way that they are constructed, there are implicit limitations present in the covariance structure of `Breadth` and `Depth` . For example, the level of `Depth` cannot exceed that of `Breadth` for a single source, so the one is in a sense a subset of the other. In order to retain maximum variance explained, and to try to further elucidate how to effectively measure external knowledge source reliance for innovation, I also broke down the summated rating scale of external knowledge sources into different categories of reliance based on principal components analysis: It was found that successful goods innovation was positively associated with all 3 categories of external knowledge sources, while successful service innovation was only linearly related to those sources not stemming from intra-industry knowledge providers. However, further analysis showed that service innovation's relationship with the reliance on these latter sources was in fact inversely quadratic at a reasonably high level of statistical significance (p<.01). For the degree of radicalness of the firms' innovations, the results are similar to that of service innovations, with pronounced linear relationships with both types of non-industry sources, and here a somewhat weak inverse quadratic relationship with intra-industry sources.

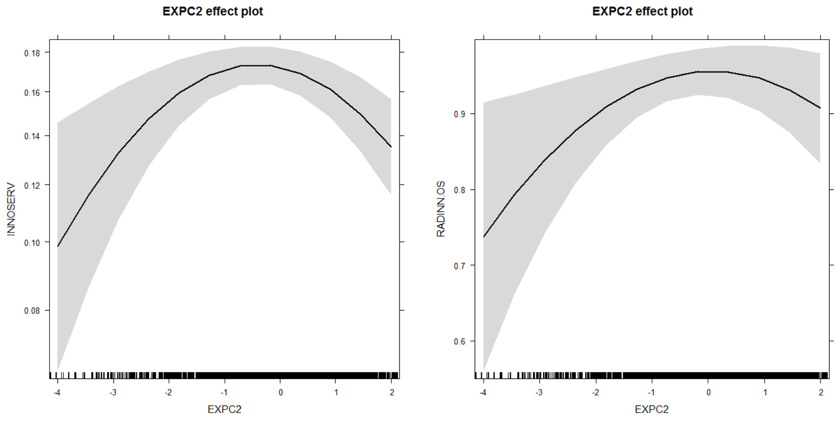
Table 6.2: Main empirical results from Model 1

Dependent variables representing innovative performance	**Main independent variables Model 1**			
	Depth	**Breadth**	**Reliance on channels for external search EXPC1,2, and 3**	**Controls**
Goods Innovation	Significantly positive but non-curvilinear Depth is beneficial at all levels, but becomes ambiguous after 6 sources	Non-significant linear relationship, but significantly and negatively quadratic: Having just some breadth is not necessarily beneficial, but high-to-mid breadth is beneficial (4 sources) and too much breadth (6-7) is not beneficial.	All sources positive and significant	R&D Intensity is curvilinearly related. International sales positive and linear Firm size positive and (log)linear Sectoral, but not regional, differences
Service Innovation	Significantly positive but non-curvilinear	Non-significant linear relationship, but significantly and negatively quadratic	Only non-industry EXPC1 (p<0.01) and academic EXPC3 (p<0.10) sources linearly significant. Business sources EXPC2 significantly and negatively curvilinear.	R&D Intensity is curvilinearly related Sectoral but not regional differences
Degree of novelty in both good and service innovation	Significantly positive but non-curvilinear	Significantly positive linear and significantly negative quadratic	Only non-industry (p<0.01) and academic (p<0.01) sources significant. EXPC2 weakly negatively curvilinear (p<0.10)	R&D Intensity is curvilinearly related. Intl Sales positive linear. Firm size positive and (log) linear. Sectoral differences, and South EU significantly higher than rest

Color coding: Green = confirmed according to hypothesis; blue = non-hypothesized result.

We can see in Figure 6.1 the effects of `EXPC2` on both service innovation and radical innovation. While no effect is discernible up to moderate levels of reliance on these sources of knowledge, i.e. there is no pattern to the effects,

Figure 6.1: Reliance on intra-industry knowledge providers' effect on service innovation and radicalness from Model 1



we can see that as levels of reliance begin to become high to extreme, a negative relationship sets in. Note though that the graph to the left displays a much more statistically significant relationship (p<.001 vs. p<.1)

What this tells us is that while successful goods innovation by these entrepreneurial firms is positively associated with higher reliance on all 3 types of knowledge sources, successful service innovation and radicalness have less straightforward relationships. While they seem to benefit from high reliance on the two non-industry source categories `EXPC1` and `EXPC3`, they do not benefit directly (as far as the data can show) from increased reliance on intra-industry sources, `EXPC2`. Nonetheless, even if there is no conclusive indication that a higher intra-industry knowledge reliance benefits innovative performance in these two indicators, having too high reliance results in lower performance for services and radicalness only.

These results suggest that, controlling for sectoral and regional differences, entrepreneurial firms benefit differently from different types of external knowledge intensity depending on what types of innovation, goods vs. services, the company produces. This is the case at least when it comes to intra-industry sources. Radicalness of innovations lies closely related with goods innovations, but it may be suggested that a more nuanced view of innovative performance may be needed for the study of entrepreneurial firms.

We can see that firms that focus more on knowledge stemming from specialist knowledge providers, component `EXPC1`, (Tether and Tajar, 2008), as well as codified and informal academic sources, `EXPC3`, seem to have a higher association with innovative performance across all three response variables. Regardless, there seems to be a high association between heightened innovative performance, and a firm going outside its network of intra-industry partners, suppliers, competitors, and customers in order to search for new forms of business opportunities, and new types of knowledge relevant to the firm's strategic direction.

What can be taken away from these results? External knowledge intensity does not relate to innovative performance of entrepreneurial firms exactly as we hypothesized. Though `Breadth` did not relate linearly as we foresaw and thus complicating H1.1a, H1.1b seemed plausible given the inverse quadratic relationships in all 3 sub-models. H1.2b was rejected, though a linear relationship is present across all sub-models for `Depth` , thus confirming H1.2a. H1.3 is rejected, as `Depth` coefficients are not always smaller than `Breadth`.

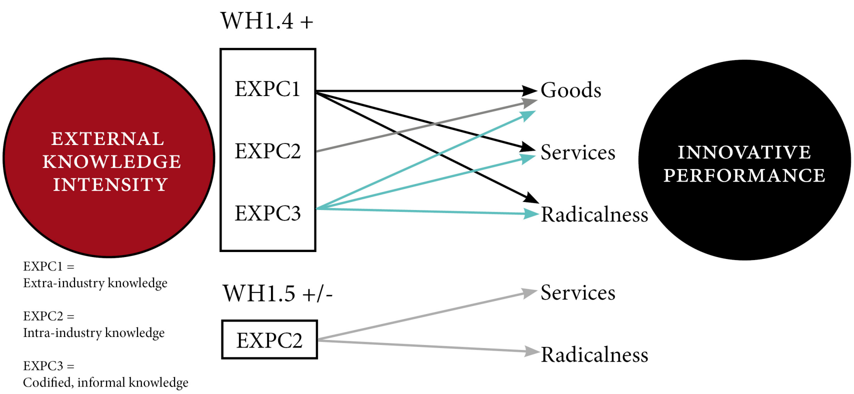
Figure 6.2 shows the portions of Model 1 that received only partial confirmation in their treatment.[2] WH1.4 was as a whole rejected, as reliance on external knowledge stemming from specialist knowledge providers and from codified, informal knowledge sources is positively associated with all 3 types of innovative performance, but intra-industry sources of knowledge were not positively associated with services. WH1.5 is also rejected, but gives some interesting food for thought, with inverse quadratic relationships between the independent and dependent variables in 2 of 3 cases. This serves to illustrate that while the model these hypotheses are not supported via all operationalizations, there are some interesting relationships at play here, and measurement error might play a part in why confirmation is not achieved.

Most surprising are the at-first-glance counter-intuitive results of the effects on innovative services. The analysis does not find a significant relationship between clients, customers, competitors and suppliers as knowledge sources, with that of innovative services as a proportion of sales. This is surprising, since the literature detailed above in the theoretical section points towards the importance of user communities, business networks, etc., in the innovation process of services (von Hippel, 1988; Miles, 2012). However, it is clear that, at least at the mean, more depth in search associates with better innovative service performance. Given the theoretical reasoning, it seems that constraints of size and newness have a real effect on what sources successful service innovations

[2]Note: The varying color of the arrows in the figure only serves to visually differentiate the sources, and is not meant to convey any specific meaning other than this. Also, these figures throughout the chapter show the partially confirmed models. I will summarize all fully confirmed hypotheses for all models at the end of the chapter.

may be drawn from and how deep collaborations need be (or can be) in the initial phases of firm development. Also, more knowledge intensive business service firms may draw more extensively on specialized sources of knowledge than previously accounted for. The prevalence of these type of firms in the sample may affect the reliance on traditional industry sources, and since these type of firms in KIBS sectors are likely more apt to apply outside knowledge in their products and services, the results are perhaps not so strange. It is also worth noting that at present I am working on a collaborative paper detailing further exploration of this result, and it would appear that to some degree a service firm's appropriation regime concerning intellectual property protection; that is, how much the firm relies on informal vs. formal IPR methods, does have an effect on how they approach their external knowledge sources. The details of this I leave for later work.

Figure 6.2: Partially confirmed hypotheses Model 1



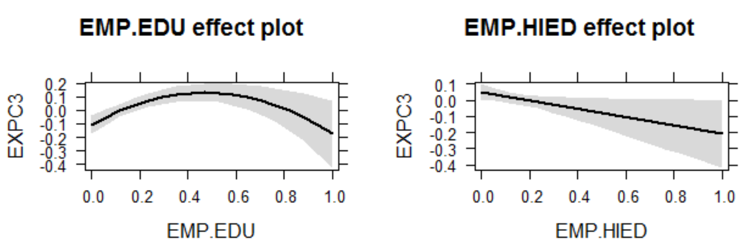
6.2 Model 2

I will turn now to interpretation and discussion of the results of the 2nd empirical model, in which I investigated the effect of internal knowledge intensity upon external knowledge intensity. The tables 6.3 and 6.4, and figure 6.5 below show the main results, direction of results, and status of the hypotheses for Model 2.

Concerning WH2.1, again the results are mixed and more nuanced than expected:

First, regarding *education levels of the firm*: Education of employees seems to be a slightly more effective variable in its association with external knowledge intensity. Firms with higher proportions of at-least tertiary educated employees on average relied more highly on specialist knowledge providers and on informal codified and academic knowledge, but relied less on intra-industry knowledge sources. Further data diagnostics revealed a curvilinear relationship between proportion of tertiary educated employees and that of reliance on how much reliance a firm places on informal, codified and academic knowledge (see Figure 6.3 below). Higher proportions of Masters and PhD education among employees was associated with lower reliance on inter-industry sources of knowledge, as well as with informal, codified and academic knowledge. The education level of the founding team was positively associated with informal and codified sources. This means that firms with higher educated employees and founders derive higher value from non-industry sources (`EXPC1` and `EXPC3`) and less value from inter-industry sources (`EXPC2`). Firms whose formation was strongly due to technical and engineering experience (`FF3`) were negatively associated with reliance on intra-industry sources of knowledge, though the relation proved following diagnostics to be convex curvilinear, meaning that the relationship is at lower levels of importance of technical and engineering knowledge, negative, with higher levels of importance of these knowledge types this becomes less detrimental (see Figure 6.3 below, far right). Regardless, the hypothesis proved over-simplified and cannot be confirmed by the regressions.

Figure 6.3: Revealed relationship between tertiary and higher education in employees and reliance on codified academic knowledge sources in Model 2



WH2.2a stated that higher university or research experience would associate with higher reliance on specialist knowledge providers and codified academic sources, and 2.2b stated that industry experience would negatively associate with these sources. Entrepreneurial experience in the team, as well as the importance of experiential and network knowledge as

a formation factor, had negative associations with specialist knowledge providers, so, founding teams that have founded prior firms or were reliant on industry experience and networks may be less likely to rely on these sources than those teams which are not. However, it was also shown that higher levels of founder entrepreneurial experience was positively associated with higher reliance on the category of codified and academic knowledge sources. So it seems that WH2.2a and WH2.2b may both be partially confirmed for specialist knowledge providers, but not for informal, codified and academic sources which did not associate entirely as hypothesized.

WH2.3a and 2.3b were in regard to the types of experience within the firm as such, founding teams containing university experience were more likely to have a lower reliance on inter-industry knowledge sources, while founding teams industrial and entrepreneurial experience would be more likely positively associated with inter-industry knowledge sources. It was found that university experience on the founding team lessened the association with intra-industry knowledge source importance, while experience and network factors driving firm formation positively associated with intra-industry knowledge source importance. The variable representing entrepreneurial experience on the part of the founding team was not significant. So, previous industry experience, but not necessarily entrepreneurial experience, drives up the importance of intra-industry sources, while university experience brings this down. WH2.3a is then confirmed, while 2.3b is partially confirmed with 1 variable behaving as hypothesized while others do not appear significant.

Concerning WH2.4a and 2.4b: Knowledge scope and disparity proved to be ineffective predictors of degree of external knowledge intensity in the targeted firms. This tells us that a founding team's functional heterogeneity, whether wide or narrow, may not adequately predict how much firms draw and rely on external knowledge sources.

WH2.5 stated that spinoff firms would have higher external knowledge intensity than non-spinoff firms. This was unconfirmed in the two models dealing with extra-industry knowledge (`EXPC1` and `EXPC3`), and disconfirmed in the one dealing with inter-industry knowledge (`EXPC2`). Indeed, all else at mean values, firms that are spinoffs of previous organizations are less reliant on business and supplier-based relationships summarized in `EXPC2`.

WH2.6 dealt with different types of opportunities. Opportunity based factors based on technical change, market-based and institutional change, and new regulations were positively associated with reliance on both specialist knowledge providers (`EXPC1`) and on inter-industry sources (`EXPC2`), but was not significantly associated with the third response variable in this model.

Like Model 1, I show the partially confirmed results of the hypotheses from Model 2 in Figure 6.5. It can be seen that despite non-confirmation of many of the hypotheses, the results were much stronger for some variables than for others. Many variables had strong results in terms of associations with EXPC1 and EXPC3, while EXPC2 associations were on the whole quite weak. This conveys that while the effects are not wholly confirmed, there is a more nuanced picture emerging of how internal knowledge intensity may relate to external knowledge intensity, and that more a more rigorous and detailed operationalization and hypotheses concerning direction of effects could aid this in future studies than what was achievable with the AEGIS survey in the context of this dissertation.

Figure 6.4: Formation factor non-linear effects

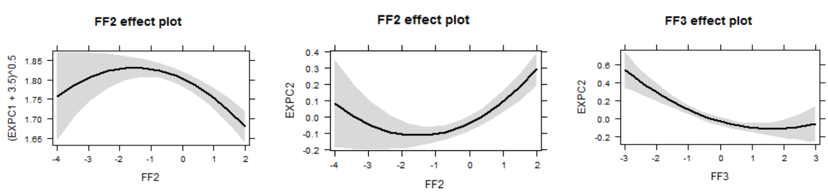


Figure 6.5: Partially confirmed hypotheses Model 2

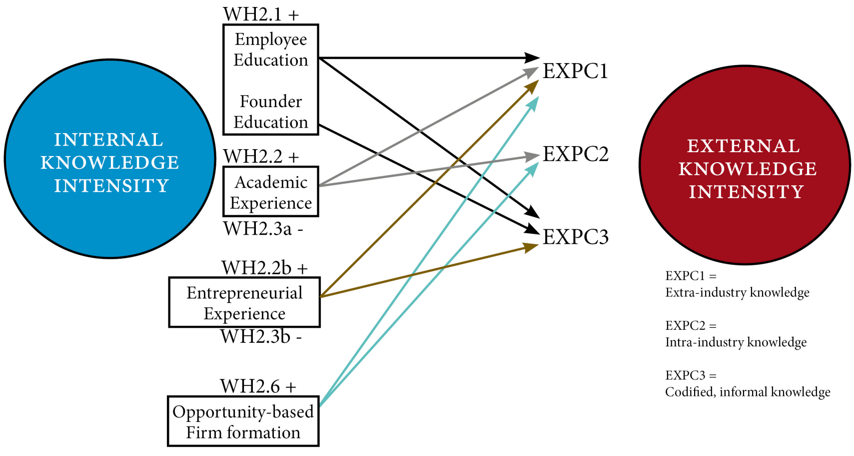


Table 6.3: Main empirical results of Model 2, part 1

Main Independent Variables Model 2	**Dependent Variables representing External Knowledge Intensity**		
	EXPC1 Specialist Knowledge Providers as sources	*EXPC2 Business Knowledge Providers as sources*	*EXPC3 Informal, codified and academic knowledge providers as sources*
EmpEdu	Significantly positive and linear (p<.01), Higher educated workforce associates with increased reliance on SKPs	Significantly negative and linear (p<.001), Higher educated workforce associates with decreased reliance on BKPs	Significantly positive and linear (p<.05), and significantly negative quadratic term (p<.001), Higher educated workforce associates with increased reliance on ICAKPs, but as this proportion increases, the marginal reliance decreases
EmpHiEdu	Non-significant	Significant and negative linear term (p<.01) Higher proportions of MSc and PhDs associates with less reliance on BKPs.	Significant and negative linear term (p<.05) Higher proportions of MSc and PhDs associates with less reliance on ICAKPs.
FoundEdu	Non-significant	NA	Significant and positive linear term (p<.01) Higher average education less reliance on ICAKPs.
FF3 Specialized knowledge formation factors	Non-significant	Significantly negative linear (p<.001), and significantly positive linear (p<.001), Higher importance of engineering and design knowledge in firm formation is negatively associated with reliance on BKPs, but as importance increases, reliance does as well	Non-significant

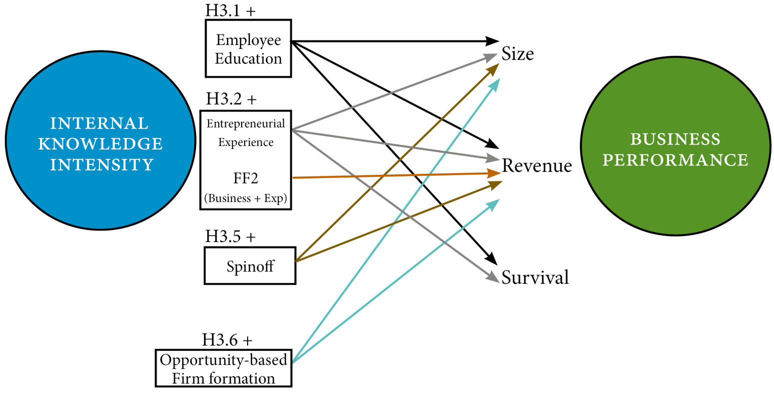
Table 6.4: Main empirical results of Model 2, part 2

Main Independent Variables Model 2	**Dependent Variables representing External Knowledge Intensity**		
	EXPC1 Specialist Knowledge Providers as sources	*EXPC2 Business Knowledge Providers as sources*	*EXPC3 Informal, codified and academic knowledge providers*
Founder University Experience	Significantly positive ($p<.001$), The presence of university exp. in the team increases reliance on SKPs	Significantly negative ($p<.05$), The presence of university exp. in the team decreases reliance on BKPs	Non-significant
Founder Entrepreneurial Experience	Marginally significant and negative ($p<.10$), but not significant in revised model Potential relationship between higher founder entrepreneurial exp. and reliance on SKPs	Non-significant	Significantly positive ($p<.05$) Higher levels of founder entrepreneurial exp. associates with higher reliance on ICAKPs.
Founder average no. years of industry experience	NA	Significantly negative ($p<.001$), Higher levels of industry experience associates with less reliance on BKPs.	Non-significant
Max age of founders	Non-significant	Non-significant	Significantly negative ($p<.10$), Potentially higher average age of founding teams associates with less reliance on ICAKPs.
FF2 Experience and network based formation factors	Significantly negative and linear ($p<.001$), and significantly negative quadratic terms ($p<.01$) Experience and networks being highly important for firm formation is negatively associated with SKPs, and as importance increases, reliance on SKPs diminishes further	Significantly positive linear ($p<.001$), and positive quadratic ($p<.001$), Experience and networks being highly important for firm formation is positively associated with higher reliance on BKPS, and as importance increases, reliance does as well	Non-significant
Spinoff Status	Non-significant	Significantly negative ($p<.05$), The firm being a spinoff is associated with less reliance on BKPs.	Non-significant
FF1 Opportunity based formation factors	Significantly positive and linear($p<.001$). Opportunity based factors importance for firm formation is positively associated with higher reliance on SKPs.	Significantly positive ($p<.001$) and linear term. Opportunity based factors being highly important for firm formation is positively associated with higher reliance on BKPs	Non-significant

6.3 Model 3

I will now address Model 3, which covered the effects of internal knowledge intensity on business performance. Tables 6.5 and 6.6, along with Figure 6.6 below convey the results of the model.

Figure 6.6: Partially confirmed hypotheses Model 3



Levels of education in the firm produced varied results on business performance, the strongest stemming from levels of employee education. Higher proportions of employees with tertiary educations were positively associated with both business performance in terms of firm size and firm revenue. Higher proportions of MSc and PhD educations among employees were associated with lower performance in terms of firm size, while higher levels of founder education (averaged across the team) was quadratically associated with firm size, suggesting that founder education may not matter until the level is moderate to high, which does have a positive effect on firm size over time. `FF3`, depicting the importance of technical and design knowledge in founding the firm, was negatively associated with firm size, and marginally negatively associated with operating revenue (p<.10 in full specification).

Regarding types of career experience among the founding team, teams with at least 1 founder with previous entrepreneurial experience had increased performance in terms of the size of the firm as well as in its operating revenue over 5 years. Additionally this variable was (p<.05) positively associated with likelihood of firm survival. University career experience was not significant for any of the dependent variables in this model. Years of industry experience and max age of founders yielded no results. `FF2`, or networks and experience being important for forming the firm, was positively associated with operating revenue only. Concerning functional experience: It was found that knowledge scope and knowledge disparity behaved largely as expected in these models, with knowledge scope producing a statistically significant linear, and inverse quadratic, effect in the first response variable (number of employees), a mild similar effect in the second (operating revenue), and only an inverse quadratic effect in the third (the survival measure denoting firm exit). So, the event of firm exit was not associated with a lack of knowledge scope, but as the index of knowledge scope became very high, the likelihood of the event happening increased. So while too high knowledge scope seems to be detrimental for all measures of business performance in these entrepreneurial firms, its beneficial association may only be observed in terms of firm size and revenue, and only ($p<.10$) marginally in the likelihood of survival. Knowledge disparity, on the other hand, produced the hypothesized effects in every model of Model 2, confirming that high levels of knowledge disparity in the sample are detrimentally associated to the business performance of entrepreneurial firms.

The founder coming from a previous organization had a positive association with both time-variant indicators of business performance, but was not significantly related to the likelihood of survival for this subset of firms. This is in line with Klepper's (2002) study which posited that spinoff performance and survival would in some cases surpass that of *de novo* entrants, although the measure that was more closely aligned with Klepper and Sleeper's (2005) definition of a spinoff, amount of same industry experience on the founding team, showed no results. Concerning the surveyed factors relevant for firm formation by respondents, those involving the exploiting of novel opportunities (`FF1`) showed statistically significant associations with the first two indicators of business performance, but not with survival. Given these results, much confirmation of the NTBF literature by Criaco at al. (2013), Colombo and Grilli (2005), and Almus and Nerlinger (1999) is to be found, and the relationships hold, in many respects, when one shifts the discussion to focus on internal knowledge intensity as I have operationalized it.

I will now look specifically at the hypotheses: H-3.1, which concerned human capital in the form of education, seems to be an oversimplified hypothesis in the model, and failed to be confirmed overall. *It seems that employee education may affect the firm in a different way than founder education, even in very small firms, and in this case to be a better predictor of business performance over time than founder education.* The proportion of post-graduate educated employees does not seem to affect every indicator for business performance. Much of this could be due to routine and competence building within the firm that is easier done by tertiary educated employees than by MSc or PhDs, who may already have self-institutionalized some academic tendencies in how they work and interact which may not be in full alignment with the firm's business model and social structures. *We also see that while formal education on the part of the founder did not seem to play a role (except in potentially making firms less profitable over time), founders having entrepreneurial experience is quite a strong predictor comparably.* Technical and design knowledge (`FF3`) as formation factors also negatively influence economic performance in terms of size and, marginally, sales.

H-3.2 looked at whether entrepreneurial, industrial and academic experience are positively related to business performance. The presence of entrepreneurial experience on the founding team was positively associated in all models. The only other variables representing these constructs that appeared significant, `FF2` (experiential and network factors' importance for firm formation) was only positively related to revenue, and ($p<.10$) marginally to firm size in one specification. H-3.3 and H-3.4 regarding knowledge scope and disparity comes closest to full confirmation. Linear and inverse curvilinear effects are present in all models for knowledge scope, and disparity negatively associates with all response variables in the model. All coefficients are significant at least at the 5% level with the exception of knowledge scope and its quadratic as affecting survival outcomes, which fell at the 10% level. *This suggests that knowledge scope does play a role in business performance in the medium-to-long term, but that its effect on whether or not a firm survives is marginally conclusive at least at lower levels of knowledge scope among the founding team. Knowledge disparity in the model behaved largely as hypothesized, negatively affecting overall performance.*

Regarding H-3.5, we see that the firm being a spinoff increases relative performance, but not necessarily affects survival for these firms, as it has a mild positive ($p<.10$) association with firm exit. Similarly, H-3.6 can be partly confirmed in that for both size and sales over time as response variables, the more important novel opportunities were in founding the firm, the better they performed, but not in terms of survival vs exit. Figure 6.6 shows the partially confirmed hypotheses of Model 3. Again, there is holistically no support for them, but a more detailed look reveals many robust relationships.

Table 6.5: Main empirical results of Model 3, part 1

Main Independent Variables Model 3	**Dependent Variables representing Business Performance**		
	Firm Growth in Employees 2010-2015	*Firm Growth in Operating Revenue 2010-2015*	*Event of firm exit (Survival indicator)*
EmpEdu	Significantly positive ($p<.001$), Higher levels of employee education associate with higher firm growth in employees	Significantly positive ($p<.05$), Higher levels of employee education associate with higher firm growth in operating revenue	Significantly negative ($p<.10$),
EmpHiEdu	Significantly negative ($p<.05$), Higher levels of MSc and PhD educated employees are negatively associated with firm size	Non-significant	Non-significant
Founder Educational Attainment	Non-significant linear but positive and significant, quadratic term, This suggests founder education does not matter until the level is ,moderate to high, which does positively affect growth.	Non-significant	Non-significant
FF3 Specialized knowledge factors	Significantly negative ($p<.001$), Higher importance of design and technical knowledge factors in forming,the firm associates with lower firm size over time	Non-significant	Non-significant
Founder Entrepreneurial Experience	Significantly positive ($p<.05$), The presence of Ent. Exp. in the founding team is positively associated with firm size over time	Significantly positive ($p<.05$), The presence of Ent. Exp. In the founding team is positively,associated with operating revenue over time	Significantly negative ($p<.10$) in revised model, Ent. Exp in the founding team is negatively associated with the event of firm exit.
FF2 Experience and network based factors	Non-significant	Significantly positive ($p<.001$), Higher importance of experience and network factors in forming the,firm associates with higher operating revenue over time.	Non-significant

Table 6.6: Main empirical results of Model 3, part 2

Main Independent Variables Model 3	**Dependent Variables representing Business Performance**		
	Firm Size in Employees 2010-2015	*Firm Sales in Operating Revenue 2010-2015*	*Event of firm exit (Survival indicator)*
Knowledge Scope	Significantly positive ($p<.001$) and curvilinear,($p<.001$). Knowledge scope is curvilinearly related to business performance in terms of firm size over time	Significantly positive ($p<.001$) and mildly curvilinear,($p<.10$). Knowledge scope is curvilinearly related to business performance in terms of operating revenue over time	significantly negative ($p<.05$) and curvilinear,($p<.05$). Knowledge scope is curvilinearly related to business performance in terms of survival over time
Knowledge Disparity	Significantly negative ($p<.001$), Knowledge disparity is negatively associated with business performance in terms of firm size over time	Significantly negative ($p<.05$), Knowledge disparity is negatively associated with business performance in terms of operating revenue over time	Significantly positive ($p<.05$)
Spinoff Status	Significantly positive ($p<.01$) The firm being a spinoff is positively associated with firm size over time	Significantly positive ($p<.05$), The firm being a spinoff is positively associated with operating revenue over time.	Non-significant
FF1 Opportunity based Factors	Significantly positive ($p<.001$), Higher importance of opportunity factors in forming the firm, associates with higher firm size over time.	Mildly significantly positive ($p<.10$), Higher importance of opportunity factors in forming the firm mildly, associates with higher operating revenue over time.	Non-significant

Education and experience appear to be generally quite good associative variables to all 3 responses: Size, revenue and survival. The lack of a statistically significant relationship with all 3 disconfirms many hypotheses in the model.

It seems then that many aspects that I have classified as *internal knowledge intensity* do positively impact *business performance* over time, but that finer spliced indicators may be warranted to truly understand the inter-relationships. Here we have shown that employee level education indicators seem to fit the problem better than only founder-based indicators. This may be because in micro-firms, the employee human capital may rival the founder human capital in terms of importance for future development of knowledge intensity, innovation and performance. At the same time, founder experience seems to play a significant role across performance indicators. More careful accounts of employee experience, which are absent here, would be useful to account for how important this is comparatively, and whether founder experience is ideal experiential indicator.

6.4 Model 4

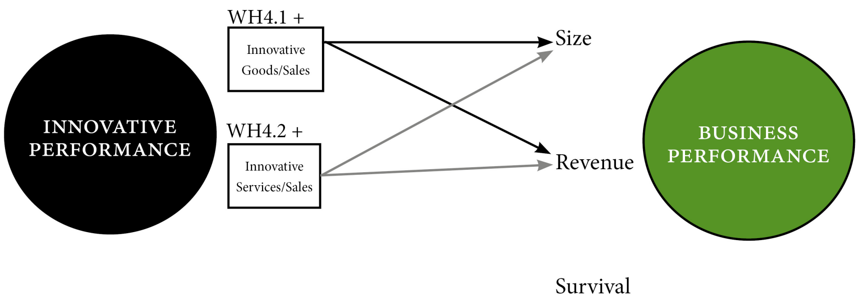
Concerning the final empirical model, Model 4. The results may be summarized quite succinctly. Table 6.7 and Figure 6.7 below convey the results of the models and of the hypotheses. Only radicalness of innovations produced by the firm seems to be an adequate predictor of business performance for this sample of firms in the fully specified models, but both other main explanatory variables showed significant associations in specifications that did not include `RadInnOS`, suggesting a collinearity problem among the explanatory variables. I found a positive association with both firm size and firm operating revenue over time (though for operating revenue the resulting relationship was curvilinear, diminishing with higher values of radicalness), and a positive association with the likelihood of the firm surviving to the present day. The other two main explanatory variables proved to be effective predictors of firm size and revenue over time (both yielding positive coefficients) as long as `RadInnOS` was not in the model. In the Cox PH model however they were ineffective, except in the fully specified model `InnoGoods` was positively associated with firm exit. Interestingly, one of the controls, R&D Intensity as a proportion of total sales, was negatively related to the firm's operating revenue over time. It is interesting here that R&D Intensity and level of radicalness of innovations do not have a similar direction of associations with the response variable.

As I pointed out in Chapter 5, I also attempted to model firm growth in size and sales regressed on innovative performance using differences in logarithmic values in the response variable from base year to target year. These were done on the dataset prior to imputation, as calculating growth rates on imputed data would likely be a spurious affair. All three explanatory variables of innovative performance yielded positive and statistically significant associations with the growth rate variables for firm size, while the innovative goods to sales ratio was negative and significantly (p<.05) associated with operating revenue growth, and services to sales as well as radicalness were positively associated (p<.10 or p<.05, see Tables 9.16 and 9.17 in the appendices) in most specifications. Meanwhile R&D Intensity seemed negatively related to growth (ranging from p<.05 to p<.10 across specifications) in operating revenue.

Table 6.7: Main empirical results of Model 4

Main Independent Variables 4	**Dependent Variables representing Business Performance**		
	Firm Size in Employees 2010-2015	*Firm Sales in Operating Revenue 2010-2015*	*Event of firm exit (Survival indicator)*
Innovative Goods/Total Sales	Sig-positive (p<.05)	Sig-positive (p<.05)	Significant and positive (p<.05)
Innovative Services/Total Sales	Sig-positive (p<.05)	Sig-positive (p<.05)	Non-significant
Radicalness of Innovations	Significantly positive (p<.001), The more radical the degree of innovations introduced the higher the size of the firm becomes over time	Significantly positive (p<.01), but with a negative and significant quadratic term (p<.001) The more radical the degree of innovations introduced the higher the operating revenue of the firm becomes over time. However at very high levels of radicalness, operating revenue experiences diminishing returns	Significantly negative (p<.05), The higher the radicalness of innovations introduced by the firm, the lower the likelihood of exit over time
Controls	Spinoff is significantly positive (p<.001) There are regional and sectoral differences	Spinoff is significantly positive (p<.05) Lagged # of employees is significantly positive (p<.001) Logit of R&D Int. is sig negative (p<.05) There are regional and sectoral differences	Also there are differences between KIBS and LTMS with that of HTMS

Figure 6.7: Partially confirmed hypotheses Model 4



6.5 The fully confirmed hypotheses

Focusing only on the fully confirmed portions of the hypotheses, three associations between the constructs can be seen. Knowledge scope showed a curvilinear relationship with business performance, while knowledge disparity was negatively and linearly related. Breadth was curvilinearly related to innovative performance, while depth was linearly related to the same construct. Moreover, innovative performance in terms of radicalness of innovation shows a positive association with business performance of the firm. In the final chapter I will reflect on these relationships, but also on the partially confirmed hypotheses that are most interesting for further research and analysis.

Chapter 7

Discussion and conclusions

> *"Innovation creates diversity and diversity, in true evolutionary fashion, makes competition feasible. Competition in turn stimulates the search for innovation based advantage and in the process... creates innovation systems from general capabilities. Thus the relation between the knowledge of growth and the growth of knowledge really is double-sided."*
>
> John Stanley Metcalfe, 2002: 14

This chapter reflects on the results of the previous chapter and links them with current theory on knowledge intensive entrepreneurship by contextualizing the relationships between knowledge, innovation and performance in entrepreneurial firms. It discusses the implications of these results, and how they might inform the future of the knowledge intensive entrepreneurship research and policy agendas.

In previous chapters, I recounted that policy has in recent years placed an amplified focus on a subset of simultaneously innovative and entrepreneurial firms, due to their assumed importance in economic growth. Stimulating small and micro firms to be more knowledge intensive and more innovative has risen to a place near the top of inter-governmental and international policy agendas. The OECD (2008), as well as the European Commission (2013) has argued that small entrepreneurial firms in key sectors are some of the most important actors driving global economic growth, and that using policy as a tool to help them overcome challenges to their size, networking potential, and competitiveness could be a strong recipe for strengthening entrepreneurship at virtually all levels of economic activity. Mostly, inter-governmental organizations' prescriptions have focused on the broadening of support programs; including better-tailored network building, greater policy awareness, and not least, encouraging more collaborative measures among these firms regarding innovation in order to become more internationally competitive (OECD, 2008; OECD, 2013). These developments are seen by Mytelka and Smith (2002) as a co-evolution of sorts between innovation theory and innovation policy, giving rise to new conceptual approaches on an international scale. EU framework programs like KEINS and AEGIS called for, and implemented, an approach that looks beyond traditionally knowledge intensive sectors, and a need to study pre-entry characteristics of entrepreneurial firms with potential for high knowledge intensity; also, they emphasized the need to study how firms use existing networks, other resources, and capabilities to sustain, grow, and to ultimately better the economy as a whole. Taking inspiration from KEINS and AEGIS, I have attempted to conduct an in-depth analysis of knowledge intensity in external and internal dimensions by making use of certain aspects of the pre-entry and external search activities, and then relate them to different types of performance using a sample of firms that come from sectors with higher potential of containing knowledge intensive entrepreneurship. This has been an exploratory attempt to further theoretical and operational understanding of KIE and knowledge intensity.

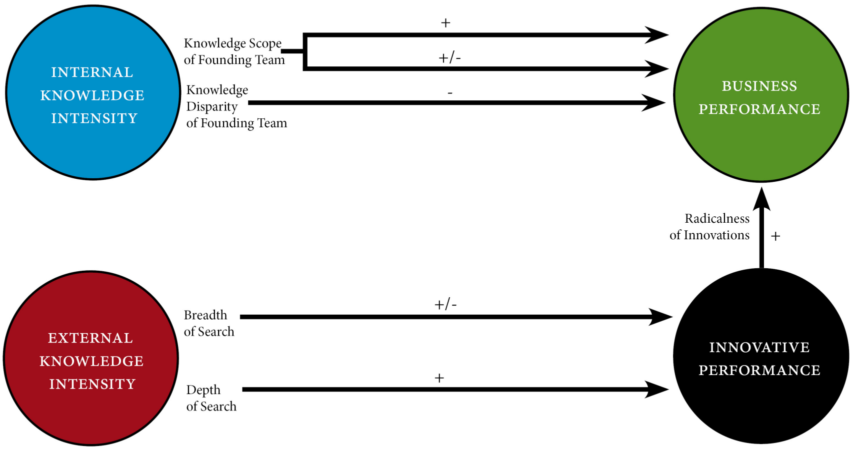
At the onset of this monograph, the research objective was made plain; *explore in what way and to what extent different forms of knowledge intensity present in entrepreneurial firms interact, and to what extent they influence overall firm performance.* I have used quantitative empirical models to answer this through a series of regression analyses. Realistically though, the statement must be broken down and its components analyzed. Also, since the objective itself is an overarching one, it is more beneficial perhaps to discuss in turn the objectives derived in Chapter 3. By discussing the extent to which they have been answered, a clearer picture of broad implications will emerge. Figure 7.1 summarizes the fully confirmed relationships, while Figure 7.2 summarizes the partially confirmed relationships (not including those with full confirmation) that I consider to be of chief interest, and will discuss in this chapter. As I will allude to in the sections to follow, many of these 'partial' relationships are quite relevant despite not fulfilling my strict criteria for full confirmation. It may be helpful for the reader to refer to these while reading the following sections, as the wealth of relationships taken up are somewhat difficult to summarize.

7.0.1 The association between external knowledge intensity and innovative performance

Our first original research objective for derived from theory was as follows: *Explore the association between external knowledge intensity and innovative performance in the entrepreneurial firm*

Based on the results, some of the hypothesized associations between external knowledge intensity; or, as characterized in Chapter 1, *the way*

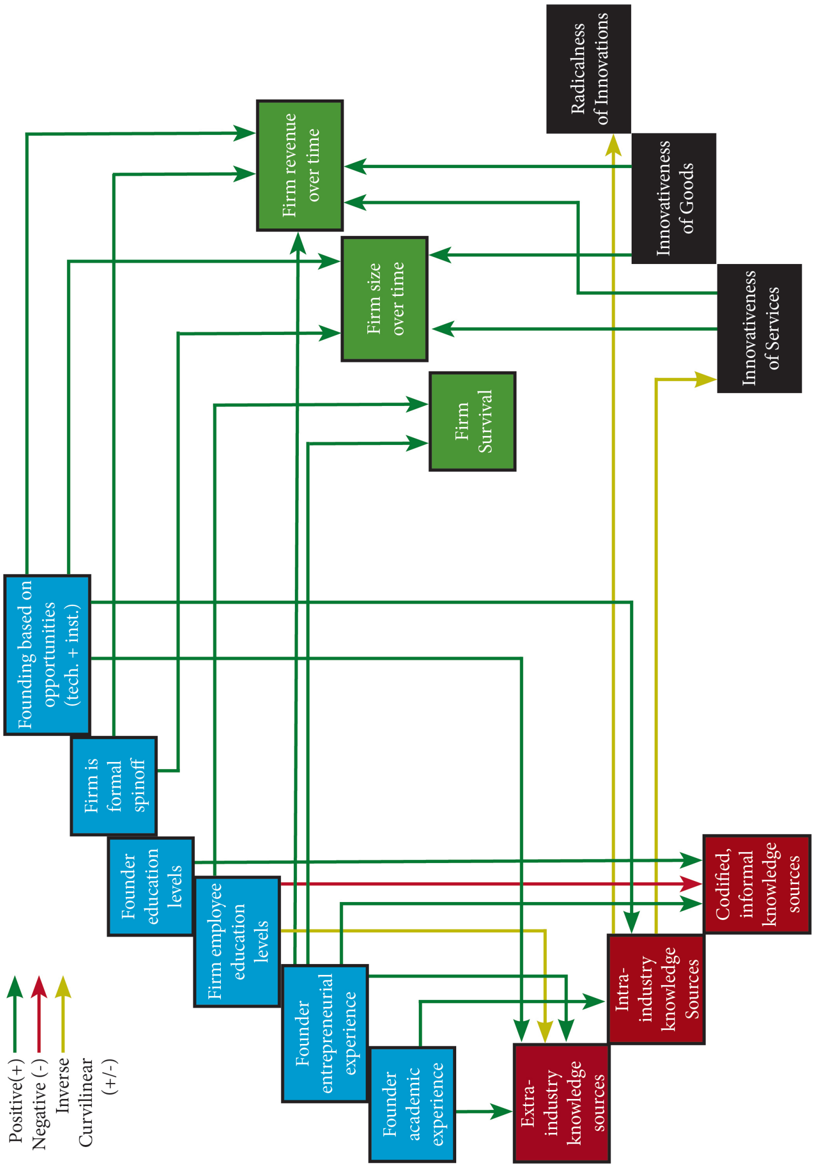
Figure 7.1: Fully confirmed associations



and extent to which a firm searches out, relies on, and valuates external knowledge post-formation, and innovative performance, which was measured in terms of sales of innovative products as well as radicalness (or novelty) of these innovations, can be confirmed. Others however were left not confirmed or only partially confirmed[1]. In this first model, I used previously established operationalizations in the form of breadth and depth of external search (Laursen and Salter, 2006), as well as in alternative models, summarizing measures for the reliance on different *types* or *categories* of external knowledge. Among the partial confirmations: Reliance on sources of knowledge stemming from non-industry actors, (including both specialist knowledge providers such as universities and research centers as well as informal, codified sources like academic and trade journals/conferences) have a discernible positive association with innovative performance, while only an over-reliance on intra-industry sources (mainly clients, customers, and suppliers) seems to (negatively) affect innovative performance in services (see Table 6.2). I found fully confirmed relationships in the following associations: The breadth of external search had a negative marginal relationship at higher values, while depth of external search yields on average a positive association with innovative performance.

[1]Moving forward, I mean 'full confirmation' in the sense that all dependent variables representing a concept had a statistically significant association with the independent variable being assessed in the hypothesis, whereas partial confirmation refers to hypotheses where only some of the dependent variables has an association

Figure 7.2: Partially confirmed associations of interest



While theory has only recently begun to address the prevalence of types of openness and search strategies of small and micro-firms (as discussed in Chapter 3), I expanded on this new area by focusing on diverse manufacturing and service sectors in a wide range of countries within the European Union. Also, as previously found by the literature, I find differences between services and manufacturing-based activities concerning innovation as an outcome of reliance on external-to-the-firm knowledge.

Breadth as well as depth of search, as measured here, did not associate entirely as hypothesized for these small entrepreneurial firms in potentially knowledge intensive sectors. *Full confirmation was provided in the analysis that depth of sources was in fact positively associated to innovative performance, without any diminishing returns detected.* This is at odds with some of the theory specifying resource constraints and periphery status in value chains as hindrances to competitiveness through increased network dependence (Forsman, 2009; 2011). Deep collaboration did not, in this study, seem to be limited in this way, as proposed by much of the extant research (Thorgren et al., 2012; Alonso and Bressan, 2014). Due to liabilities of newness and of smallness experienced by these firms, an increase of depth is more effective than an increase of breadth, since the interaction costs may be lower for a small new venture to deepen its existing relationships rather than forge new ones. It might also be so that depth does not experience diminishing returns because the size and resource constraints do not permit this level of excessive depth that has been documented in larger firms. Since these two concepts share variance per definition (see figure 4.9 and the surrounding discussion in Chapter 4), relying solely on them as indicators was problematic. This is one of the main reasons I chose to supplement the analysis with principal components. *Speaking broadly though, it would seem that higher depth of search in potentially-KIE heavy sectors seem to lead to higher innovative performance, rather than broader search. And overall, breadth and depth associated differently with innovative performance than has been common for studies of both large manufacturing and service firms, with depth providing more linear benefits to innovativeness and breadth lying in the margin, negatively affecting innovativeness when its levels become moderately high.* Indeed the results of this model are less similar to Laursen and Salter's (2006) popularized conception of breadth and depth of search and now it affects performance, and more in line with more recent work on Open Innovation that finds only positive effects for depth (Bengtsson et al., 2015).

Some aspects of the partially confirmed results from the first model warrant further study: Part of the model addressed the firm's reliance on different groupings of knowledge sources, and their association with innovative performance. There is much more to be done here, and this conceptualization is just the beginning. The finding that service innovations and intra-industry sources are quadratically and inversely related, but not linearly related, is quite striking. It could be that service oriented innovators are already so entrenched in customer/client relationships that a deepening or broadening is not tangible, or relevant. However, since knowledge intensive business services vary highly in their transient nature of delivery, some being highly based on local and cultural contexts (Tether and Hipp, 2002), the value of added depth in relationships might be less relevant beyond the networks in place during the innovation process itself, and even detrimental. In any case, these relationships uncovered here warrant further investigation and theory testing, among which the analysis of intellectual property regimes of service firms might prove particularly elucidating.

Comparing these results with the theory put forth in Chapter 3, I find that cooperative networks do seem to enhance learning and innovativeness in small and micro-firms (Chell and Banes, 2000; Mäkinen, 2002; Reinl and Kelliher, 2010). Also the drawbacks related to newness, smallness and competition (Kotey and Sheridan, 2004; Forsman, 2009; Forsman, 2011; Franco and Haase, 2010; Alonso and Bressan, 2014) are visible through the diminishing marginal returns of breadth of search. KIE firms as captured by the sample do seem to, as theorized (Malerba et al., 2015), extensively use networks and external sources to overcome such limitations. Also, regarding a new venture's reliance on intra-industry sources of knowledge for innovative development, den Hertog et al. (2010) and Kindström (2013) previously argued that service innovations are co-created, often with clients, suppliers, and other stakeholders. These results show that in addition to this, the nature of services being co-created in general makes it difficult to distinguish the associations without more specific measurement tools.

7.0.2 The association between internal knowledge intensity and external knowledge intensity

My second research objective was to: *Explore the association between internal knowledge intensity and external knowledge intensity in the entrepreneurial firm.*

My results indicate that internal knowledge intensity as it has been used on the construct and operational levels[2] might warrant more fine-grained

[2] Initially, I defined internal knowledge intensity as the knowledge intensity that is largely

analysis, as none of our working hypotheses were fully confirmed. There are differences in how different types of internal knowledge intensity factors of entrepreneurial firms affect the reliance on different types of sources of knowledge. The strongest support of the directions of effects that we have hypothesized stems from *education levels of employees*, and from the *founding team having entrepreneurial and/or university experience.* However, we found evidence counter to our hypotheses regarding education levels and how a firm draws on intra-industry knowledge: Tertiary and PhD education levels both associated with *lower* reliance on clients, customers and suppliers as important knowledge sources. Also particularly relevant for knowledge source reliance levels across the three categories is the presence of technical and institutional opportunities as formation factors for the firm[3].

As hypothesized, I found that some aspects of internal knowledge intensity positively impact external knowledge intensity, while others have no effect or have counter-hypothetical effects. *The main takeaway is then that different sources of knowledge are affected much differently by the types of internal knowledge intensity in the firm, and future hypotheses should not be unidirectional when applied to different components or categories of external sources of knowledge.*

Neither the positive or negative aspects of functional heterogeneity[4](measured by knowledge scope and disparity, respectively) were significantly associated with external knowledge intensity in any operationalization, as was hypothesized in part due to motivation by Beckman (2006) and Classen et al. (2012), who suggest that functional heterogeneity in the form of knowledge scope of the founding team may encourage search diversity among firms. This may be related to the fact that the construction of the variables accounts for variation within quite strict *functional backgrounds*[5], not variety in previous sectoral experience or the like. Also, these effects tended to be obscured by the importance of the opportunity based formation factors (see above), suggesting that functional heterogeneity alone does not drive expansion of search, but that experience and opportunity awareness and -exploitation of founders may provide such incentive.

inherent in a firm when it comes into being, rooted in different types of human capital investments and outcomes, as well as other knowledge-based factors that have driven the firm to formation.

[3]This was captured using a principal component combining the importance of opportunities deriving from technological change, new market needs, or, new regulations or institutional requirements as factors influencing the formation of the company

[4]Or, the diversity (or lack of) of the backgrounds of the founders (Hambrick and Mason, 1984)

[5]The scale combined across the following backgrounds: Technical and engineering management, General management, Product design, Marketing, and Finance

Indeed, employee education and founder experience are again here quite good predictors of external search category reliance, at least concerning non-industry sources of knowledge. This is in line with the view that enhanced human capital leads to more expansive network ties in more knowledge intensive areas (Yli-Renko et al., 2001; Thorpe et al., 2005), and that an increased average level of education of a firms' employees leads to higher absorptive capacity and capabilities to navigate more distant search spaces (Classen et al., 2012). I find that for these firms, employee education levels on the whole prove more influential than founder education in the models, and the former might be a better approximation of the most important educational components of KIE firms. Indeed, small and micro-firm employee skills and education may be just as, if not more, important than that of the founding team for how they value external knowledge for innovation and new business opportunities. According to my results, firms with a higher educated employee base are more likely to reach outside the industry that they are active in, but my results also show that this effect may turn detrimental as education levels within the firm become very high, as informal and codified knowledge sources are less valued by firms with more highly educated employees. *Founders* with higher education levels, conversely, tended to rely more on informal and codified knowledge sources than those with lower education levels.

One striking result here again is that in my results, founder experience seems to be more associated with external search categories than founder education. The presence of university experience of the founder is related to a higher valuation of extra-industry actors, as well as intra-industry actors, as important knowledge sources. These results are echoed by those of the formation factor-specific indicators. I find that the importance of technical and institutional opportunities for founding the firm leads to heightened reliance on both extra- and intra-industry sources, signifying that when entrepreneurs in KIE-rich sector are aware of new opportunities, that they often reach out to all types of external sources (even though the results concerning codified sources was ambiguous, the other two were relatively robust). Meanwhile work and market experience within a sector perhaps leads firms to be less reliant on specialist knowledge providers and more reliant on intra-industry sources. This conveys a sort of path dependency in the way of thinking about external knowledge for entrepreneurs in all of these sectors. I find that experience, routines, and deemed relevance of outside knowledge all may be factors shaping decisions to interact more heavily with outside sources.

Theoretically, since this second model consisted mainly of working hypotheses, there is not much in the way of relating back to previous studies for these results. One can see however that, as argued by Dahlin et al. (2005), educational diversity enhances the use of diverse information for this subset of firms. Classen et al. (2015) argued that higher levels of cognitive diversity lead to better absorptive capacity for external knowledge, and this also holds for this model. Basu et al. (2015) argued for academic founders having a distinctive impact on resource diversity in the new venture, which these results also support. Finally, notions of experience of the entrepreneur within the industry guiding venture development (Agarwal et al., 2004) could be confirmed.

7.0.3 The association between internal knowledge intensity and business performance

The third research objective was to: *Explore the association between internal knowledge intensity and business performance of the entrepreneurial firm.* Internal knowledge intensity positive associated with performance through the following constructs: Functional heterogeneity of experience; Entrepreneurial and relevant industry experience; The importance of technical and institutional opportunities for founding the firm; High (up to a point) levels of education of the employee base; And, whether or not the firm formally came from a previous organization. Business performance was operationalized as the year-on-year size and revenue of the venture, and its survival during the period of study. Internal knowledge intensity negative associated with business performance through excessively high levels of knowledge scope, and high levels of knowledge disparity, among the founders in most instances, as well as a predominantly highly educated employee base in one case.

I found full confirmation only for the knowledge scope and disparity indices of the founding team, which behaved largely as expected, even in the broad sectoral and international setting employed in this study. Through knowledge scope and knowledge disparity, functional heterogeneity seems to contribute to the *internal knowledge intensity* construct as a strong predictor of economic performance. This lends credence to the assertion that high levels of positive heterogeneity combined with low levels of negative heterogeneity amplifies its business performance. It was also hypothesized that too broad a scope of founding team knowledge would have a negative effect, which was confirmed by the analysis. Concerning specific results to each construct: size, revenue and survival odds of the firm, there are some interesting take-aways as well:

Knowledge scope does play a role in business performance in terms of size and revenue in the medium-to-long term period studied here, but its effect on whether or not a firm survives is marginally conclusive, at least at lower levels of knowledge scope among the founding team. Overall, these results largely confirm those of Cantner et al. (2010), and additionally it informs that these relationships may largely hold for entrepreneurial firms in Europe irrespective of sector or country of origin as far as the data can capture, as I am unaware of this result as of yet being found on such a large scale.

Separating founding team heterogeneity is still in an early stage of theoretical and operational development, and the results here have confirmed it to be an effective tool for measuring its complex role in firm performance outcomes. These results show that the implications of scope and disparity of founding team knowledge reach beyond previous research on more high-tech manufacturing-oriented entrepreneurial firms, also encompassing those active in knowledge intensive business services and low and mid-tech manufacturing industries. The more variety and diversity in the founding team's functional backgrounds are associated with higher performance, until they become highly varied and diverse, which may hinder collaboration by the team in running the venture. It stands to reason then that the more dissimilar and redundant (knowledge disparity's root constructs) the functional backgrounds of the team are, the lower the performance, and the higher the likelihood of firm failure during the period studied.

The partially confirmed models also offer some interesting conclusions: In terms of education and experience; Employee education, founder entrepreneurial experience, opportunities driving firm formation, and being a spinoff seem to be most relevant, but the others have more mixed results. Employee education in other words was found to be associated with firm size, revenue as well as survival. This is interesting since higher education has been under fire recently, with the value of a (western) college education being called into question in terms of its benefit to entrepreneurial success (cf. Forbes, 2013; 2015; The Economist, 2015). Indeed, some are discounting the importance of the educated workforce nowadays, and especially the importance of formal founder education[6]. While my results regarding the latter are inconclusive, an educated workforce seems to play a role in determining the performance outcomes of firms in these potentially knowledge intensive sectors in the EU. Note also that there is a blurred distinction between education of founders and

[6]However, recent findings point to the crucial role of the endogeneity involved in dropout entrepreneurs' decisions to exit college in order to start a business as a predictor for entrepreneurial activity as well as performance (Buenstorf et al., 2016).

of employees in such small firms as shown throughout. These people often share office space, and work very closely with one another in many roles, especially directly following the firm founding. Future studies might drop founder indicators in favor of employee or entire firm education levels especially concerning new ventures, since this broader view of entrepreneurial firm-level knowledge might weigh heavier than individual knowledge in these environments.

Founder formal education does not seem to be associated with higher business performance. However, founder experience and opportunity awareness on the part of the founder do associate with higher business performance: This means that firms may be better off to work towards building entrepreneurial competence organically, and that experienced and driven founders combined with better-educated and motivated employees might be a good fit for success regardless of industrial character or national background. This recommendation directed at entrepreneurial firms of course must be accompanied by the caveat that different industries obviously require much different base competences and educational levels on the part of employees as well as founders. Additionally, being a spinoff (formally) or the firm being formed highly thanks to innovative opportunities does not associate with survival, but do associate with size and revenue. These offer some confirmation of Klepper's (2002) ideas about spinoff success rates and survival. And at least partially, the model is in agreement with previous studies arguing that founder human capital benefits firm size and revenue over time (Almus and Nerlinger, 1999; Klepper, 2002; Colombo and Grilli, 2005; Criaco et al., 2013).

Overall, it seems like out of all the indicators for internal knowledge intensity, founder experience and founder team heterogeneity, combined with employee education levels and the nature of the opportunities which spawned the firm, have the strongest association with heightened business performance. This shows, I argue, that knowledge intensive entrepreneurial firms are indebted to their employees as well as their founders for creating an effective, sustainable and progressive business organization.

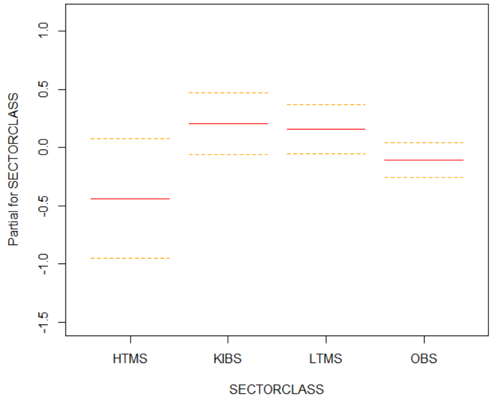
7.0.4 The association between innovative performance and business performance

My fourth research objective was to: *Explore the association between innovative performance and business performance of entrepreneurial firms.* While the working hypotheses constructed were quite broad in nature, they largely yielded the expected directions of results. Fully confirmed hypotheses indicate that radicalness of innovation in these firms positively impacts business performance in all three indicators; sales, revenue, and likelihood of survival. Sales of innovative products and services did not prove to be a conclusive indicator in fully specified models, but without radicalness included, they yielded overall similar results, with the exception of innovative products sales being negatively related to survival[7].

The role of share of innovative goods to that of total sales in the survival model was insignificant, and problematic in achieving full confirmation. One interpretation, when one combines this result with that of radicalness having a positive association with survival, is that the more heavily one invests in new innovations in terms of product portfolio by having a large percentage of revenue being based on innovations, the more at risk for exit the firm may become over time. This relationship is not apparent for firms that rely heavily on service innovations in their revenue stream, suggesting that manufacturing firms are potentially more at risk for exit. Additionally, a larger share of innovative sales could be an indicator of a firm early in the commercialization process, before more standardized goods become the dominant form of revenue. I interpret this to mean that manufacturing firms may be particularly at risk for exit in sectors with a high potential for KIE firms. By viewing the effect plot in Figure 7.3 from the full Cox survival model for the sectoral control variable, we can see a clearer picture of what might be going on: There is a visible and statistical difference between survival rates in low- to mid-tech manufacturing sectors (LTMS) and high- to mid-tech manufacturing sectors (HTMS), in that the former are distinguishably more at risk for exit. What may be the case for low- to mid-tech firms is that they meet with more stringent expectations from the market concerning their goods and services, and often, implementation of new products in a low tech market will result in competition with already established products in fairly well-saturated markets. In any case, low- to mid-tech firms may be experiencing higher failure rates related to their attempts to introduce more innovativeness in goods.

[7] As well as to firm growth in sales in a pre-imputation growth model

Figure 7.3: Generalized hazard rates for entrepreneurial firms by sector from Model 4.3



Summarily, innovativeness in goods is positively associated with performance, but is also connected to a higher likelihood of failure, or risk. Survival rates however are improved if the firm moves towards more new-to-the-world innovations. These results move beyond the studies of innovation's benefit for entrepreneurial firms in both goods (Hughes, 2001; Van Auken et al., 2008; Gunday et al., 2011) and services (Cainelli et al., 2006; Ariana Mansury and Love, 2008), and also warns of its riskiness. As Coad et al. (2016) have argued with R&D investments, more innovativeness in ratio to sales output can have a strong positive effect if successful, otherwise there can be a dangerous risk for decline and/or failure. While the innovation process of KIE firms is often beneficial in the long run for the economy as a whole, it can be hazardous to the firm itself. Overall, the results show that on average the more 'innovative' the entrepreneurial firm, the higher the performance (all else equal) for this subset. This could mean that those firms in KIE likely sectors that are more novel in their goods and services tend to succeed more, though firms that invest too heavily in new goods in their product portfolio could be at risk. In other words; investing in, creating, and applying new knowledge through novelty in innovations seems to associate with both survival, financial and size-based performance. And in the limited pre-imputation sample, growth rates seem potentially positively related to degree of radicalness as well for both size and revenue (see Appendix for regression models).

7.1 Knowledge, innovation, performance and the entrepreneurial firm: A re-visitation

I argue from this doctoral thesis that *entrepreneurial firms' differing forms of knowledge intensity do not entirely associate positively with overall firm performance, though the majority of indicators used point towards this relationship. It is as so much else, not entirely a simple matter. Additionally, I find that innovativeness in terms of level of radicalness of innovations affects, or consistently associates with, heightened business performance over time. I do not find any systematic hypothesized associations between external and internal knowledge intensity. However, this could be related to the wide range of variables used to assess overall constructs.*

In Figure 7.1 previously shown, my fully confirmed results show that the functional heterogeneity as well as radicalness of innovations affects the business performance over time of entrepreneurial firms, while breadth and depth have respectively a curvilinear and linear effect on the innovative performance of these firms. Therefore, *some* aspects of internal knowledge intensity associate with external knowledge intensity, which in turn associates with innovative performance, some aspects of which in turn drive business performance, the latter also being directly affected by some aspects of internal knowledge intensity. This results in a semi-circular loop where internal knowledge intensity is the exogenous predictor, and all other concepts are in one way or another a result of this.

I argue then that how a firm draws on or augments its knowledge intensity can then be an internal- or external-to-the-firm process, and is always a combination of the two. In my view, *internal knowledge intensity* <u>on the firm level</u> stems from the concentration of manifest knowledge produced by new combinations of individual knowledge that, when solidified into new methods of interaction between individuals, creates something new on the organizational level in the form of collective or shared understanding, information, resources or capabilities.[8] I have proxied these creative combinations by looking at firm pre-history conditions and employee and founder human capital. I therefore define *external knowledge intensity* <u>on the firm level</u> as; how and to what extent

[8] Also, I don't entirely discount some crucial aspects of knowledge intensity across industries, such as novelty or innovation of organizational practice or routine, and increased intensity of interface with end users in devising co-terminal services. I have left these important factors are left out of the current build of the definition due to a lack of sufficient testability in the data.

a firm both applies and values external-to-the-firm know-how or knowledge, and combines it with its inherent resources to create new combinations, also resulting in new organizational behavior or action manifested in firm-level knowledge or understanding.[9]

Both of these two above-described types of knowledge intensity are inter-related, and may amplify the likelihood of an entrepreneurial firm being more innovative on different levels. I argue that this knowledge intensity can be seen as stocks and flows in character. By this I mean that at any given time there is a base level of knowledge intensity found within a firm which is dynamically augmented by the firms' inter-actions and intra-actions through which it recombines resources to create something resembling the Schumpeterian notion of 'new combinations'. Much, but unfortunately not all, of the empirical material confirms my approximations of these constructions and their inter-relationships, but one must hold in mind that the criteria for confirmation in this aspect were quite conservative (a fact of which I am aware, and I leave the reader to assess the strength of these conclusions based on this). For instance, I have found that different types of internal knowledge intensity will have differentiated associations with external knowledge sources, and how intensely they are valued. The education levels of the firm are important, and not just that of the founders. Firms with high proportions of tertiary educated personnel (the typical indicator for knowledge intensity on the sectoral level is 33% tertiary education) seem to be more adept at reaching outside their industries to other actors for new business opportunities, and less reliant on inter-industry actors.

These conclusions have helped me to illustrate that knowledge intensity, though it does have its idiosyncrasies dependent on the environment it occurs in, can be mapped in a more or less consistent ways in terms of entrepreneurship and how it relates to performance; at least in a European context.. By controlling for sectoral and regional differences, I have established a generalized picture of knowledge intensive entrepreneurial activity in the European Union circa 2011, and some indication of development over time of these firms in the years that followed until the present day. Knowledge, innovation and performance in entrepreneurial firms are clearly inter-related and -dependent which serves to confound any sense of causality, but some patterns of association are distinct.

[9]In lieu of the ability to quantify the actual intensity of how these external processes actively augment resource and capability stocks of a firm, I relied mainly on how much the firm relies on such external sources for new opportunity exploitation.

7.1.1 Implications for policy

Policy to support knowledge intensive entrepreneurship will likely benefit for the work carried out in this text. Much of the findings here indicate that there are indeed patterns visible across sectors and regions in terms of knowledge intensive activities. I find that there is an indication that higher education among employees is associated with successful performance of new ventures, as is entrepreneurial experience of the founding team. This is also the case for high levels of beneficial functional heterogeneity like diversity, while maintaining low levels of redundancy and dissimilarity among the founding team. Since policy is not often able to address so many specific variables simultaneously, I argue that a simpler approach may be warranted. Indeed, many question the validity of a specialized sectoral approach to growth and development promotion through strategic governmental action. There have been calls for policy measures to become broader and less specialized. For example, recently Feldman and Choi (2015: 291) expressed that, based on work done towards re-conceptualizing what economic development actually means as opposed to economic growth:

> *"The paradox of place-specific economic development policy is that broad-based government investments in education and infrastructure are critical to future economic growth. Targeting certain sectors, specific industries, or isolated components of the innovation ecosystem is unlikely to succeed if basic capacity is lacking."*

This sentiment (while informed much more by economic geography, agglomeration externality theory, and a more spatial view of innovation based on knowledge spillover theory than my own contributions) echoes a point similar to some of the empirical findings that I demonstrate here. Prevalence of tertiary (and at times, even more advanced) education of employees, across sectors, does seem to matter for new venture performance across sectors. The experience of entrepreneurs also has been a prominently effective component in my analyses. It could be that by broadening investment strategies, human capital components of companies' internal knowledge intensity could be more effectively augmented. The same work by Feldman and Choi (2015) emphasizes that:

> *"(t)he best economic development strategy is to enable as many actors to productively participate in the economy to the fullest of their ability. This prioritizes improving quality of life and well-being by enhancing capabilities and ensuring that agents have freedom to achieve."*

Perhaps instead of trying to promote certain sectoral activities, which traditionally has been focused on high-tech areas of the economy, I argue that it might be better for policy to reduce risk and increase incentives for already active and budding entrepreneurs, so that they might increase the experiential component of their companies' knowledge intensity, and by investing in education, improve the overall pool of educated employees available to the founding teams. While admittedly a more long run strategy, it seems topical. The understanding of the nature of innovation, the process and learning and knowledge creation, and the role of government in facilitation and supporting these things (Smith, 2005) remain crucial for designing future innovation policy.

7.1.2 Future research

To point out a few more tangible additional interesting avenues:

Carrying out a more thorough mapping of research and development, and how it structurally relates to internal and external knowledge intensity in KIE sectors seems prudent. The presence of curvilinearity in the results regarding R&D intensity could also benefit from more research. One striking result from this dissertation is the near ubiquitous importance of research and development in the first two empirical model sets, and its declining marginal effect. Given the overwhelming focus on external sources of knowledge as important for innovation in the literature since the Open Innovation paradigm gained notoriety with Chesbrough's seminal book (2003), researchers have laid less focus on internal R&D as crucial knowledge activities. My results suggest not only is it still important in the Open Innovation paradigm, but its effect often eclipses that of the main explanatory variables that I have used.

Conducting a follow up survey more tailored to investigating internal and external knowledge intensity as derived by resource based theory. Since my own conceptualizations occurred separately of data gathering and survey design, I was somewhat limited in my choice of operational definitions and variables to represent underlying concepts. Additional summated rating scales of knowledge intensity and innovative performance indicators in low- and mid-tech sectors as well as services as I have revisited in Chapter 2 would be immensely useful. Since low- and mid-tech have tangibly different knowledge and innovation landscapes at times, more detailed information would be welcome.

Digging deeper into external sources of knowledge for innovation used by different types of sectoral actors in both services and manufacturing in the EU and beyond, and how these are influential for innovation outcomes. I feel I have only scratched the surface here, and more sophisticated techniques might see application here, including more rigorous use of multidimensional scaling and factor analysis of more carefully constructed questionnaires.

Looking more into organizational innovation and training, as these are commonly cited components of knowledge intensity in low tech, medium tech, and service sectors. Again, the data was a bit restrictive in allowing for full scale analysis of these types of variables in conjunction with the ones that were chosen.

Working towards the creation of less problematic overlap in innovation performance indicators in survey design. In lieu of this, one thing that has received limited attention but could be quite fruitful is to apply some type of quantile regression approach the proportion variables representing innovative performance in terms of sales ratios. Quantile regression, a highly effective form of robust regression, could possibly show marked differences in the data space. Unfortunately this technique is not readily applicable to the multiple imputed datasets like the one utilized here. This brings me to the next point quite precisely.

Collecting more detailed and robust longitudinal data about potentially knowledge intensive entrepreneurial firms. My own data collection was somewhat plagued with missingness, and some crucial assumptions about the nature of this missingness had to be made in order to render certain variables usable. More thorough treatment and retention of key variables of interest is advised in future works.

Continuing to unpack the importance of educational attainment of the entire firm in KIE sectors, and whether founder educations matter so much compared to this (as I have suggested above the two have a likely overlap).

Using what we now know about KIE, sampling a new group of firms using a more sophisticated method than merely sectoral likelihood of KIE existing. This is admittedly difficult, time consuming, and expensive, but there is worth in it for future projects and research.

7.1.3 In closing

Knowledge intensive entrepreneurship as a research concept is a challenging one, but it is also a very important one. For it is the essence of what many judge to be one of the main explanatory factors of economic change, growth, and well being in the modern age. I have made some rather large assumptions about the nature of the connection between KIE success and economic development, prosperity, growth and well being, and chosen to focus on the nature of knowledge intensity itself, and what it might mean if a resource based view is applied. It may be helpful to view it as a problem of resources and output when viewed from a policy standpoint as well, with firms as the population, and some type of natural selection mechanism choosing the winners instead of the nation-states and their policy makers. Providing fertile ground for knowledge intensive entrepreneurship in all types of sectors may be best achieved through broader policy measures than those that are popular at present. The AEGIS project represented an initial explorative venture into the phenomenon of knowledge intensive entrepreneurship, and many lessons have been learned about how this phenomenon might best be approached and grappled with. We have already begun to see an increased focus on prescribing more systemic approaches to fostering economic growth, urging policy makers to take a step back and consider the broad and complex situation involving promoting and retaining knowledge intensive activities. AEGIS and KEINS have prescribed much the same medicine, and I can only add my voice to the tumult. Already, though, it seems as though policy is shifting gears towards towards new concepts: At the time of writing this section, the OECD is shifting to a focus on global productivity, and firms at the global productivity frontier. A paper by Andrews et al. (2015) advocates framework policies to aid productivity diffusion by sharpening incentives for firms to adopt new technologies, and the reallocation of resources to more productive firms. These authors lay quite a heavy focus on both turnover as well as patenting stock as indicators of productivity, and the OECD (2015: 7) professes it expects productivity *"to be the main driver of economic growth and well being over the next 50 years, via investment in innovation and knowledge-based capital."* I view this to be a bit of a misstep, in that productivity concerns were a large driving factor in investigating knowledge intensive activities, and it seems that policy makers may be falling back on a more, for them at least, easily understood characterization of economic growth agents. It is my sincere hope that this direction does not disregard all that has been learned about knowledge-based, or knowledge-intensive entrepreneurship in the past few years, and that it is kept in context that knowledge intensity is made up of more than just education, although this does seem empirically to be one of the more robust components in terms of firm performance. Additional complications may arise due to the decreasing stability of liberal and globalizing economic policies that are occurring at the time of this writing. The co-evolution of innovation policy and innovation research has brought about many changes in the way governments, researchers and entrepreneurs interact with science and technology in order to create knowledge in the past decades (Smith, 2005), and for the next step to occur there is a need for clarity of purpose and the ability to act among all involved. Learning to distinguish between different types of knowledge, and different types and degrees of knowledge intensity, is a crucial step in this development. In this respect, a more careful approach focused on the application and use of such knowledge, and how it diffuses in different context, rather than just an overall motivation or incentive to adopt new technologies for new firms, might be of higher worth.

Post Scriptum

I would like to include a few key points regarding the research covered in these pages that I found hard to place, so much of it found its way here. Throughout this dissertation I have grappled with ideas concerning knowledge, the growth of knowledge, and innovation. What I hope I have achieved is at least to stress the complexity of assigning simplified indicators to complex problems. For instance, equating a sector's level of knowledge intensity to its mean level of educational achievement is something that should not be done lightly. J.S Metcalfe (1998:122) said that:

> *"[in] an evolutionary system, the role of policy is to facilitate the ongoing development of innovative variety, not by second guessing the market but by creating the conditions under which innovations flow more easily. In this the policy making is not seeking to optimize the exploitation of given opportunities, but rather to adaptively create the conditions for the emergence of new opportunities. Government can neither predict which are the likely innovations nor the promising markets. Rather its proper role is to build an infrastructure in support of firms and let the innovations follow from the market process."*

This message is one that lies near my own in completing this work. Hopefully, my work has shown that despite the fact that the EU is made up of diverse countries with differing institutional settings and heritage, and therein diverse sectoral systems, the constructs that I have used to represent knowledge intensive activity by and large are at least in some ways consistent across the region, regardless of sector and place. We see differences in magnitude of course, but often the effects are present anyway. Taking an evolutionary perspective, policy makers should be hesitant to 'pick winners' based on sectoral systems or to target entrepreneurial support based on 'knowledge intensity', for it can be seen here that the concept flows into all sectors, in all regions, and it is on average important for new woodworking ventures, logistic service providers, as well as biotech startups. Baseline support like improving educational infrastructure and employability of graduates ought to take priority over tailoring programs to drive up the amount of entrepreneurship. In the Schumpeterian sense, true entrepreneurship involves by definition innovation, and by stimulating the base infrastructure on which industries are built instead of existing "high potential" industries, more of this might come into being. From an evolutionary standpoint, the knowledge intensity of the firm is idiosyncratic: A unique combination of resources of capabilities driving firm performance. We need to recognize how complex inter-relationships between sectors and firms bring about societal progress and change from within the business system, and that new knowledge may find its end application in an entirely different context that that within which it was devised.

Regarding my own next steps: Much has been learned about the inter- and intra-relationships between the concepts and constructs employed in this study, but the picture still lacks cohesion. With that in mind, the next step in this research is to employ more advanced empirical techniques to construct a structural model which may simultaneously map the effects of all models at once. Some methods used here, for instance, the MI-GEE models, are more limiting than a properly conceived latent growth curve model might be, which can take into account all the inter-relationships between the latent constructs, as well as introducing a time dimension only in certain linkages in the structural model (see Bollen and Curran, 2006). This might better capture the essence of my arguments, since it solved much of the potential endogeneity issues that plague this type of associational research. Moreover, this would allow for measurement of coefficients representing not only the latent variables that I glimpsed in the principal component analyses, but also those of internal and external knowledge intensity, and economic and innovative performance more directly. The first step will be to assess the fully specified path models of each of the broader concepts that I draw on, that is, both types of knowledge intensity, innovation, and performance, separately at first, and analyze the different covariances and correlations that may exist between the residuals, before combining all the components into a full structural model. That, I intend, will be forthcoming following the publication of this work. In doing this it will also be possible to prune the number of variables in the analysis to those that fit the structural model. While much of this has been done already, it could not be fit into the framework of this dissertation. Also, statistically significant results in the non-hypothesized direction that are counter-intuitive or contrary to the theory drawn upon, will be more closely examined, by way of gathering more comprehensive data, if possible, from these same firms.

The next research steps taken will likely to be prune those inefficient predictors in order to shape a more systematic model of associations between variables, and thus between constructs. Great care must be taken here in ensuring that we come as close to the concept as we can with our measurement. This is no easy task, but it seems that in some places, some degree of faith must be put into the effectiveness of the operationalization. For instance, perhaps relying solely on percentage of tertiary employees would have been better than additionally including founder education levels (ordinally scaled) and the percentage of more advanced degree holders in the firm. This is because the founder variable might also be captured in the employee variable, as the latter largely indirectly reflects the founder's own education, and the advanced degree variable is actually a subset of the tertiary degree variable. True, it captures nuances the other cannot, but these cannot effectively be separated from already accounted for variance. Additionally, the spinoff measures employed here are perhaps not as aligned as they could be conceptually. I chose to utilize two indirect and one direct definition for spinoffs; the university experience (dummy) and the industry experience (age of founder), and the actual corporate origin of the firm (stemming from a previous firm). Simplification might also have been of worth in this area.

I would also like to use this space to address further the absence of what could be a considered a key theoretical relationship in the framework developed in this thesis. That is namely the link between my concept of internal knowledge intensity, and that of innovative performance. There are a number of reasons, as well as potential problems, that I should like to make apparent to the reader. The first being the following: Of all that could be analyzed given this particular framing of the knowledge intensive entrepreneurial firm, the link between entrepreneurial venture pre-history and human capital with that of innovative performance is one of the most researched relationships to date. Though not directly applied to knowledge intensive entrepreneurship, there are numerous studies in the field that tackle how founder and employee development, education, experience and administrative heritage influence the outcomes of how innovative a new firm might be as a result (e.g. Helfat and Lieberman, 2002; Caloghirou et al., 2004; Leiponen, 2005; Weterings and Koster, 2006; Marvel and Lumpkin, 2007; Cantner et al., 2010; Andersson and Lööf, 2012; Basu et al., 2015; Kristinsson et al., 2016). Doing so in the context of the present study would not only be mainly replicative in nature, but largely confirmatory rather than exploratory work. There is of course nothing wrong with this, but given the wealth of complexity present in the analysis without including this dimension, it seemed best to leave this linkage out of the full model.

Nonetheless, by measuring the association between internal knowledge intensity and external knowledge intensity, and the association between external knowledge intensity and innovative performance, there is in effect a link between internal knowledge intensity and innovative performance. Granted, this is in indirect effect, and there is likely some degree of effects one fails to account for by excluding the direct relationship between these two, which will present a potential endogeneity bias in the empirical analysis[10]. One treatment of the problem that might alleviate this is the use of structural equation modeling or two stage least squares regression, where more complex models, correlations, and covariance structures especially between residuals may be specified. It may be said then that including direct effects between the two constructs would likely moderately weaken the statistical power of the indirect effects.[11]

The way I have conceptualized internal and external knowledge intensity has been influenced by current concepts in social science and economics like human capital and search, respectively. It should be made clear that while there may exist other constructs that might aptly measure knowledge intensity in addition to these, I have made the choice to emphasize those that appear in this work. This came not only from limitations placed on me, the researcher, by the available data, but also in terms of clarity of measurement. Given the data, I was limited to the different variables available therein, and also the way in which the information was structured in terms of summated rating scales which one could then analyze.

Of course, external knowledge intensity could potentially go beyond the idea of search; much knowledge that exists external to a firm is not sought or searched for, and yet is absorbed into or used by various mediums nonetheless. For instance, in accordance with innovation systems approaches, institutions and regulations that steer the development of the firm before it is even conceived of likely play a major role in shaping the outcomes of a venture. The best I can do here is to control for sectoral and regional differences, which while on the whole may be unsatisfactory, it is still a step closer to getting at the essence of the definition of the KIE firm and what it means for society at the general level.

Much of what I focus on here as external knowledge intensity is concerned with importance of external knowledge sources for generating new business opportunities, but the *intensity* could vary across multiple dimensions, not merely how important these sources are, how they contribute to specific types of innovation and business development, and how much these themselves affect firm performance, etc. The survey data

[10] This is of course a problem in most areas of social science, as there may always be unaccounted for effects in regression analyses

[11] Indeed, in a simplified trial structural equation model mapping these specific relationships using latent variables, I find exactly this. During robustness checks I have added this relationship into large scale test models and not found any substantive differences

used looks at aggregated phenomena and does not trace any specific innovation pathways from specific invention to a specific product. Regardless, my attempt here, as detailed in Chapter 3, was to assess this type of firm as a unique product of its resources and capabilities, rationality, and its external and internal environment. To the extent possible I have used this unique data source to frame KIE activity in a resource-based context.

Another potentially problematic insight is that knowledge intensity, as I measure it, does not always allow for the measure of some type of magnitude, or relative change, which would lend further credence to the term *intensity*. Many of the operationalizations do deal with magnitudes: depth and breadth of search, as well as the principal components used to approximate external knowledge sources and factors important to firm formation do so. To some extent, it could be argued that levels of education, proportions of skilled or educated employees, and functional heterogeneity capture some type of continuum which may be more or less intense. Admittedly, the use of dummy variables complicates this, since something like a member of the founding team having prior entrepreneurial or university experience is basically an all or nothing indicator. More useful might have been some more complementary information on patenting or startup failure rates of founders, venture capital backing, etc. This type of data was however not available for these firms in the given time frame, and sometimes if it was present, the quality was quite poor (especially concerning patenting and venture capital).

One final word on the levels of performance and the time-frame analyzed in this thesis: Concerning the business performance metrics of the 3rd and 4th models, I have mainly focused on the growth intercepts, that is, levels of different variables between 2010 and 2014, and how these have been affected by, or associated with, various explanatory variables. As I have noted, I also carried out logarithmic change-based growth regressions for the whole period, which yielded some results but only for innovative performance as affecting growth (See Appendix, Tables 8.14 - 8.17). What this tells us is that it is more likely that firm pre-history sets the trajectory of growth, but does not necessarily influence year by year change. There are an abundance of other effects that the models I was able to construct here have become endogenous to the residual terms, that is, I cannot separate them from randomness. What is clear though is that while internal knowledge intensity does not seem to systematically affect firm growth in terms of sales, size and survival, innovative performance does have some tangible, if somewhat inconsistent, effects. One major takeaway from this is that, if one extrapolates so that high-growth firms in sectors with higher likelihoods for KIE further enrich the economy and provide societal benefits, then more knowledge intensive entrepreneurial firms also may dramatically impact economic and societal growth and well being over time. This conjecture requires of course much more thorough data collection before any sort of concrete hypothesis might be tested.

References

Abernathy, W.J., Clark, K.B, (1984) “Innovation: Mapping the winds of creative destruction”, *Research Policy*, 14(1), 3-22.

Acock, A.C. (2005) “Working With Missing Values”, *Journal of Marriage and Family*, 67, 1012-1028.

Acs, Z., Audretsch, D.B. (eds.) (2003) *Handbook of Entrepreneurship Research*, Springer Science & Business Media.

Acs, Z.J., Audretsch, D., Braunerhjelm, P., Carlsson, B. (2009) “The knowledge spillover theory of entrepreneurship”, *Small Business Economics*, 32(1), 15-30.

AEGIS Research Project (2013) Advancing Knowledge-Intensive Entrepreneurship and Innovation for Economic Growth and Social Well-being in Europe, Grant Agreement number 225134, European Commission, DG Research, Brussels.

Agarwal, R., Echambadi, R., Franco, A.M., Sarker, M.B. (2004) “Knowledge Transfer through Inheritance: Spin-out Generation, Development, and Survival”, *The Academy of Management Journal*, 47(4), 501-522.

Agarwal, R., Braguinsky, S. (2015) "Industry evolution and entrepreneurship: Steven Klepper’s contributions to industrial organization, strategy, technological change, and entrepreneurship", *Strategic Entrepreneurship Journal*, 9(4), 380-397.

Almus, M., Nerlinger, E.A. (1999) “Growth of new technology-based firms: which factors matter?”, *Small Business Economics*, 13, 141-154.

Alonso, A., Bressan, A. (2014) "Collaboration in the context of micro businesses: The case of Terracotta artisans in Impruneta (Italy)", *European Business Review*, 26(3), 254-270.

Alvesson, M. (1995) *Management of knowledge-intensive companies*, de Gruyter: Berlin; New York.

Amit, R., Schoemaker, P.J.H. (1993) “Strategic assets and organizational rent”, *Strategic Management Journal*, 14, 33-46.

Andersson, M., Lööf, H. (2012) "Small business innovation: firm level evidence from Sweden" *The Journal of Technology Transfer*, 37(5), 732-754.

Andrews, D., Criscuolo, C., Gal, P (2015) “Frontier firms, technology diffusion and public policy: Micro-evidence from OECD countries”, OECD Productivity Working Papers, ISSN 2413-9424.

Ariana Mansury, M., Love, J.H. (2008) “Innovation, productivity and growth in US business services: A firm-level analysis”, *Technovation*, 28, 52-62.

Arora, A., Cohen, W.M, Walsh, J.P. (2016) "The acquisition and commercialization of invention in American manufacturing: Incidence and impact", *Research Policy*, 45(6), 1113-1128

Audretsch, D.B., Keilbach, M. (2006) “Entrepreneurship: A Survey of the Literature,” Report compiled for KEINS, Work package 1, Paper 2, Deliverable No.3.

Audretsch, D.B., Thurik, R. (2001) “What’s New about the New Economy? Sources of Growth in the Managed and Entrepreneurial Economies,” *Industrial and Corporate Change*, 10(1), 267-315.

Arrow, K. (1962) “Economic welfare and the allocation of resources for invention”, in *National Bureau of Economic Research, The Rate and Direction of Inventive Activity*, 609-625, Princeton University Press: Princeton, NJ.

Attneave, F. (1959) *Applications of Information Theory to Psychology*, Holt, Rinehart and Winston: New York.

Autio, E. (1997) “New, technology-based firms in innovation networks symplectic and generative impacts”, *Research Policy*, 26(3), 263-281.

Autio, E., Sapienza, H.J., Almeida, J.G. (2000) “Effects of Age at Entry, Knowledge Intensity, and Imitability on International Growth”, *The Academy of Management Journal*, 43(5): 909-924.

Autio, E., Acs, Z.J. (2007) “Individual and country-level determinants of growth aspiration in new ventures. In A. Zacharakis (ed.), *Frontiers of Entrepreneurship Research* 2007. Babson Park, MA: Babson College.

Bain, J.S. (1968) *Industrial Organization*, Wiley: New York

Baines, T.S., Lightfoot, H.W., Benedettini, O., Kay, J.M. (2009) “The servitization of manufacturing: A review of literature and reflection on future challenges”, *Journal of Manufacturing Technology Management*, 20, 547-567.

Baptista, R., Karaöz, M., Mendonça, J. (2014) "The impact of human capital on the early success of necessity versus opportunity-based entrepreneurs", *Small Business Economics*, 42, 831-847.

Barbaroux, P. (2014) "From market failures to market opportunities: managing innovation under asymmetric information", *Journal of Innovation and Entrepreneurship* 3(1).

Barker, V.L., III, Mueller, G.C. (2002) "CEO Characteristics and Firm R&D Spending", *Management Science*, 48(6), 782-801.

Barney, J.B. (1991) "Firm Resources and Sustained Competitive Advantage," *Journal of Management*, 17(1), 99-120.

Barney, J.B. (1996) "The Resource-Based Theory of the Firm", *Organization Science*, 7(5), 469.

Basu, S., Sahaym, A., Howard, M.D., Boeker, W. (2015) "Parent inheritance, founder expertise, and venture strategy: Determinants of new venture knowledge impact", *Journal of Business Venturing*, 30, 322-337.

Becker, G.S (1964) *Human Capital*, Columbia University Press: New York.

Beckman, C.M. (2006) "The Influence of Founding Team Company Affiliations on Firm Behavior", *The Academy of Management Journal*, 49(4), 741-758.

Bell, D. (1973) *The Coming of Post-Industrial Society*, Basic Books: New York.

Bengtsson, L., Lakemond, N., Lazzarotti, V., Manzini, R., Pellegrini, L. and Tell, F. (2015) "Open to a Select Few? Matching Partners and Knowledge Content for Open Innovation Performance", *Creativity and Innovation Management*, 24(1), 72-86.

Birch, D. (1979) *The job generation process*: MIT Program on Neighborhood and Regional Change, Cambridge, MA.

Birch, J. (1989) "Change, innovation, and job generation", *Journal of Labor Research*, 10(1), 33-39.

Blalock, H.M. Jr. (1982) *Conceptualization and Measurement in the Social Sciences*, Sage: London.

Blau, P.M. (1977) *Inequality and heterogeneity*, Free Press: New York.

Bloch, H., Metcalfe, J.S. (2011) “Restless Knowledge, Capabilities and the Nature of the Modern Firm”, Centre for Research in Applied Economics Working Paper Series: no 201103, Curtin University of Technology, School of Economics and Finance.

Block, J.H., Fisch, C.O., van Praag, M. (2016) "The Schumpeterian entrepreneur: a review of the empirical evidence on the antecendents, behaviour, and consequences of innovative entrepreneurship", *Industry and Innovation*, available online at: http://dx.doi.org/10.1080/13662716.2016.1216397.

Boeker, W. (1987) “Strategic Origins: Entrepreneurial and Environmental Imprinting at Founding”, *Academy of Management Proceedings.* 1987(1). Academy of Management.

Boeker, W. (1988) “Organizational origins: entrepreneurial and environmental imprinting at the time of founding”, in Carrol, G. (ed.), *Ecological Models of Organizations*, p. 33-51, Ballinger: Cambridge, MA.

Boeker, W. (1997) “Executive migration and strategic change: The effect of top manager movement on product-market entry”, *Administrative Science Quarterly*, 42(2), 213-236.

Bollen, K.A, Curran, P.J. (2006) *Latent Curve Models - A Structural Equation Perspective*, Wiley-Interscience: Hoboken, NJ.

Boone, C., Hendriks, W. (2009) “Top Management Team Diversity and Firm Performance: Moderators of Functional-Background and Locus-of-Control Diversity”, *Management Science*, 55(2), 165-180.

Bosma, N., van Praag, M., Thurik, R., de Wit, G. (2004) “The Value of Human and Social Capital Investments for the Business Performance of Startups”, *Journal of Small Business Economics*, 23, 227-236.

Box, G.E.P, Cox. D.R. (1964) “An analysis of transformations”, *Journal of the Royal Statistical Society*, Series B (Methodological), 26(2), 211-252.

Brant, R. (1990) "Assessing proportionality in the proportional odds model for ordinal logistic regression", *Biometrics*, 46, 1171-1178.

Brüderl, J., Prisendörfer, P., Zeigler, R. (1992) “Survival chances of newly founded business organizations”, *American Sociological Review*, 57(2), 227-242.

Bruni, D.S., Verona, G. (2009) "Dynamic Marketing Capabilities in Science-based Firms: an Exploratory Investigation of the Pharmaceutical Industry", *British Journal of Management*, 20, 101-117.

Brunswicker, S., Vanhaverbeke, W. (2015) "Open innovation in small and medium-sized enterprises (SMEs): External knowledge sourcing strategies and internal organizational facilitators." *Journal of Small Business Management*, 53(4), 1241-1263.

Brush, C., Chaganti, R. (1999) "Businesses without glamour? An analysis of resources on performance by size and age in small service and retail firms", *Journal of Business Venturing*, 14, 233-257.

Brush, C.G., Greene, P.G., Hart, M.M. (2001) “From initial idea to unique advantage: the entrepreneurial challenge of constructing a resource base”, *Academy of Management*, 15, 64-78.

Buenstorf, G., Nielsen, K., Timmermans, B. (2016) "Steve Jobs or No Jobs? Entrepreneurial activity and performance among Danish college dropouts and graduates", *Small Business Economics*, online Early View.

Bunderson, J.S., Sutcliffe, K.M. (2002) “Comparing alternative conceptualizations of functional diversity in management teams: Process and performance effects”, *Academy of Management Journal*, 45, 875-893.

Burns, A.F. (1934) *Production Trends in the united states since 1870*, MIT Press: Boston.

Bygrave, W., Hunt, S. (2005) *Global Entrepreneurship Monitor*, 2004 Financing Report, Wellesley – London: Babson College and London Business School.

Cainelli, G., Evangelista, R., Savona, M. (2006) ”Innovation and economic performance in services: a firm-level analysis”, *Cambridge Journal of Economics*, 30(3), 435-458.

Caloghirou, Y., Kastelli, I. Tsakanikas, A. (2004) "Internal capabilities and external knowledge sources: complements or substitutes for innovative performance?", *Technovation*, 24(1), 29-39.

Caloghirou, Y., Protogerou, A., Tskanikas, A. (2011) AEGIS D7.1.5 Final report summarizing survey methods and results, 1-134.

Caloghirou, Y., Protogerou, A., Tskanikas, A. (2015) "Knowledge Intensive Entrepreneurship: Exploring a taxonomy based on the AEGIS survey", in Malerba, F., Caloghirou, Y., McKelvey, M., Radosevic, S. (eds.) (2015) *Dynamics of Knowledge-Intensive Entrepreneurship: Business Strategy and Public Policy.* Routledge: London.

Camacho, J.A., Rodriguez, M. (2005) "How innovative are services? An empirical analysis for Spain", *Service Industries Journal*, 25(2): 253-271.

Campbell, C.A. (1992) "A Decision Theory Model for Entrepreneurial Acts", *Entrepreneurship Theory and Practice*, 17(1), 21-27.

Campbell, C.A. (1995) "An empirical test of a decision theory model for entrepreneurial acts", *Entrepreneurship and Regional Development*, 7, 95-103.

Caniels, M.C.J. (2005) "What drives innovativeness in industrial clusters? Transcending the debate", *Cambridge Journal of Economics*, 29, 497-515.

Cantner, U., Goethner, M, Stuetzer, M. (2010) "Disentangling the effects of new venture team functional heterogeneity on new venture performance", Jena economic research papers, N 2010, 029, Max Planck Institute of Economics.

Cantwell, J. (2002) "Innovation, profits, and growth: Schumpeter and Penrose", in Penrose, E., Pitelis, C., *The growth of the firm: the legacy of Edit Penrose*, Oxford University Press on Demand.

Carayannis, E.G., Alexander, J. (1999) "Winning by co-opeting in strategic government-university-industry partnerships: the power of complex, dynamic knowledge networks", *Journal of Technology Transfer*, 24(2-3), 1-14.

Carroll, G., Hannan, M. (2000) *The Demography of Corporations and Industries*, Princeton University Press: Princeton, NJ.

Cattell, R.B (1966) "The scree test for the number of factors", *Multivariate Behavioral Research*, 1, 245-276.

Chambers, J.M., Hastie, T.J. (1992) *Statistical Models in S*, Wadsworth & Brooks/Cole.

Chandler, G.N., Hanks, S.H. (1993) "Measuring the performance of emerging businesses: A validation study", *Journal of Business Venturing*, 8, 391-408.

Chandler, G.N., Hanks, S.H. (1994) "Market attractiveness, resource-based capabilities, ventures strategies, and venture performance”, *Journal of Business Venturing*, 9, 331-349.

Chandler, G.N., Honig, B., Wiklund, J. (2005) "Antecedents, moderators, and performance consequences of membership change in new venture teams", *Journal of Business Venturing*, 8(5), 391-408.

Chell, E., Baines, S (2000) "Networking, entrepreneurship and micro-business behavior", *Entrepreneurship and Regional Development*, 12(3), 195-215.

Chesbrough, H. (2003) *Open Innovation: The new imperative for creating and profiting from technology*, Harvard University Press: Cambridge, MA.

Chowdhury, S. (2005) "Demographic diversity for building an effective entrepreneurial team: Is it important?" *Journal of Business Venturing*, 20(6), 727-746.

Clarysse, B., Tartari, V., Salter, A. (2011) "The impact of entrepreneurial capacity, experience and organizational support on academic entrepreneurship", *Research Policy*, 40(8), 1084-1093.

Classen, N., Van Gils, A., Bammens, Y., Carree, M. (2012) "Accessing Resources from Innovation Partners: The Search Breadth of Family SMEs", *Journal of Small Business Management*, 50(2), 191-215.

Cliff, J.E., Devereaux Jennings, P., Greenwood, R. (2006) "New to the game and questioning the rules: The experiences and beliefs of founders who start imitative versus innovative firms", *Journal of Business Venturing*, 21, 633-663.

Coad, A., Segarra, A., Teruel, M. (2013) "Like milk or wine: does firm performance improve with age?" *Structural Change and Economic Dynamics*, 24, 173-189.

Coad, A., Segarra, A., Teruel, M. (2016) "Innovation and firm growth: Does firm age play a role?" *Research Policy*, 45(2), 387-400.

Coff, R.W. (2002) "Human capital, shared expertise, and the likelihood of impasse in corporate acquisitions", *Journal of Management*, 28, 107-128.

Cohen, W.M., Levinthal, D.A. (1989) "Innovation and Learning: The Two Faces of R&D", *Economic Journal*, 99: 569-596.

Cohen, W.M., Levinthal, D.A. (1990) "Absorptive Capacity: A New Perspective on Learning and Innovation," *Administrative Science Quarterly*, 35(1), 128-152.

Colombo, M.G., Delmastro, M., Grilli, L. (2004) "Entrepreneurs' human capital and the start-up size of new technology-based firms", *International Journal of Industrial Organization*, 22(8), 1183-1211.

Colombo, M.G., Grilli, L. (2005) "Founders' human capital and the growth of new technology-based firms: a competence-based view", *Research Policy*, 34, 795-816.

Colombo, M.G., Mustar, P., Wright, M. (2010) "Dynamics of Science-Based Entrepreneurship", *Journal of Technology Transfer*, 35(1), p. 1-15.

Conner, K.R., Prahalad, C.K. (1996) "A Resource-Based Theory of the Firm: Knowledge versus Opportunism", *Organization Science*, 7(5), 477-501.

Cook, D.O., Kieschnick, R., McCullough, B.D. (2008) "Regression analysis of proportions in finance with self-selection", *Journal of Empirical Finance*, 15, 860-867.

Cook, R.D., Weisberg, S. (1997) "Graphics for assessing the adequacy of regression models", *Journal of the American Statistical Association*, 92(438), 490-499.

Cox, D.R. (1972) "Regression models and life tables (with discussion)", *Journal of the Royal Statistical Society*, Series B, 34, 187-220.

Cox. D.R., Oakes, D. (1984) *Analysis of Survival Data*, Chapman and Hall: London.

Cressy, R., Olofsson, C. (1997) "The financial conditions for Swedish SMEs: Survey and research agenda", *Small Business Economics*, 9, 179-192.

Criaco, G., Minola, T., Migliorini, P. Serarols-Tarrés, C. (2013) "'To have and have not': founders" human capital and university start-up survival", *Journal of Technology Transfer*, 39(4), 567-593.

Criscuolo, P., Nicolaou, N., Salter, A. (2012) "The elixir (or burden) of youth? Exploring differences in innovation between start-ups and established firms", *Research Policy*, 41, 319-333.

Cronbach, L., Meehl, P.E. (1955) "Construct Validity in Psychological Tests", *Psychological Bulletin*, 52, 281-302.

Cyert, R.M., March, J.G. (1963) A Behavioral Theory of the Firm, Prentice Hall: Englewood Cliffs, N.J. Czarnitzki, D., Delanote, J. (2012) "Young Innovative Companies: the new high-growth firms?" *Industrial and Corporate Change*, 22(5), p. 1315-1340.

Dahlander, L., Gann, D. (2010) "How open is innovation?" *Research Policy*, 39(6), 699-709.

Dahlin, K.B., Weingart, L.R., and Hinds, P.J. (2005) "Team Diversity and Information Use", *The Academy of Management Journal*, 48(6), 1107-1123.

Das, T.K., Teng, B.S. (2000) "A Resource-Based Theory of Strategic Alliances", *Journal of Management*, 26(1), 31-61.

David, P.A., and Foray, D. (2003) "Economic fundamentals of the knowledge society", *Policy futures in education*, An e-Journal 1(1) special issue, January.

de Jong, J.P., Marsili, O. (2006) "The fruit flies of innovations: A taxonomy of innovative small firms," *Research Policy*, 35(2), 213-229.

Delmar, F., Shane, S. (2004) "Legitimating first: organizing activities and the survival of new ventures", *Journal of Business Venturing*, 19, 385-410.

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

Delmar, F., Wennberg, K. (2010) *Knowledge Intensive Entrepreneurship: The Birth, Growth and Demise of Entrepreneurial Firms*, Edward Elgar Publishing: Cheltenham, UK.

Demsetz, H. (1991) "The theory of the firm revisited", in *The Nature of the Firm*, 159-178, Oxford University Press: New York.

den Hertog, P., van der Aa, W., de Jong, M.W. (2010) "Capabilities for managing service innovation: towards a conceptual framework", *Journal of Service Management*, 21, 490-514.

Dencker, J.C., Gruber, M., Shah, S.K. (2009) "Pre-Entry Knowledge, Learning, and the Survival of New Firms", *Organization Science*, 20(3), 561-537.

Dencker, J.C., Gruber, M. (2014) "The Effects of Opportunities and Founder Experience on New Firm Performance", *Strategic Management Journal*, available online at:
http://onlinelibrary.wiley.com/doi/10.1002/smj.2269/abstract.

Di Gregorio, D., Shane, S. (2003) "Why do some universities generate more start-ups than others?" *Research Policy*, 32(2), 209-227.

Dodgson, M, Gann, D.M., Phillips, N. (eds.) (2014) *The Oxford Handbook of Innovation Management*, Oxford University Press: Oxford.

Dollinger, M.J. (1984) "Environmental Boundary Spanning and Information Processing Effects on Organizational Performance", *The Academy of Management Journal*, 27(2), 351-368.

Drucker, P. (1994) *Post-Capitalist Society*, Harper Business: New York.

Dunning, J.H., Lundan, S. (2008) *Multinational Enterprises and the Global Economy*, 2nd Edition, Edward Elgar: Cheltenham, UK.

Dunteman, G. (1989) *Principal Components Analysis*, Series: Quantitative Applications in the Social Sciences, 69. SAGE Publications: Newbury Park, London.

Economist, The (2015) "Universities: The world is going to university", Vol. 411, Print Edition, March 28th, 2015.

Edquist, C. (2006) "Systems of Innovation: Perspectives and Challenges", in Fagerberg, J., Mowery, D.C, Nelson, R.R. (eds.) *The Oxford Handbook of Innovation*, Oxford University Press: New York.

Edquist, C., and McKelvey, M. (2000) (Eds.). *Systems of Innovation: Growth, Competitiveness and Employment*, Edward Elgar Publishing: Cheltenham, UK.

Eisenhardt, K.M., Schoonhoven, C.B. (1990) "Organizational growth: Linking founding team, strategy, environment, and growth among US semiconductor ventures, 1978-1988", *Administrative science quarterly*, 35(3), 504-529

Eisenhardt, K.M., Martin, J.A. (2000) "Dynamic Capabilities: what are they?" *Strategic Management Journal*, 21(10-11), 1105-1121.

Eliasson, G. (1996) "Spillovers, integrated production and the theory of the firm", *Journal of Evolutionary Economics*, 6(2), 125-140.

EC-DG ENTR (2009) *European competitiveness report 2008*, Luxemburg.

EMCC (2005) "Sector Futures: The knowledge intensive business service sector,"http://www.eurofound.europa.eu/emcc/publications /2005/ef0559en.pdf.

European Commission (2012) *Knowledge-intensive (business) services in Europe*, prepared by Schricke, E., Zenker, A., Stahlecker, T., EC: Brussels.

European Commission (2013) *Research and Innovation performance in EU Member States and Associated countries: Innovation Union progress at country level*, Brussels, Belgium: RTD Publications.

Eurostat (2014) Eurostat indicators of high-tech industry and knowledge-intensive services. http://ec.europa.eu/eurostat/cache/metadata/Annexes/htec_esms_an3.pdf. Retrieved 2014-04-14.

Evans, D.S, Jovanovic, B. (1989) "An estimated model of entrepreneurial choice under liquidity constraints", *Journal of Political Economy*, 97, 806-827.

Fagerberg, J. (2006)" Innovation: A guide to the literature", in Fagerberg, J., Mowery, D.C, Nelson, R.R. (eds.) *The Oxford Handbook of Innovation*, Oxford University Press: New York.

Fann, G.L., Schmeltzer, L.R. (1989) "The Use of Information from and about Competitors in Small Business Management", *Entrepreneurship Theory and Practice*, 13(4), 35-46.

Feldman, M., Choi, J. (2015) "Harnessing the Geography of Innovation: Toward Evidence Based Economic Development Policy", in Archibugi, D., Fillippetti, A. (eds.) *The Handbook of Global Science, Technology, and Innovation*, Wiley-Blackwell: New York.

Fleming L. (2001) "Recombinant uncertainty in technological search", *Management Science* 47(1), 117–132.

Fleming, L., Sorenson, O. (2001) "Technology as a complex adaptive system: Evidence from patent data", *Research Policy*, 30(7), 1019-1039.

Fleming, L., Sorenson, O. (2004) "Science as a map in technological search", *Strategic Management Journal*, 25(8-9), 909-928.

Florin, J., Lubatkin, M., Schulze, W. (2003) "A social capital model of high growth ventures", *Academy of Management Journal*, 46(3), 374-384.

Fontana, R., Malerba, F., Marinoni, A. (2015) "Knowledge intensive entrepreneurship in different sectoral systems: a taxonomy", in Malerba, F., Caloghirou, Y., McKelvey, M., Radosevic, S. (eds.) (2015). *Dynamics of Knowledge-Intensive Entrepreneurship: Business Strategy and Public Policy.* Routledge: UK.

Forbes Magazine (2013) "The Secret to Entrepreneurial Success: Forget College", online, retrieved 2016-10-10.

Forbes Magazine (2015) "Education vs. Entrepreneurship: Which Path Wins?", online, retrieved 2016-10-10.

Forsman, H. (2009) "Improving innovation capabilities of small enterprises. A cluster strategy as a tool", *International Journal of Innovation Management*, 13(2), 1-23.

Forsman, H. (2011) "Innovation capacity and innovation development in small enterprises. A comparison between the manufacturing and service sectors", *Research Policy*, 40, 739-750.

Forsman, H., Ratanen, H. (2011) “Small manufacturing and service enterprises as innovators: a comparison by size”, *European Journal of Innovation Management*, 14(1), 27-50.

Foss, N.J. (1996a)"Knowledge-Based Approaches to the Theory of the Firm: Some Critical Comments", *Organization Science* 7(5), 470-476.

Foss, N.J. (1996b) "More Critical Comments on Knowledge-Based Theories of the Firm", *Organization Science* 7(5), 519-523.

Fox, J. (2003) "Effects Displays in R for Generalised Linear Models", *Journal of Statistical Software*, 8(15), 1-27.

Fox. J., Andersen, R. (2006) “Effect displays for multinomial and proportional-odds logit models”, *Sociological Methodology*, 36(1), 225-255.

Fox, J., Weisberg, S. (2010) "Time-Series Regression and Generalized Least Squares in R: An appendix to An R Companion to Applied Regression, 2nd Edition".
https://socserv.socsci.mcmaster.ca/jfox/Books/Companion/appendix/Appendix-Timeseries-Regression.pdf.

Fox, J., Weisberg, S (2011) *An R Companion to Applied Regression 2nd edition*, SAGE Publications: Thousand Oaks, CA.

Fox, J., Weisberg, S. (2015) "Using car functions in other functions", available at:
https://cran.r-project.org/web/packages/car/vignettes/embedding.pdf

Fox. J. (2016) *Applied Regression Analysis and Generalized Linear Models, 3rd edition*, SAGE Publications: Thousand Oaks, CA.

Franco, M., Haase, H. (2010) "Failure factors in small and medium-sized enterprises: qualitative study from an attributional perspective", *International Entrepreneurship and Management Journal*, 6, 503-521.

Freel, M. (2006) "Patterns of technological innovation in knowledge-intensive business services", *Industry and Innovation*, 13(3): 335-358.

Freeman, C., Soete, L. (1997) *The Economics of Industrial Innovation*, 3rd edition, Pinter: London.

Gabriel, K.R. (1971) "The biplot graphic display of matrices with application to principal component analysis", *Biometrika*, 58(3): 453-467.

Gassmann, O., Enkel, E., Chesbrough, H. (2010) "The future of open innovation", *R&D Management*, 40(3), 213-221.

Gallego, J., Rubalcaba, L., Hipp, C. (2013) "Organizational Innovation in Small European Firms: A multidimensional approach", *International Small Business Journal*, 31(5): 563-579.

Gimeno, J., Folta, T.B., Cooper, C., Woo, Y. (1997) "Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms", *Administrative Science Quarterly*, 42(4), 750-783.

Gimmon, E., Levie, J. (2010) "Founder's human capital, external investment, and the survival of new high-technology ventures", *Research Policy*, 39(9), 1214-1226.

Glaister, K.W., Buckley, P.J. (1996) "Strategic Motives For International Alliance Formation", *Journal of Management Studies*, 33, 301-332.

Gort, M., Klepper, S. (1982) "Time Paths in the Diffusion of Product Innovations", *Economic Journal*, 92(3), 630-655.

Granstrand, O (2000) "The Shift Towards Intellectual Capitalism - the Role of Infocom Technologies," *Research Policy*, 29(9), 1061-1080.

Graham, J.W, Olchowski, A.E, Gilreath, T.D. (2007) "How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory", *Prevention Science*, 8(3),206-213.

Graham, J.W. (2009) "Missing data analysis: Making it work in the real world," *Annual review of psychology*, 60, 549-576.

Grant, R. (1996) "Toward a knowledge-based theory of the firm," *Strategic Management Journal*, 17, 109-122.

Grant, R. (2002) "The Knowledge-Based View of the Firm", in Wei Choo, C., Bontis, N (eds.) *The Strategic Management and Intellectual Captial and Organizational Knowledge* (2002) Oxford University Press: Oxford

Grueber, M., Studt, T. (2011) *2012 Global R&D Funding Forecast*, Advantage Business Media and R&D Magazine.

Gunday, G., Ulusoy, G., Kilic, K., Alpkan, L., (2011) “Effects of innovation types on firm performance”, *International Journal of Production Economics*, 133(2), 662-676.

Guttman, L. (1954) "Some necessary conditions for common factor analysis", *Psychometrika*, 30, 179-185.

Götzfried, A. (2004) "European employment increasing in services and especially in knowledge-intensive services", *Statistics in Focus: Science and Technology*, 10, Eurostat, Catalogue number KS-NS-04-010-EN-N.

Haati, A., Madupu, V., Yavas, V., Babakus, E. (2005) "Cooperative strategy, knowledge intensity and export performance of small and medium sized enterprises", *Journal of World Business*, 40, 124-138.

Hagedoorn, J. (1993) "Understanding the rationale of strategic technology partnering: Interorganizational modes of cooperation and sectoral differences", *Strategic Management Journal*, 14(5), 371-385.

Hagedoorn, J., Cloodt, M. (2003) "Measuring innovative performance: is there an advantage in using multiple indicators?" *Research Policy*, 32(8), 1365-1379.

Hambrick, D.C., Mason, P.A. (1984) "Upper echelons: The organization as a reflection of its top managers", *Academy of Management Review*, 9(2), 193-206.

Hatzichronoglou, T. (1997) "Revision of the High-Technology Sector and Product Classification," *OECD Science, Technology and Industry Working Papers*, 1997/02, OECD Publishing.

Hauknes J. (1999) "Knowledge intensive services – what is their role?", Paper presented at the OECD Forum on Realising the Potential of the Service Economy. Paris: 28 September 1999.

Hayek, F.A. (1945) "The use of knowledge in society", *American Economic Review*, 35, 519-532

Headd, B., Kirchhoff, B. (2009) "The growth, decline and survival of small business: an exploratory study of life cycles", *Journal of Small Business Management*, 47(4), 531-550.

Heckman, J. (1979) "Sample selection bias as a specification error", *Ekonometrika*, 47, 153-161.

Helfat, C.E. (1994) ”Evolutionary trajectories in petroleum firm R&D”, *Management Science*, 40, 1730-1747.

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

Henrekson, M., Johansson, D. (2010) "Gazelles as job creators: a survey and interpretation of the evidence", *Small Business Economics*, 35, 227-244.

Hipp, C., Grupp, H. (2005) "Innovation in the service sector: The demand for service-specific innovation measurement concepts and typologies", *Research Policy*, 34, 517-535.

Hirsch-Kreinsen, H., Hahn, K., Jacobson, D., (2008) "The low-tech issue", in Hirsch-Kreinsen, H. and Jacobson, D., (eds.) *Innovation in Low-tech Firms and Industries* (2008), 3-24. Edward Elgar: Cheltenham, UK; Northampton, MA.

Hirsch-Kreinsen, H., Jacobson, D. (eds.) (2008) *Innovation in Low-Tech Firms and Industries*, Edward Elgar: Cheltenham, UK; Northampton, MA, USA.

Hirsch-Kreinsen, H., Schwinge, I. (2015) "Knowledge intensive entrepreneurship in low-technology industries", in Malerba, F., Caloghirou, Y., McKelvey, M., Radosevic, S. (eds.) (2015) *Dynamics of Knowledge-Intensive Entrepreneurship: Business Strategy and Public Policy*. Routledge: London.

Hitt, M. A., Tyler, B.B. (1991) "Strategic Decision Models: Integrating Different Perspectives", *Strategic Management Journal*, 12(4), 327–351.

Hitt, M.A., Bierman, L., Shimizu, K., Kochhar, R. (2001) "Direct and Moderating Effects of Human Capital on Strategy and Performance in Professional Service Firms: A Resource-Based Perspective", *The Academy of Management Journal*, 44(1): 13-28.

Hoffman, K., Parejo, M., Bessant, J., Perren, L. (1998) "Small firms, R&D, technology and innovation in the UK: a literature review", *Technovation*, 18(1), 39-55.

Holmén, M., Magnusson, M., McKelvey, M. (2007) "What are innovative opportunities?" *Industry and Innovation*, 14(1), 27-45.

Honaker, J., King, G. (2010) "What to do about missing values in time series cross-section data", *American Journal Political Science Review*, 95(1), 49-69.

Honaker, J., King, G., Blackwell, M. (2011) "Amelia II: A Program for Missing Data", *Journal of Statistical Software*, 45(7), 1-47.

Højsgaard, S., Halekoh, U., Yan, J. (2006) "The R Package geepack for Generalized Estimating Equations", *Journal of Statistical Software*, 15(2), 1-11.

Hubbard, A.E., Ahern, J., Fleischer, N.L., Van der Laan, M., Lippman, S.A., Jewell, N., Bruckner, T., Satariano, W.A. (2010) "To GEE or Not to GEE", *Epidemiology*, 21(4), 467-474.

Huber, G.P. (1991) "Organizational Learning: the contributing processes and the literatures". *Organization Science*, 2, 88-115.

Hughes, A. (2001) "Innovation and business performance: small entrepreneurial firms in the UK and the EU", *New Economy*, 8(3), 157-163.

Imai, K., King, G., Lau, O. (2007) "Zelig: Everyone's Statistical Software", http://gking.harvard.edu/zelig.

Imai, K., King, G., Lau, O. (2008) "Towards A Common Framework for Statistical Analysis and Development", *Journal of Computational and Graphical Statistics*, 17(4), 892-913.

Jacoby, W.G. (1991) *Data Theory and Dimensional Analysis*, Series: Quantitative Applications in the Social Sciences, 78. SAGE Publications: Newbury Park, London.

Jacoby, W.G. (1999) "Levels of measurement and political research: An optimistic view", The American *Journal of Political Science*, 43(1), 271-301.

Jacoby, W.G. (f/c). "opscale: A function for Optimal Scaling", *The R Journal*, vol x/y, available at: http://polisci.msu.edu/jacoby/icpsr/scaling /computing/alsos/Jacoby,%20opscale%20MS.pdf.

Jantunen, A. (2005) "Knowledge-processing capabilities and innovative performance: an empirical study", *European Journal of Innovation Management*, 8(3), 336-349.

Jöreskog, K.G. (1973) "A general method for estimating a linear structural equation system", in Goldberger, A.S., Duncan, O.D. (eds.), *Structural Equation Models in the Social Sciences*, Academic Press: New York, 85-112.

Kaiser, H.F. (1960) "The application of electronic computers to factor analysis", Educational and *Psychological Measurement*, 20, 141-151.

Kaplan, E.L., Meier, P. (1958) "Nonparametric estimation from incomplete observations", *Journal of the American Statistical Association*, 53(282), 457-481.

Katila, R., Ahuja, G. (2002) "Something old, something new: A longitudinal study of search behavior and new product introduction," *Academy of Management Journal*, 45(6), 1183-1194.

Kerlinger, F.N. (1986) *Foundations of Behavioral Research*, Holt, Rinehart & Winston: New York.

Keupp, M.M., Gassmann, O. (2013) "Resource constraints as triggers of radical innovation: Longitudinal evidence from the manufacturing sector," *Research Policy*, 42(8), 1457-1468.

Keupp, M.M., Gassmann, O. (2007) "The competitive advantage of early and rapidly internationalising SMEs in the biotechnology industry; a knowledge based view", *Journal of World Business*, Special Issue: The Early and Rapid Internationalisation of the Firm, 42(3), 350-366.

Kim, D-J., Kogut, B. (1996) "Technological Platforms and Diversification", *Organization Science*, 7(3), 283-301.

Kirzner, I., (1997) "Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach", *Journal of Economic Literature*, 35, 60-85.

Klepper, S. (1996) "Exit, Growth, and Innovation over the Product Life Cycle", *The American Economic Review*, 86(3), 562-583.

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

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

Klepper, S (2016) Experimental Capitalism: The Nanoeconomics of American High-Tech Industries, Braguinsky, S., Hounshell, D, Miller, J. (eds.), Princeton University Press: USA.

Klevorick, A. K., Levin, R., Nelson, R., Winter, S. (1995) “On the Sources and Significance of Inter-Industry Differences in Technological Opportunities”, *Research Policy*, 24, 185-205.

Knight, F. (1921) *Risk, Uncertainty, and Profit*, Hart, Schaffner and Marx: New York.

Kogut, B., Zander, U. (1992) "Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology", *Organization Science*, 3(3), 383-397.

Kor, Y.Y., Leblebici, H. (2005) "How do interdependencies among human-capital deployment, development and diversification strategies affect firms' financial performance?" *Strategic Management Journal*, 26: 967-985.

Kotey, B., Sheridan, A. (2004) "Changing HRM practices with firm growth", *Journal of Small Business and Enterprise Development*, 11, 474-485.

Kindström, D., Kowalkowski, C., Sandberg, E. (2013) "Enabling service innovation: A dynamic capabilities approach", *Journal of Business Research*, 66, 1063-1073.

Kraaijenbrink, J., Spender, J.-C., Groen, A.J. (2010) "The Resource-Based View: A Review and Assessment of Its Critiques", *Journal of Management*, 36(1), 349-372.

Kristinsson, K., Candi, M., Sæmundsson, R. (2016) "The Relationship between Founder Team Diversity and Innovation Performance: The Moderating Role of Causation Logic", *Long Range Planning*, 49, 464-476.

Kuznets, S (1954) *Economic Change*, Heinermann: London.

Lam, P. (2007) "normal.gee: Generalized Estimating Equation for Normal Regression", in Imai, K., King, G., Lau, O. (2007) "Zelig: Everyone's Statistical Software", http://gking.harvard.edu/zelig

Lasagni, A. (2013) "How can external relationships enhance innovation in SMEs? New evidence from Europe", *Journal of Small Business Management*, 50(2), 310-339.

Lassen, A.H., McKelvey, Maureen (2012) AEGIS D7.2.1 Conceptualizing knowledge-intensive entrepreneurship: A Literature Review for analyzing case studies and defining policy implications. 1-83.

Laursen, K., Foss, N. (2003) "New human resource management practices, complementarities and the impact on innovation performance", *Cambridge Journal of Economics*, 27(2), 243-263.

Laursen, K. (2012) "Keep searching and you'll find: what do we know about variety creation through firms' search activities for innovation?" *Industrial and Corporate Change*, 21(5), 1181-1220.

Laursen, K., Salter, A (2004) "Searching high and low: what types of firms use universities as a source of innovation?" *Research Policy*, 33, 1201-1215.

Laursen, K., Salter, A. (2006) "Open for innovation: the role of openness in explaining innovation performance among U.K. manufacturing firms," *Strategic Management Journal*, 27(2), 131-150.

Laursen, K., Salter, A.J. (2014) “The paradox of openness: Appropriability, external search and collaboration”, *Research Policy*, 43(5), 867-878.

Lee, S., Park, G., Yoon, B., Park, J. (2010b) "Open innovation in SMEs—an intermediated network model," *Research Policy*, 39(2), 290-300.

Leiponen, A. (2005) "Skills and Innovation", *International Journal of Industrial Organization*, 23, 303-323.

Leiponen, A., Helfat, C.E. (2011) "Location, decentralization, and knowledge sources for innovation”, *Organization Science*, 22, 641-658.

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G. (1987) “Appropriating the Returns from Industrial-Research and Development”, *Brooking papers on economic activity*, 783-831.

Leonard, W. (1971) "Research and Development in Industrial Growth", *Journal for Political Economy*, 79, 232-256.

Liang, K-Y., Zeger, S.L. (1986) "Longitudinal data analysis using generalized linear models", *Biometrika*, 73(1), 13-22.

Liebeskind, J.P. (1996) "Knowledge, Strategy, and the Theory of the Firm", *Strategic Management Journal*, 17, 93-107.

Little, J.R., Rubin, D. (2002) *Statistical analysis with missing data*, Wiley: New York.

Loasby, B.J. (1999) *Knowledge, Institutions and Evolution in Economics: The Graz Schumpeter Lectures v.2*, Routledge: London.

Loasby, B.J. (2000) "How do we know?" in Earl, S.F., Frowen, S.F. (eds.) *Economics as an Art of Thought*, Routledge: London.

Long, J.S. (1997) *Regression Models for Categorical and Limited Dependent Variables. A volume in the Sage Series for Advanced Quantitative Techniques*, Sage Publications: Newbury Park, CA.

Love, J.H., Roper, S., Vahter, P. (2014) “Learning from openness: The dynamics of breadth in external innovation linkages”, *Strategic Management Journal*, 35(11), 1703-1716.

Lumpkin, G.T., Dess, G.G. (1996) "Clarifying the entrepreneurial orientation construct and linking it to performance," *Academy of Management Review*, 21(1), 135-172.

Lundvall, B.-Å., Borrás, S. (1997) *The Globalising Learning Economy*, European Communities: Luxembourg.

Machlup, F. (1967) "Theories of the firm: Marginalist, behavioral, managerial," *American Economic Review*, 57, 201-220.

Makadok, R. (2001) "Toward a Synthesis of the Resource-Based and Dynamic-Capability Views of Rent Creation", *Strategic Management Journal*, 22(5), 387-401.

Malerba, F (2010) "Knowledge-intensive entrepreneurship and innovation systems in Europe", in Malerba, F. (ed.) *Knowledge-Intensive Entrepreneurship and Innovation Systems: Evidence from Europe*, Abington, Routledge: New York.

Malerba, F., McKelvey, M (2010) AEGIS D1.1.1. Conceptualizing knowledge-intensive entrepreneurship: Concepts and models, 1-46.

Malerba, F., McKelvey, M. (2015) "Knowledge Intensive Entrepreneurship: Conceptual Framework", in Malerba, F., Caloghirou, Y., McKelvey, M., Radosevic, S. (eds.) (2015). *Dynamics of Knowledge-Intensive Entrepreneurship: Business Strategy and Public Policy*. Routledge: London.

Malerba, F., Caloghirou, Y., McKelvey, M., Radosevic, S. (eds.) (2015) *Dynamics of Knowledge-Intensive Entrepreneurship: Business Strategy and Public Policy*. Routledge: London.

Mamede, R., Mota, D., Godinho, M.M.(2010 "Are the dynamics of knowledge-based industries any different?" in Malerba, F. (ed.) *Knowledge-Intensive Entrepreneurship and Innovation Systems: Evidence from Europe*, Abington, Routledge: New York.

Mansfield, E. (1972) "Contribution of R&D to economic growth in the United States", *Science* 175 (4021), 477-486.

March, J. (1991) "Exploration and exploitation in organizational learning," *Organization Science*, 2(1), 71-87.

March, J., Simon, H. (1958) *Organizations*, Wiley: New York.

Marshall, A. (1949) *Principles of economics*: 8th edition. Macmillan: London.

Marvel, M.R, Lumpkin, G.T. (2007) "Technology Entrepreneurs' Human Capital and its Effects on Innovation Radicalness", *Entrepreneurship Theory and Practice*, 31(6), 807-828.

Mason, E.S. (1939) "Price and Production Policies of Large-Scale Enterprises", *The American Economic Review*, 29(1), 61-74.

McKelvey, M. (1996) *Evolutionary Innovations*, Oxford University Press: Oxford and New York.

McKelvey, M., Lassen, A. (2013) *Managing Knowledge Intensive Entrepreneurship*, Edward Elgar: Cheltenham, UK.

McKelvey, M., Rake, B. (2015) "Product innovation success based on cancer research in the pharmaceutical industry: co-publication networks and the effects of partners", *Industry and Innovation*, 23(5), 383-406.

Menger, Carl (1871/1975/2002) *Principles of Economics*, Ludwig von Mises Institute: Auburn, AL, USA.

Metcalfe, J.S. (1998) *Evolutionary Economics and Creative Destruction*, Routledge: London.

Metcalfe, J.S. (2002) "Knowledge of growth and the growth of knowledge", *Journal of Evolutionary Economics*, 12: 3-15.

Metcalfe, J.S., Ramlogan, R. (2005) "Limits to the economy of knowledge and knowledge of the economy", *Futures*, 37: 655-674.

Metcalfe, J.S., Foster, J., Ramlogan, R. (2006) "Adaptive economic growth", *Cambridge Journal of Economics*, 30, 7-32.

Miles, I. (2005) "Innovation in Services", in Fagerberg, J., Mowery, D.C., Nelson, R. R. (eds.) *The Oxford Handbook of Innovation* (2005), Oxford University Press: Oxford; New York.

Miles, I. (2012) "Introduction to Service Innovation", in: Macaulay, L., Wilby, J., Tan, Y., Zhao, L., Theodoulidis, B. (eds.), *Case Studies In Service Innovation.* Springer Science + Business Media, New York, 1-18.

Millan, J.M., Congregado, E., Román, C., van Praag, M., van Stel, A. (2014) "The value of an educated population for an individual's entrepreneurship success" *Journal of Business Venturing*, 29, 612-632.

Mises, L.V. (1949) *Human Action*, Ludvig von Mises Institute: Auburn, AL, USA.

Muller, E., Zenker, A. (2001) "Business services as actors of knowledge transformation: the role of KIBS in regional and national innovation systems", *Research Policy*, 30, 1501-1516.

Muller, E., Doloreux, D. (2009) "What we should know about knowledge-intensive business services", *Technology in Society*, 31, 64-72.

Mytelka, L.K., Smith, K. (2002) "Policy learning and innovation theory: an interactive and co-evolving process", *Research Policy*, 31, 1467-1479.

Mäkinen, H. (2002) "Intra-firm and inter-firm learning in the context of start-up companies", *International Journal of Entrepreneurship and Innovation*, 3(1), 5-34.

Nelson, R.R., Phelps, E.S. (1966) “Investments in Humans, Technological Diffusion, and Economic Growth”, *The American Economic Review*, 56(1/2), 69-75.

Nelson, R.R., Peck, M.J., Kalacheck, J. (1967) *Technology, economic growth, and public policy*; a Rand Corporation and Brookings Institution Study.

Nelson, R.R., Winter, S.G., Schuette, H.L. (1976) “Technical change in an evolutionary model”, *The Quarterly Journal of Economics*, 90-118.

Nelson, R.R., Winter, S. (1977) "In search of useful theory of innovation", *Research Policy*, 6(1), 36-76.

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

Nelson, R.R. (1991) "Why do firms differ, and how does it matter?” *Strategic Management Journal*, 12, 61-74.

Nelson, R.R. (1993) *National innovation systems: a comparative analysis*, Oxford University Press: New York, Oxford.

Nickerson, J.A., Zenger, T.R. (2004) "A Knowledge-Based Theory of the Firm - The Problem-Solving Perspective", *Organization Science*, 15(6), 617-632.

Nieto, M., Santamaría, L. (2010) “Technological Collaboration: Bridging the Innovation Gap between Small and Large Firms”, *Journal of Small Business Management*, 48(1), 44-69.

Nonaka, I., Takeuchi, H. (1995) *The Knowledge-Creating Company - How Japanese Companies Create the Dynamics of Innovation*, Oxford University Press: Oxford; New York.

Oakey, R. (1991) "High technology small firms: their potential for rapid industrial growth", *International Small Business Journal*, 9, 30-42.

OECD (1994) *The Measurement and Scientific and Technical Activities: Proposed Standard Practice for Surveys of Research and Experimental Development* (Frascati Manual: 1993), OECD: Paris.

OECD (1996) *The Knowledge Based Economy*, OECD/GD(96)102, OECD:Paris.

OECD (2002) *OECD Frascati Manual: Proposed Standard for Surveys on Research and Experimental Development*, 6th revision, OECD: Paris.

OECD (2003) *Science, Technology, and Industry Scoreboard 2003 - Towards a Knowledge-Based Economy*, OECD: Paris.

OECD (2005) *Oslo-Manual: Proposed Guidelines for Collecting and Interpreting Technological Innovation Data, 3rd edition*, OECD: Paris.

OECD (2008) *Fostering entrepreneurship for innovation*, OECD Science, Technology and Industry Working Papers 2008/5, OECD: Paris.

OECE (2012) *Policy Report on Service R&D and Innovation*, Working Party of National Experts on Science and Technology Indicators and Working Party on Innovation and Technology Policy, OECD: Paris.

OECD (2013) *Science, Technology and Industry Scoreboard: Innovation for Growth.* OECD: Paris.

OECD (2015) *The Future of Productivity*, OECD: Paris.

O'Regan, N., Sims; M. (2008) "Identifying high technology firms: A sectoral analysis", *Technovation*, 28, 408-423.

Pan. W. (2001) "Akaike's information criterion in generalized estimating equations", *Biometrics*, 57, 120-125.

Papke, L.E., Wooldridge, J.M. (1996) "Econometric methods for fractional response variables with an application to 401(k) plan participation rates", *Journal of Applied Econometrics*, 11, 619-632.

Parker, S.C. (2004) *The economics of self-employment and entrepreneurship*, Cambridge University Press: Cambridge.

Pavitt, K. (1984) "Sectoral patterns of technical change: towards a taxonomy and a theory" *Research Policy*, 13(6), 343-373.

Penrose, E (1959) *The Theory of the Growth of the Firm*, Blackwell: Oxford, UK.

Penrose, E (1985) "The Theory of the Growth of the Firm Twenty-Five Years Later", *Acta Universitatis Upsaliensis: Studia Oeconomicae Negotiorum* (Uppsala Lectures in Business), 20, 1-16.

Perkmann, M., Walsh, K. (2007) "University -industry relationships and open innovation: Towards a research agenda", *International Journal of Management Review*, 9(4), 259-280.

Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D'Este, P., Fini, R., Geuna, A., Grimaldi, R., Hughes, A., Krabel, S., Kitson, M., Llerena, P., Lissoni, F., Salter, A., Sobrero, M. (2013) "Academic engagement and commercialisation: A review of the literature on university-industry relations", *Research Policy*, 42(2), 423-442.

Pfeffer, J. (1994) *Competitive Advantage Through People*, Harvard Business School Press: Boston.

Phelps, E.S. (1966) "Models of Technical Progress and the Golden Rule of Research", *The Review of Economic Studies*, 33(2), 133-145.

Pitelis, C.N. (2007) "A Behavioral Resource-Based View of the Firm: The Synergy of Cyert and March (1963) and Penrose (1959)", *Organization Science*, 18(3), 478-490.

Pitelis, C.N (2009) "Edith Penrose's 'The Theory of the Growth of the Firm' Fifty Years Later", in Penrose, E. *The Theory of the Growth of the Firm, 4th Edition*, Oxford University Press: Oxford.

Ployhart, R.E., Moliterno, T.P. (2011) "Emergence of the human capital resource: A multilevel model", *Academy of Management Review*, 36(1): 127-150.

Polanyi, M. (1962) *Personal Knowledge: Towards a Post-critical Philosophy.* University of Chicago Press: Chicago.

Polanyi, M. (1966) *The Tacit Dimension*, Doubleday: New York.

Popper, K. (1959/2009) *The Logic of Scientific Discovery*, Routledge: New York.

Porter, M.E. (1980) *Competitive strategy: Techniques for analyzing industries and competitors*, Free Press: New York.

Porter, M.E. (1985) *Competitive advantage*, Free Press: New York.

Powell, W., Grodal, S. (2005) "Networks of Innovators", in Fagerberg, J., Mowery, D.C., Nelson, R. R. (eds.) *The Oxford Handbook of Innovation* (2005), Oxford University Press: Oxford; New York.

Protogerou, A., Caloghirou, Y. (2015) "Dynamic capabilities in young knowledge-intensive firms: an empirical approach", in Malerba, F., Caloghirou, Y., McKelvey, M., Radosevic, S. (eds.), *Dynamics of Knowledge-Intensive Entrepreneurship: Business Strategy and Public Policy.* Routledge: London.

Reinl, L., Kelliher, F. (2010) "Cooperative micro-firm strategies: Leveraging resources through learning networks", *Entrepreneurship and Innovation*, 11(2), 141-150.

Rickne, A., Jacobsson, S. (1999) "New Technology-Based Firms In Sweden - A Study of Their Direct Impact On Industrial Renewal", *Economics of Innovation and New Technology*, 8(3), 197-223.

Romer, P.M. (1986) "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 94(5), 1002-1037.

Romer, P.M. (1990) “Endogenous Technical Change”, *The Journal of Political Economy*, 98(5): S71-S10.

Rosenberg, N. (1974) “Science, Invention and Economic Growth”, *The Economic Journal*, 84(333), 90-108.

Rosenberg, N. (1976) *Perspectives on Technology*, Cambridge University Press: Cambridge, UK.

Rosenberg, N. (1982) *Inside the Black Box: Technology and Economics*, Cambridge University Press: Cambridge, UK

Rosenbusch, N., Brinkmann, J., Bausch, J. (2011) “Is innovation always beneficial? A meta-analysis of the relationship between innovation and performance in SMEs”, *Journal of Business Venturing*, 26(4), 441-457.

Rothwell, R. (1989) "Small firms, innovation and industrial change", *Small Business Economics*, 1, 51-64.

Rothwell, R., Zegveld, W. (1982) *Innovation and the small and medium sized firm*, University of Illinois at Urbana. Pinter: London, UK.

Rubin, D.B. (1976) "Inference and missing data", *Biometrika*, 63(3), 581-592.

Salter, A., McKelvey, M (2016) "Evolutionary analysis of innovation and entrepreneurship: Sidney G. Winter - recipient of the 2015 Global Award for Entrepreneurship Research", *Small Business Economics*, 47(1), 1-14.

Santarelli, E., Vivarelli, M. (2007) "Entrepreneurship and the process of firms’ entry, survival and growth", *Industrial and Corporate Change*, 16(3), 455-488.

Schafer, J.L. (1997) *Analysis of Incomplete Multivariate Data*, Chapman & Hall: New York.

Schafer, J.L., Olsen, M.K. (1998) "Multiple imputation for multivariate missing-data problems: A data analyst’s perspective", *Multivariate Behavioral Research*, 33(4), 545-571.

Scherer, F.M. (1965) "Firm Size, Market Structure, Opportunity, and the

Output of Patented Inventions", *The American Economic Review*, 55(5), Part 1, 1097-1125.

Schoenfeld, D. (1982) "Partial residuals for the proportional hazards regression model", *Biometrika*, 69, 239-241.

Schneider, C., Veugelers, R. (2010) "On young highly innovative companies: why they matter and how (not) to policy support them", *Industrial and Corporate Change*, 19(4), 1-39.

Schumpeter, J. (1934) *The Theory of Economic Development*, Harvard University Press: Cambridge, MA.

Schumpeter, J.A. (1939) *Business Cycles*, Volumes 1 and 2, McGraw-Hill: New York.

Schumpeter, J. (1942) *Capitalism, Socialism, and Democracy*, Harper & Row: NY.

Schrader, S. (1991) "Informal Technology Transfer between Firms: Cooperation through Information Trading", *Research Policy*, 20(2), 153-170.

Shane, S. (2000) "Prior Knowledge and the Discovery of Entrepreneurial Opportunities", *Organization Science*, 11(4), 448-469.

Shane, S. (2003) *A general theory of entrepreneurship: The individual-opportunity nexus*, Edward Elgar: Northampton, MA:. Shane, S., Stuart, T. (2002) "Organizational endowments and the performance of university start-ups", *Management Science*, 48, 154-170.

Shane, S., Venkatraman, S. (2000) "The promise of entrepreneurship as a field of research", *Academy of Management Review*, 25(1), 217-226.

Shane, S., Khurana, R. (2003) "Career Experience and Firm Founding", *Industrial and Corporate Change*, 12(3), 519-543.

Shannon, C.E, Weaver, W. (1948) "A mathematical theory of communication", *The Bell Technical System Journal*, 27, 379-423 & 623-656.

Sharif, N. (2006) "Emergence and development of the National Innovation Systems concept", *Research Policy*, 35, 745-766.

Shen, C-W., Chen, Y-H. (2013) "Model selection of generalized estimating equations with multiple imputed longitudinal data", *Biometrical Journal*, 55(6), 899-911.

Siepal, J., Cowling, M., Coad, A. (2016) "Non-founder human capital and the long-run growth and survival of high-tech ventures", *Technovation*, http://dx.doi.org/10.1015/j.technovation.2016.09.001

Simon, H.A. (1955) "A Behavioral Model of Rational Choice", *Quarterly Journal of Economics*, 69, 99-118.

Smith, A. (1776/1887) *An Inquiry Into the Nature and Causes of the Wealth of Nations*, Nelson and Sons.

Smith, K.H. (2002) "What is the 'Knowledge Economy'? Knowledge intensity and distributed knowledge bases". Discussion Paper. United Nations University, Institute for New Technologies, Maastricht, The Netherlands.

Smith, K. (2005) "Changing economic landscapes: liberalisation and knowledge infrastructures", *Changing environments*, 32(5), 339-347.

Smith, K.H., Collins, C.J, Clark, K.D. (2005) "Existing knowledge, knowledge creation capability, and the rate of new product introduction in high-technology firms", *Academy of Management Journal*, 48(2), 346-357.

Smith, K.H. (2009) "Innovation in low-and medium-technology industries", *Research Policy*, 38, 441-446.

Solow, R.M. (1957) "Technical Change and the Aggregate Production Function", *The Review of Economics and Statistics*, 39(3), 312-320.

Spender, J.-C. (1992) "Limits to learning from the West: how Western management advice may prove limited in Eastern Europe", *International Executive*, 34(5), 389-410.)

Spender, J.-C. (1994) "Organizational Knowledge, Collective Practice and Penrose Rents", *International Business Review,* 3(4), 353-367.

Spender, J.-C. (1996) "Making Knowledge the Basis of a Dynamic Theory of the Firm", *Strategic Management Journal*, 17, 45-62.

Spender, J.-C., Grant, R. (1996) "Knowledge and the Firm: Overview", *Strategic Management Journal*, 17, 5-9.

Stehr, N. (1994) *Knowledge Societies*, Sage Publications: London.

Stinchcombe, A. L. (1965) "Social structure and organizations", in: March, J.G. (ed.), *Handbook of organizations*, 153-193. Chicago: Rand McNally

Strambach S. (2001) "Innovation processes and the role of knowledge-intensive business services", in: Koschatzky, K., Zulicke, M., Zenker, A. (eds.) I*nnovation networks: concepts and challenges in the European perspectives*, Physica-Verlag: Heidelberg.

Strambach, S. (2008) "Knowledge-Intensive Business Services (KIBS) as drivers of multilevel knowledge dynamics", *International Journal of Services Technology and Management*, 10, 152-174.

Storey, D.J., Tether, B.S. (1998) "New technology-based firms in the European Union: an introduction", *Research Policy*, 26(9), 933-946.

Stuart, T.E., Ding, W.W. (2006) "When do scientists become entrepreneurs? The social structural antecedents of commercial activity in the academic life sciences", *Academy of Management Journal of Sociology*, 112, 97-144.

Sundbo, J., Gallouj, F. (2000) "Innovation as a loosely coupled system in services", International *Journal of Services Technology and Management*, 1(1), 15-36.

Szeto, E. (2000) "Innovation capacity: working towards a mechanism for improving innovation within an inter-organizational network", *The TQM Magazine*, 12, 149-158.

Teece, D.J. (1980) "Economics of Scope and the Scope of the Enterprise", *Journal of Economic Behavior and Organization*, 1(3), 223-247.

Teece, D.J. (1986) "Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy", *Research Policy*, 15, 285-305.

Teece, D.J., Pisano, G.P., Shuen, A. (1997) "Dynamic capabilities and strategic management", *Strategic Management Journal*, 18, 509-533.

Teece, D.J. (2007) "Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance", *Strategic Management Journal*, 28, 1319-1350.

Tell, F. (2004) "What do Organizations know? Dynamics and Justification Contexts in R&D activities", *Organization*, 11(4), 443-471.

Tether, B. (2002) "Who Co-Operates for Innovation, and Why: An Empirical Analysis", *Research Policy*, 31(6), 947-967.

Tether, B.S. (2005) "Do services innovate (differently)? Insights from the European Innobarameter survey", *Industry and Innovation*, 12(2), 153.

Tether, B. (2014) "Services, Innovation, and Managing Service Innovation", in Dodgson, M., Gann, D.M., Phillips, N. (eds.) *The Oxford Handbook of Innovation Management* (2014), Oxford University Press: Oxford.

Tether, B.S., Hipp, C. (2002) "Knowledge intensive, technical, and other services: patterns of competitiveness and innovation compared", *Technology Analysis & Strategic Management*, 14(2), 163-182.

Tether, B., Tajar, A. (2008) "Beyond industry-university links: Sourcing knowledge for innovation from consultants, private research organisations and the public science-base", *Research Policy*, 37(6-7), 1079-1095.

Thorgren, S., Wincent, J., Boter, H. (2012)."Small firms in multipartner R&D alliances: Gaining benefits by acquiescing", *Journal of Engineering and Technology Management*, 29, 453-467.

Thorpe, R., Holt, R. Macpherson, A., Pittaway, L. (2005) "Using knowledge within small and medium-sized firms: A systematic review of the evidence", *International Journal of Management Reviews*, 7(4), 257-281.

Tödtling, F., Lehrer, P., Trippl, M. (2006) "Innovation in knowledge intensive industries: The nature of geography and knowledge links", *European Planning Studies*, 14(8), 1035-1058.

Tsai, W. (2001) "Knowledge Transfer in Intraorganizational Networks: Effects of Network Position and Absorptive Capacity on Business Unit Innovation and Performance", *The Academy of Management Journal*, 44(5), 996-1004.

Tukey, J. (1949) "One degree of freedom for non-additivity", *Biometrics*, 5(3), 232-242.

Ucbasaran, D., Lockett, A., Wright, M., Westhead, P. (2003) "Entrepreneurial founder teams: Factors associated with member entry and exit", *Entrepreneurship Theory and Practice*, 28(2), 107-127.

Unger, J.M., Rauch, A., Frese, M., Rosenbusch, N. (2011) "Human capital and entrepreneurial success: a meta-analytical review", *Journal of Business Venturing*, 26, 341-358.

Utterback, J.M., Abernathy, W.J. (1975) "A Dynamic Model of Process and Product Innovation", *International Journal of Management Science*, 3(6), 639-656.

Uzzi, B. (1999) "Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing", *American Sociological Review*, 64(4), 481-505.

Vahter, P., Love, J., Roper, S. (2012) "Openness and innovation performance: are small firms different?" Working Paper No. 113, *CSME Working Papers*, Warwick Business School.

Van Auken, H., Madrid-Guijarro, A, Garcia-Perez-de-Lema, D. (2008) “Innovation and performance in Spanish manufacturing SMEs”, *International Journal of Entrepreneurship and Innovation Management*, 8(1), 36-56.

van de Vrande, V., de Jong, J., Vanhaverbeke, W., de Rochemont, M. (2008) “Open Innovation in SMEs: Trends, motives and management challenges”, *Technovation*, 29(6-7), 423-437.

van de Vrande, V., Vanhaverbeke, W., Gassmann, O. (2010) “Broadening the scope of open innovation; past research, current state and future directions”, *International Journal of Technology Management*, 52(3-4), 221-235.

Vanhaverbeke, W. (2012) “Open innovation in SMEs: how can small companies and start-ups benefit from open innovation strategies”, Research report, available at: www.Flandersdc.Be.

Venables, W.N., Ripley, B.D. (2002) *Modern Applied Statistics with S, 4th edition*, Springer: NY, USA.

Venkatraman, N., Ramanujam, V. (1986) ”Measurement of Business Performance in Strategy Research: A Comparison of Approaches”, *The Academy of Management Review*, 11(4), 801-814.

von Hippel, E. (1988) *The Sources of Innovation*, Oxford University Press, New York and London.

Von Tunzelmann, N., Acha, V. (2005) “Innovation in “Low-Tech” Industries”, in Fagerberg, J., Mowery, D.C., Nelson, R. R. (eds.) *The Oxford Handbook of Innovation* (2005), Oxford University Press: Oxford; New York.

Wally, S., Baum, J. (1994) “Personal and Structural Determinants of the Pace of Strategic Decision Making”, *Academy of Management Journal*, 37(4), 932-956.

Wang, C.L., Ahmed, P.K. (2007) “Dynamic capabilities: A review and research agenda”, International *Journal of Management Reviews*, 9, 31-51.

Wennberg, K., Wiklund, J., Wright, M. (2011) “The effectiveness of university knowledge spillovers: Performance differences between university spinoffs and corporate spinoffs,” *Research Policy*, 40, 1128-1143.

Wernerfelt, B. (1984) “The Resource-Based View of the Firm,” *Strategic Management Journal*, 5(2), 171-180.

Weterings, A., Koster, S. (2007) "Inheriting knowledge and sustaining relationships: What stimulates the innovative performance of small software firms in the Netherlands?" *Research Policy*, 36(3), 320-335.

Williamson, O.E. (1979) “Transaction-Cost Economics: The Governance of Contractual Relations”, *Journal of Law and Economics*, 22(2), 233-261.

Williamson, O.E. (1981) “The Economics of Organization: The Transaction Cost Approach”, *American Journal of Sociology*, 87(3), 548-577.

Wilke, H. (1998) *Systemisches Wissensmanagement*, Lucius and Lucius: Stuttgart.

Wooldridge, J.M. (2002/2012) *Econometric Analysis of Cross Section and Panel Data*, The MIT Press: Cambridge, MA.

Yeo, I-K., Johnson, R. (2000) “A new family of power transformations to improve normality or symmetry”, *Biometrika*, 87, 954-959.

Yli-Renko, H., Autio, E., Sapienza, H. (2001) “Social capital, knowledge acquisition and knowledge exploitation in young technology-based firms”, *Strategic Management Journal*, 22(6-7), 587-613.

Yli-Renko, H., Autio, E., Tontti, V. (2002) “Social capital, knowledge, and the international growth of technology-based new firms”, *International Business Review*, 11(3), 279-304.

Young, A.A. (1928) “Increasing returns and economic progress”, *Economic Journal*, 38, 527-542.

Young, F. (1981) “Quantitative analysis of qualitative data”, *Psychometrika*, 46(4), 357-388.

Zahra, S., George, G. (2002) “Absorptive Capacity: A Review, Reconceptualization, and Extension”, *The Academy of Management Review*, 27(2), 185-203.

Zahra, S., Van de Velde, E., Larraneta, B. (2007) “Knowledge conversion capability and the performance of corporate and university spin-offs”, *Industrial and Corporate Change*, 16(4), 569-608.

Zollo, M., Winter, S.G. (2002) “Deliberate Learning and the Evolution of Dynamic Capabilities”, *Organization Science*, 13(3), 339-351.

Appendix

8.1 Alternating Least Squares Transformations

Figure 8.1: Alternating least squares optimal scaling (ALSOS) transformation of RadInn variable

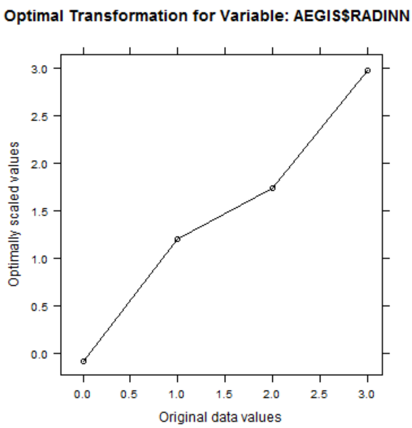


Figure 8.1 and Table 8.1 show the results of the rescaling process. We can see that, aside from making the ordinal variable optimized for OLS regression, the R^2 of the rescaled variable is about a 0.041 improvement over the ordinal untransformed variable. A small improvement to be sure. But, nonetheless preferable to the alternative, and the rescaling optimizes the variable for use in OLS regression as a response variable!

Table 8.1: Optiscale R-squared iterations and improvements

Original Value	**0**	**1**	**2**	**3**
Scaled value	-0.07893	1.20908	1.74038	2.98457
# of obs.	1364	1261	588	216

Iteration	**R-squared**	**Improvement**
1	0.1238525	0.12385
2	0.1274784	0.00362
3	0.1279307	0.00045

8.2 Graphical interpretation, diagnostics and robustness checks for Models 1 and 2

Table 8.2: Brant test of parallel regression assumptions for 1.3, spec.VI

Variable	**chi2**	**p>chi2**	df
All	156.95	0.000	28
Breadth	3.78	0.151	2
$Breadth^2$	4.20	0.122	2
Depth	3.36	0.186	2
$Depth^2$	4.78	0.092	2
Firm_age	1.63	0.442	2
logEmp	17.77	0.000	2
IntlSales	12.95	0.002	2
R&DInt	34.52	0.000	2
2.Sector	28.11	0.000	2
3.Sector	25.91	0.000	2
4.Sector	33.21	0.000	2
2.SectorCLASS	5.57	0.062	2
3.SectorCLASS	4.62	0.099	2
4.SectorCLASS	11.39	0.003	2

A significant test statistic provides evidence that the parallel regression assumption has been violated

Table 8.3: Brant test of parallel regression assumptions for 1.6, spec. VII

Variable	**chi2**	**p>chi2**	df
All	142.86	0.000	32
EXPC1	0.10	0.952	2
$EXPC1^2$	2.54	0.281	2
EXPC2	3.81	0.149	2
$EXPC^2$	4.29	0.117	2
EXPC3	10.26	0.006	2
$EXPC3^2$	0.79	0.674	2
Firm_age	1.85	0.396	2
logEmp	18.51	0.000	2
IntlSales	12.29	0.002	2
R&DInt	33.21	0.000	2
2.Sector	26.03	0.000	2
3.Sector	23.63	0.000	2
4.Sector	30.81	0.000	2
2.SectorCLASS	6.82	0.033	2
3.SectorCLASS	4.85	0.089	2
4.SectorCLASS	12.32	0.002	2

A significant test statistic provides evidence that the parallel regression assumption has been violated

Figure 8.2: Marginal model plots of 1.1 and 1.2

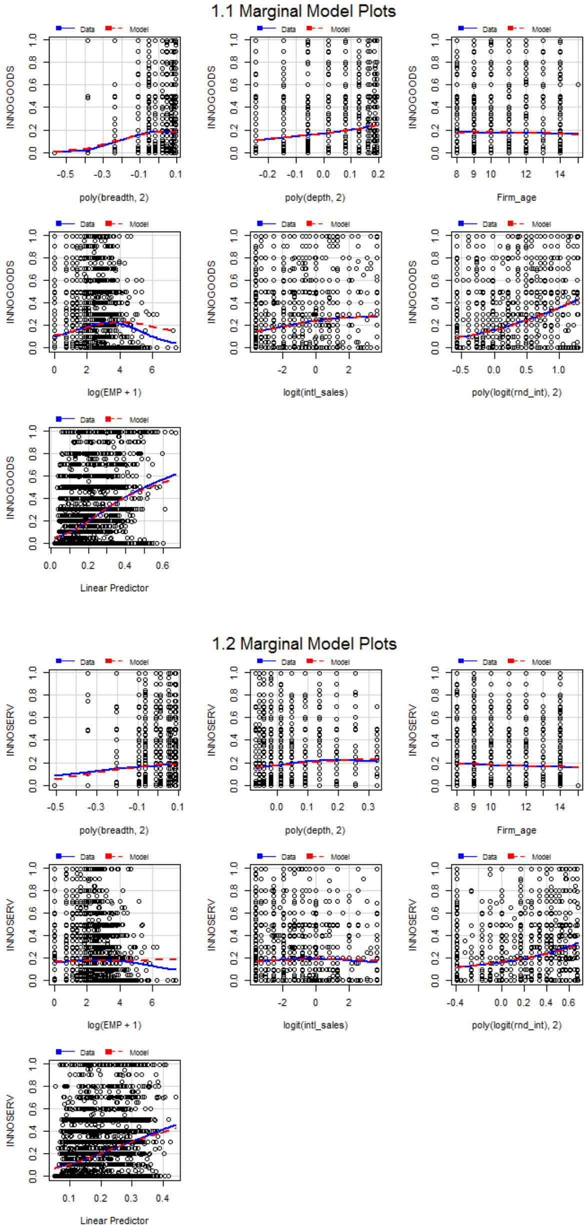
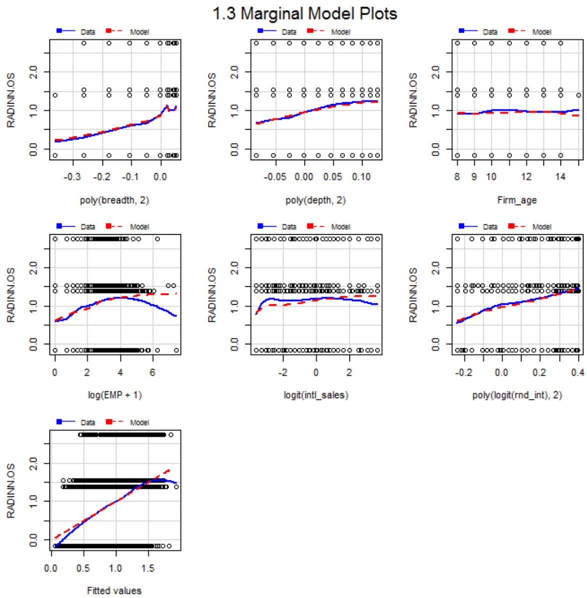


Figure 8.3: Marginal model plots of 1.3 and 1.4



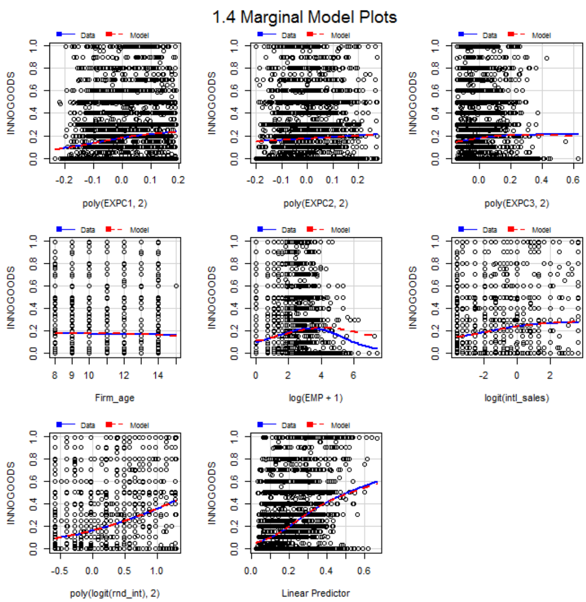


Figure 8.4: Marginal model plots of 1.5 and 1.6

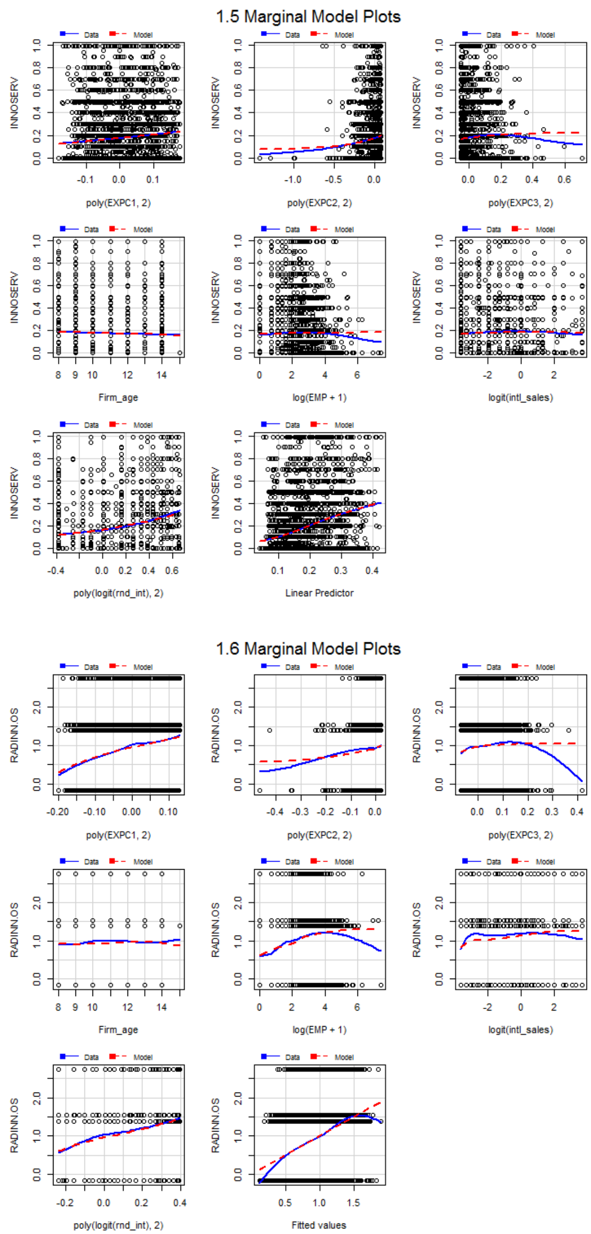


Table 8.4: Variance Inflation Factor tests for Models 1.1 – 1.3

VIF for 1.1

	GVIF	Df	$GVIF^{1/2*Df}$
poly(Breadth, 2)	1.70	2.00	1.14
poly(Depth, 2)	1.91	2.00	1.18
FirmAge	1.03	1.00	1.01
log(Emp + 1)	1.15	1.00	1.07
logit(IntlSales)	1.10	1.00	1.05
poly(logit(r&d_int), 2)	1.19	2.00	1.05
SectorCLASS	1.19	3.00	1.03
Sector	1.29	3.00	1.04

VIF for 1.2

	GVIF	Df	$GVIF^{1/2*Df}$
poly(Breadth, 2)	1.78	2.00	1.15
poly(Depth, 2)	1.99	2.00	1.19
FirmAge	1.03	1.00	1.01
log(Emp + 1)	1.13	1.00	1.07
logit(IntlSales)	1.10	1.00	1.05
poly(logit(r&d_int), 2)	1.21	2.00	1.05
SectorCLASS	1.17	3.00	1.03
Sector	1.28	3.00	1.04

VIF for 1.3

	GVIF	Df	$GVIF^{1/2*Df}$
poly(Breadth, 2)	1.78	2.00	1.16
poly(Depth, 2)	2.00	2.00	1.19
FirmAge	1.03	1.00	1.01
log(Emp + 1)	1.15	1.00	1.07
logit(IntlSales)	1.10	1.00	1.05
poly(logit(r&d_int), 2)	1.21	2.00	1.05
SectorCLASS	1.18	3.00	1.03
Sector	1.30	3.00	1.05

Table 8.5: Variance Inflation Factor tests for Models 1.4 – 1.6

VIF for 1.4

	GVIF	Df	$\text{GVIF}^{1/2*Df}$
poly(pca1, 2)	1.34	2.00	1.08
poly(pca2, 2)	1.81	2.00	1.16
poly(pca3, 2)	1.15	2.00	1.03
FirmAge	1.03	1.00	1.01
log(Emp)	1.11	1.00	1.06
logit(IntlSales)	1.10	1.00	1.05
poly(logit(r&d_int), 2)	1.22	2.00	1.05
SectorCLASS	1.23	3.00	1.04
Sector	1.29	3.00	1.04

VIF for 1.5

	GVIF	Df	$\text{GVIF}^{1/2*Df}$
poly(pca1, 2)	1.37	2.00	1.08
poly(pca2, 2)	1.74	2.00	1.15
poly(pca3, 2)	1.14	2.00	1.03
FirmAge	1.03	1.00	1.01
log(Emp + 1)	1.13	1.00	1.06
logit(IntlSales)	1.10	1.00	1.05
poly(logit(r&d_int), 2)	1.24	2.00	1.05
SectorCLASS	1.22	3.00	1.03
Sector	1.30	3.00	1.04

VIF for 1.6

	GVIF	Df	$\text{GVIF}^{1/2*Df}$
poly(pca1, 2)	1.41	2.00	1.09
poly(pca2, 2)	1.83	2.00	1.16
poly(pca3, 2)	1.14	2.00	1.03
FirmAge	1.03	1.00	1.01
user	1.77	1.00	1.33
log(Emp + 1)	1.14	1.00	1.07
logit(IntlSales)	1.10	1.00	1.05
poly(logit(r&d_int), 2)	1.23	2.00	1.05
SectorCLASS	1.22	3.00	1.03
Sector	1.32	3.00	1.05

Figure 8.5: Effects plots for 2.1 spec. VI – EXPC1 as response variable

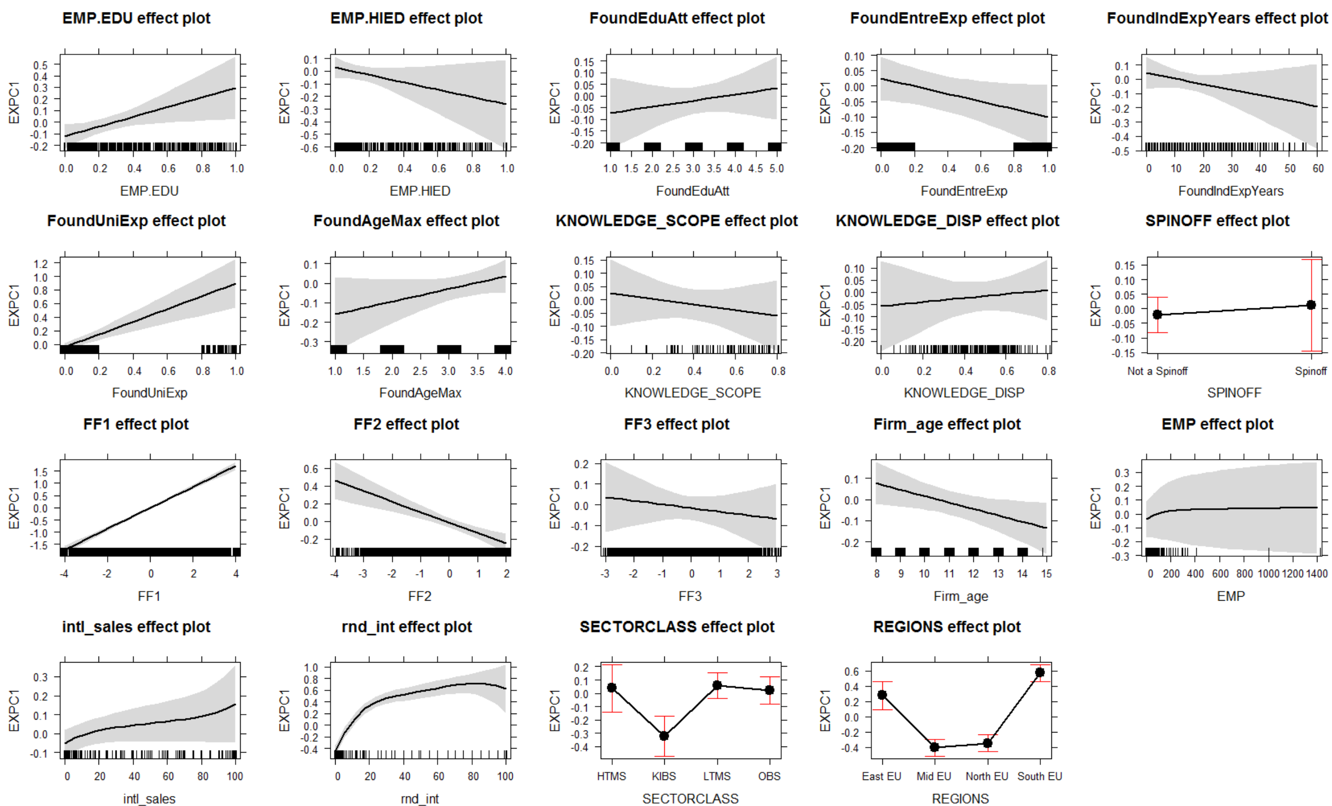


Figure 8.6: Effects plots for 2.2 spec. VI – EXPC2 as response variable

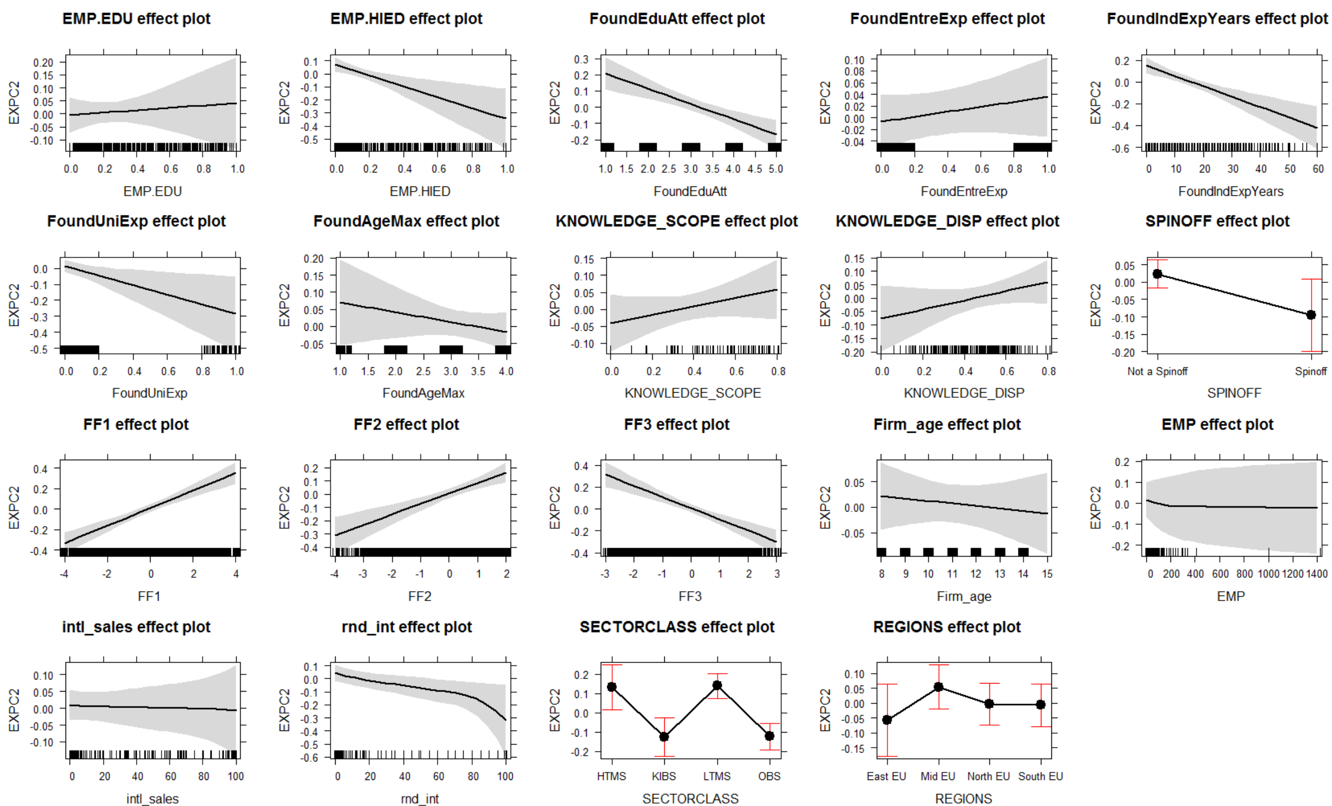


Figure 8.7: Effects plots for 2.3 spec. VI – EXPC3 as response variable

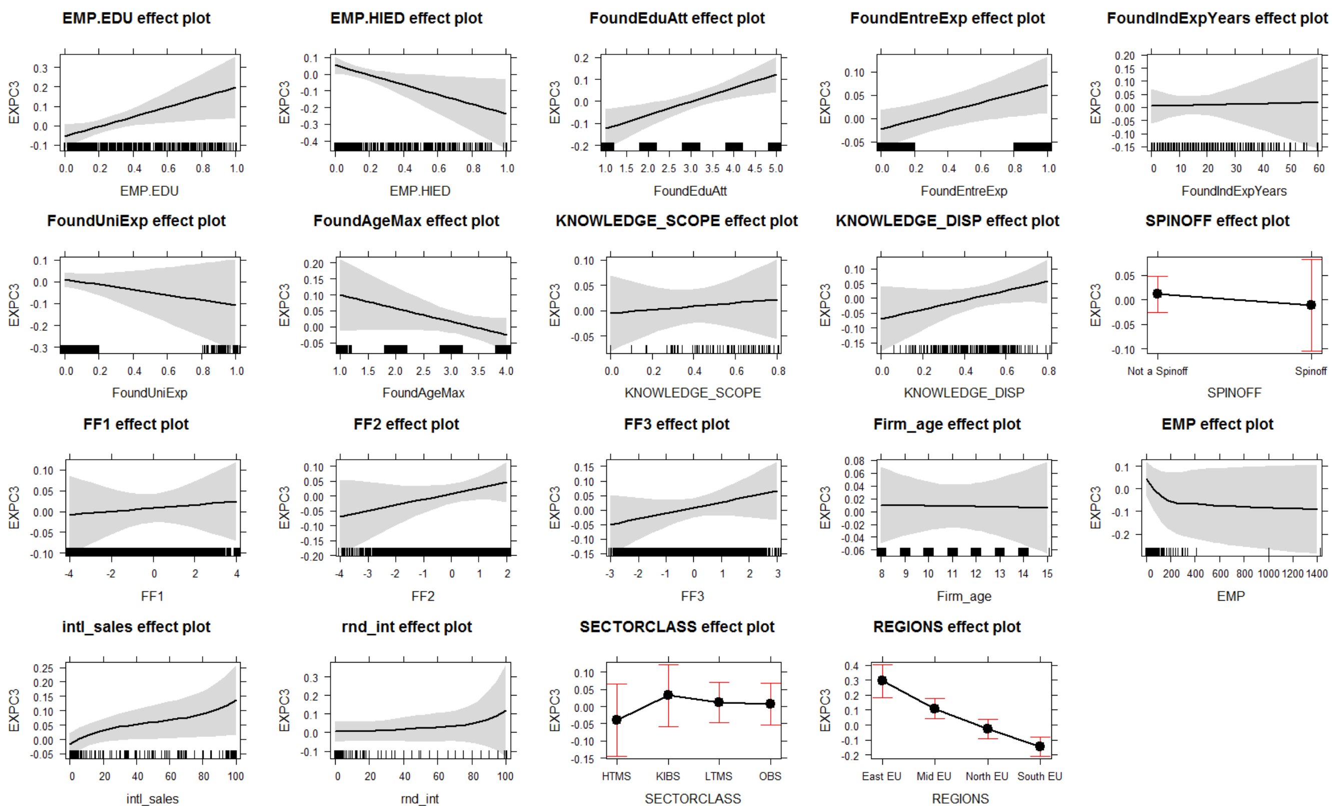


Figure 8.8: Residual Plots of 2.1 spec VI – EXPC1 as response

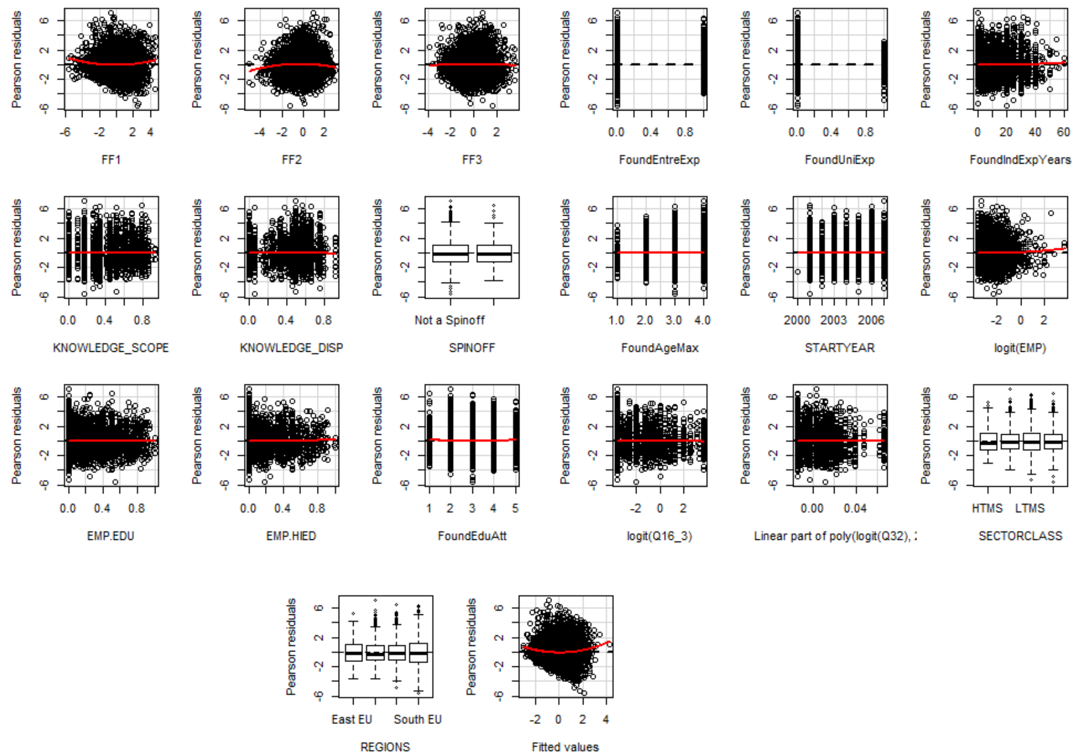


Figure 8.9: Residual Plots of 2.2 spec VI – EXPC2 as response

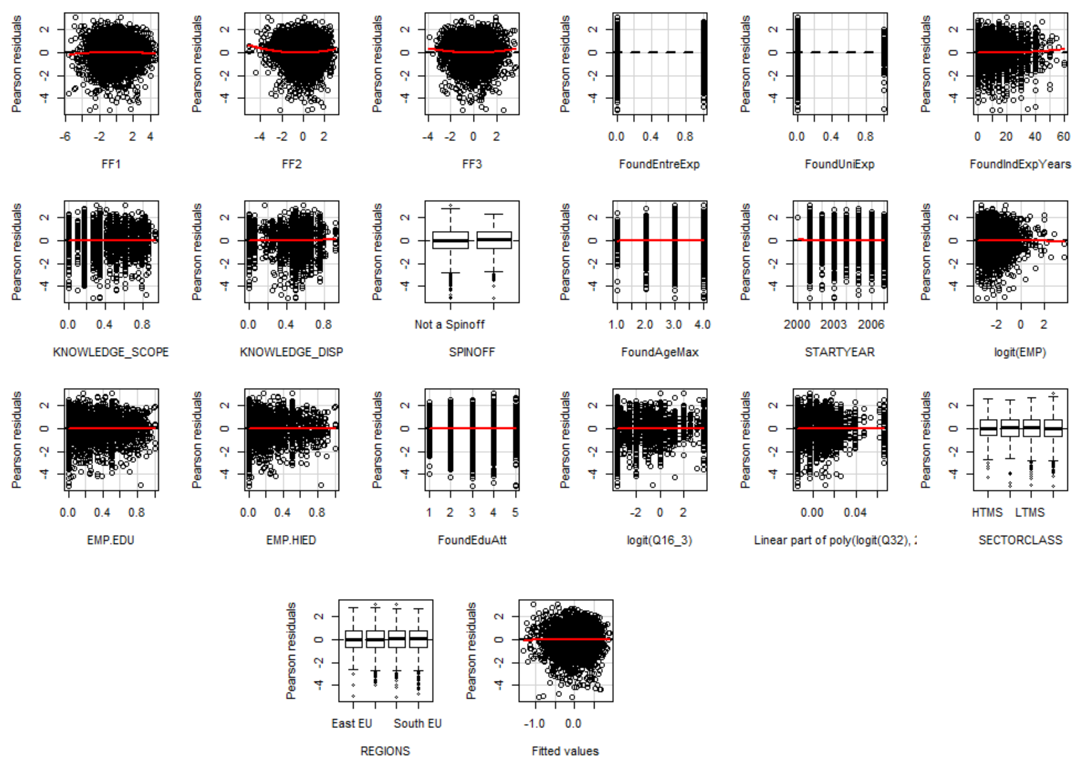


Figure 8.10: Residual Plots of 2.3 spec. VI – EXPC3 as response

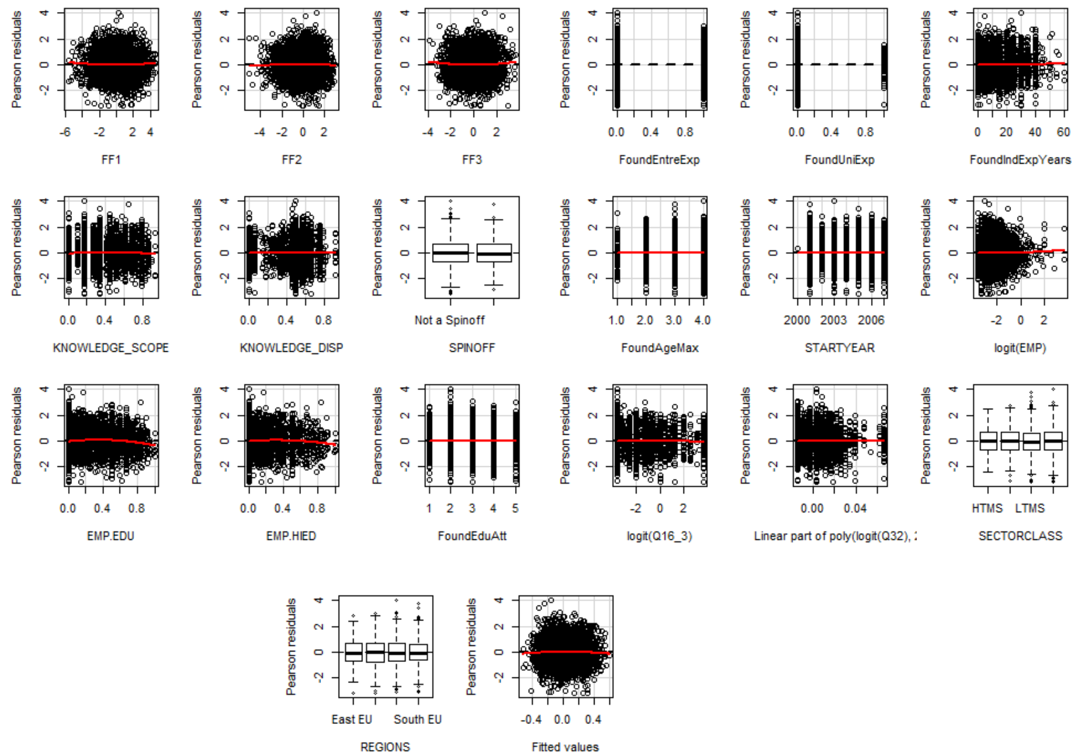


Table 8.6: Lack-of-fit and Tukey tests

2.1

	Test stat	Pr(>|t|)
EmpEdu	0.37	0.71
EmpHiEdu	0.64	0.52
FoundEdu	2.15	0.03
FoundEnt	-0.77	0.44
FoundUni	-1.13	0.26
AgeMax	-0.60	0.55
KScope	0.28	0.78
KDisp	-1.28	0.20
Spinoff		
FF1	4.10	0.00
FF2	-3.02	0.00
FF3	-0.65	0.52
Firm_age	0.60	0.55
log(Emp + 1)	1.20	0.23
logit(IntlSales)	-0.71	0.48
poly(logit(R&DInt), 2)		
SectorCLASS		
Sector		
Tukey test	4.50	0.00

2.2

	Test stat	Pr(>|t|)
EmpEdu	0.76	0.45
FoundEnt	0.67	0.50
FoundUni	0.52	0.60
AgeMax	-1.24	0.22
KScope	0.04	0.96
KDisp	0.89	0.37
Spinoff		
FF1	-2.18	0.03
FF2	2.71	0.01
FF3	2.17	0.03
Firm_age	1.59	0.11
log(Emp + 1)	-1.82	0.07
logit(IntlSales)	-1.23	0.22
poly(logit(R&DInt), 2)		
SectorCLASS		
Sector		
Tukey test	-0.21	0.83

2.3

	Test stat	Pr(>|t|)
EmpEdu	-3.90	0.00
EmpHiEdu	-2.26	0.02
FoundEdu	0.62	0.54
FoundEnt	0.17	0.86
FoundUni	1.55	0.12
AgeMax	-0.67	0.50
KScope	-1.15	0.25
KDisp	-0.08	0.93
Spinoff		
FF1	1.44	0.15
FF2	-0.51	0.61
FF3	1.79	0.07
Firm_age	0.42	0.67
log(Emp + 1)	-0.44	0.66
logit(IntlSales)	-1.25	0.21
poly(logit(R&DInt), 2)		
SectorCLASS		
Sector		
Tukey test	-1.85	0.06

Table 8.7: Variance Inflation tests for Model 2 following modification

VIF for 2.1 Modified

	GVIF	Df	$GVIF^{1/2*Df}$
FF1	1.22	1.00	1.11
poly(FF2, 2)	1.20	2.00	1.05
FF3	1.13	1.00	1.06
KScope	1.31	1.00	1.14
KDisp	1.27	1.00	1.13
Spinoff	1.07	1.00	1.04
EmpEdu	2.56	1.00	1.60
EmpHiEdu	2.58	1.00	1.61
FoundEdu	1.55	1.00	1.25
FoundEnt	1.11	1.00	1.05
FoundUni	1.09	1.00	1.04
FoundInd	1.51	1.00	1.23
AgeMax	1.44	1.00	1.20
FirmAge	1.07	1.00	1.03
log(Emp)	1.19	1.00	1.09
logit(IntlSales)	1.12	1.00	1.06
poly(logit(r&d_int), 2)	1.27	2.00	1.06

VIF for 2.2 Modified

	GVIF	Df	$GVIF^{1/2*Df}$
EmpEdu	1.27	1.00	1.13
FoundEnt	1.11	1.00	1.05
FoundUni	1.06	1.00	1.03
AgeMax	1.10	1.00	1.05
KScope	1.30	1.00	1.14
KDisp	1.27	1.00	1.13
Spinoff	1.07	1.00	1.04
FF1	1.24	1.00	1.11
poly(FF2, 2)	1.14	2.00	1.03
poly(FF3, 2)	1.23	2.00	1.05
Firm_age	1.06	1.00	1.03
log(Emp + 1)	1.21	1.00	1.10
logit(IntlSales)	1.11	1.00	1.06
poly(logit(R&DInt), 2)	1.26	2.00	1.06
SectorCLASS	1.47	3.00	1.07
Sector	1.39	3.00	1.06

VIF for 2.3 Modified

	GVIF	Df	$GVIF^{1/2*Df}$
FF1	1.16	1.00	1.08
FF2	1.11	1.00	1.05
FF3	1.13	1.00	1.06
KScope	1.31	1.00	1.14
KDisp	1.27	1.00	1.13
Spinoff	1.07	1.00	1.04
poly(EmpEdu, 2)	2.76	2.00	1.29
EmpHiEdu	2.60	1.00	1.61
FoundEdu	1.63	1.00	1.28
FoundEnt	1.11	1.00	1.05
FoundUni	1.09	1.00	1.04
FoundInd	1.50	1.00	1.23
AgeMax	1.44	1.00	1.20
FirmAge	1.07	1.00	1.03
log(Emp)	1.19	1.00	1.09
logit(IntlSales)	1.12	1.00	1.06
poly(logit(r&d_int), 2)	1.27	2.00	1.06

8.3 Graphical interpretation, diagnostics and robustness checks for Models 3 and 4

Figure 8.11: Overimputation plots for response variables: Models 3 and 4

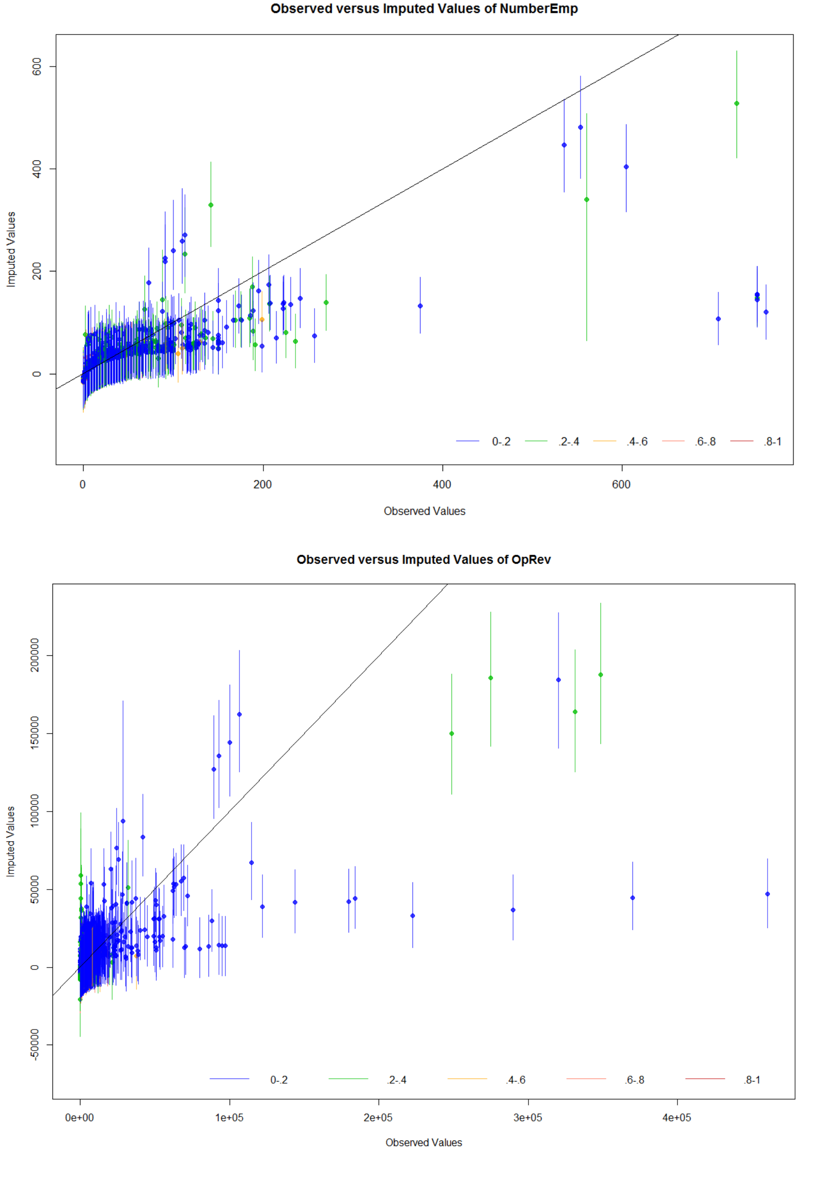
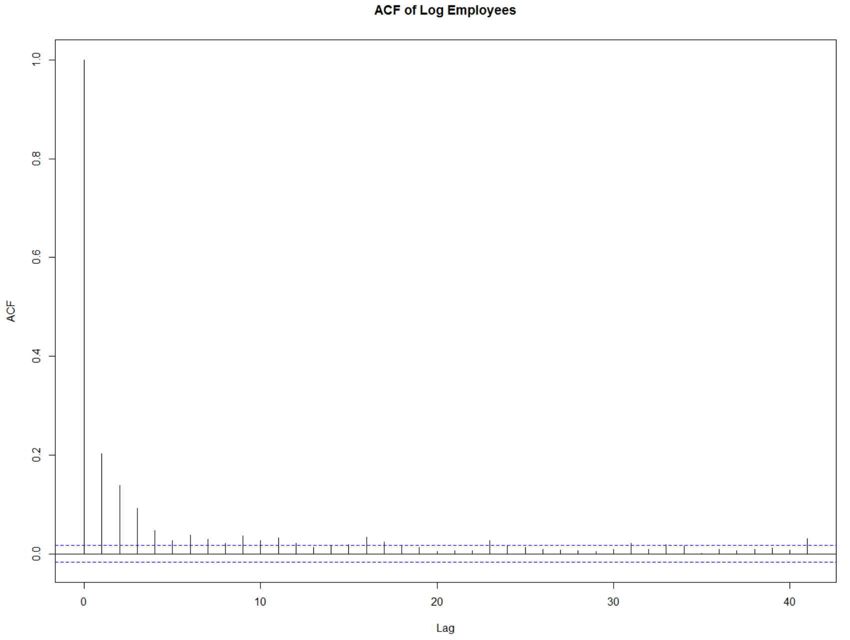


Figure 8.12: Auto- and partial-correlation functions of response variable: # of Employees



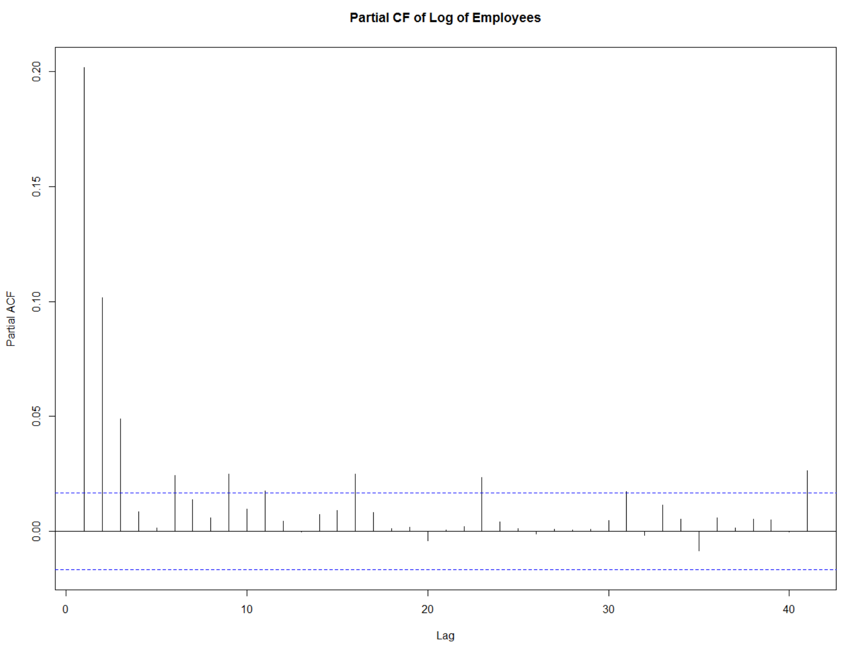


Figure 8.13: Auto- and partial-correlation functions of response variable: Operating Revenue

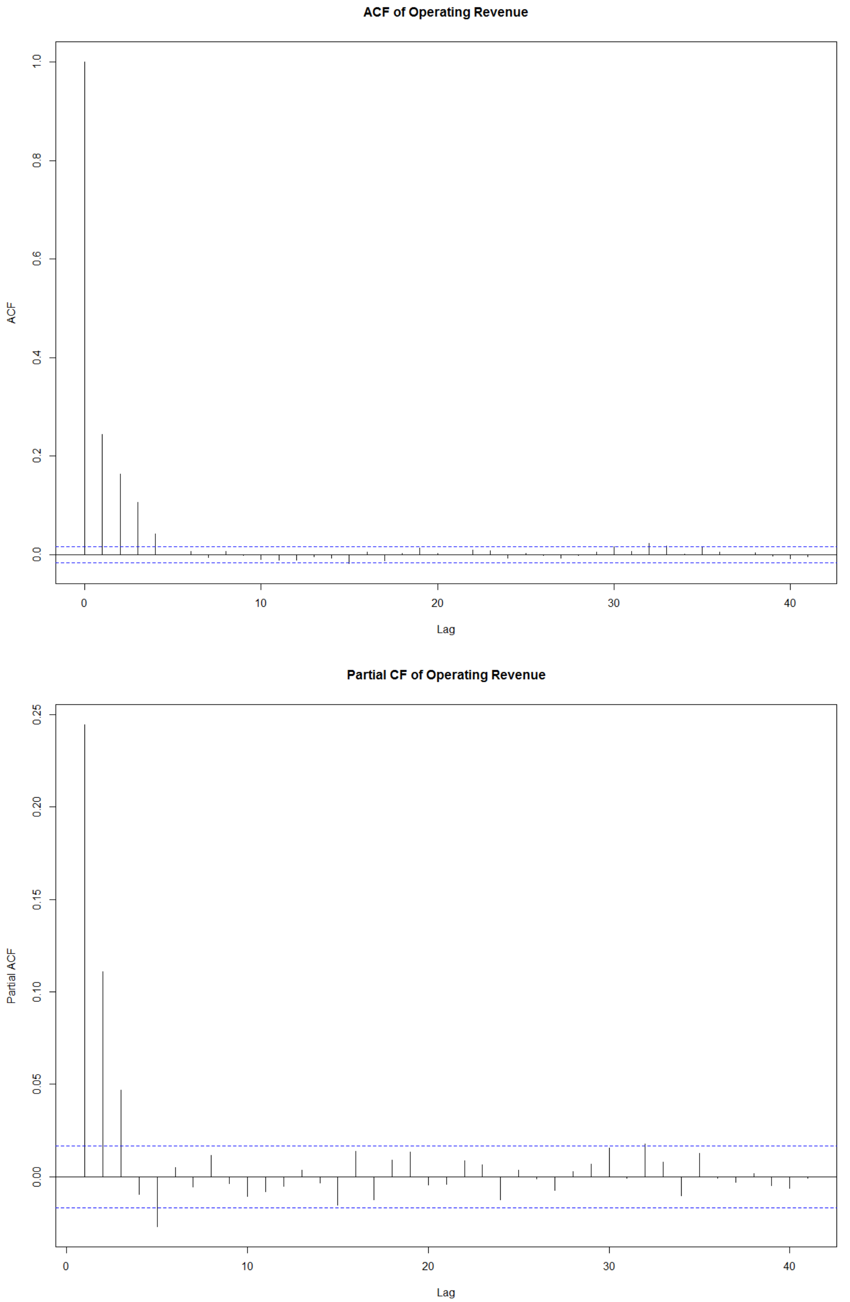


Table 8.8: GEEGLM (intercept) regressions pre-imputation for Models 3 and 4 full specifications

	I log(Emp)	II log(OpRev)	III log(Emp)	IV log(OpRev)
EmpEdu	0.6855^{***}	-0.1357		
	(0.1790)	(0.1984)		
EmpHiEdu	-0.3333	0.3717		
	(0.1936)	(0.2231)		
FoundEdu	-0.0183	0.0128		
	(0.0322)	(0.0406)		
KDisp	-1.1247^{***}	-0.5639^{*}		
	(0.1999)	(0.2381)		
KScope	2.4064^{***}	1.7762^{**}		
	(0.5366)	(0.6525)		
KScope^2	-2.1431^{***}	-1.5852^{*}		
	(0.5662)	(0.6877)		
FoundEnt	0.0219	0.1929^{**}		
	(0.0586)	(0.0731)		
FoundUni	0.0498	-0.4279		
	(0.1631)	(0.2997)		
FoundInd	0.0006	0.0003		
	(0.0032)	(0.0041)		
AgeMax	0.0610	0.0485		
	(0.0391)	(0.0505)		
Spinoff	0.3378^{***}	0.2103^{*}	0.3091^{***}	0.2046^{*}
	(0.0839)	(0.1029)	(0.0843)	(0.1018)
FF1	0.0415^{*}	0.0176		
	(0.0181)	(0.0206)		
FF2	0.0295	0.0450		
	(0.0248)	(0.0316)		
FF3	-0.0600^{*}	-0.0589		
	(0.0239)	(0.0364)		
logit(InnoGoods)			0.0199	0.0590^{*}
			(0.0213)	(0.0253)
logit(InnoServ)			-0.0141	0.0133
			(0.0225)	(0.0244)
RadInnOS			0.0673	-0.0318
			(0.0503)	(0.0589)
logit(R&DInt)	0.0291	-0.0155	0.0131	-0.0245
	(0.0221)	(0.0271)	(0.0228)	(0.0256)
FirmAge	0.0203	0.0033	0.0353^{**}	0.0076
	(0.0121)	(0.0161)	(0.0118)	(0.0150)
lag(log(NumberEmp), 1)		1.0606^{***}		1.0857^{***}
		(0.0477)		(0.0436)
(Intercept)	2.1190^{***}	3.6527^{***}	2.1503^{***}	4.0299^{***}
	(0.2502)	(0.3193)	(0.2328)	(0.2794)
Sectoral controls	Yes	Yes	Yes	Yes
Regional controls	Yes	Yes	Yes	Yes
Scale parameter: gamma	0.8964	1.4004	0.9719	1.3668
Scale parameter: SE	0.0478	0.0934	0.0525	0.0826
Correlation parameter: alpha	0.9618	0.8636	0.9584	0.8565
Correlation parameter: SE	0.0184	0.0206	0.0145	0.0197
Num. obs.	4415	4037	4758	4337
Num. clust.	1306	1118	1415	1195

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 8.9: Model 3.1 Pre-imputation: GEEGLM for all specifications

	N.Emp1	N.Emp	N.Emp	N.Emp	N.Emp	N.Emp	N.Emp
EmpEdu	0.6290^{***}					0.6972^{***}	0.6855^{***}
	(0.1793)					(0.1803)	(0.1790)
EmpHiEdu	$-0.3345^{\dagger}$					$-0.3710^{\dagger}$	$-0.3333^{\dagger}$
	(0.1999)					(0.1966)	(0.1936)
FoundEdu	0.0197					−0.0058	−0.0183
	(0.0332)					(0.0321)	(0.0322)
FoundEnt		0.0283				0.0155	0.0219
		(0.0594)				(0.0585)	(0.0586)
FoundUni		0.1356				0.0694	0.0498
		(0.1722)				(0.1657)	(0.1631)
FoundInd		0.0021				0.0006	0.0006
		(0.0032)				(0.0033)	(0.0032)
AgeMax		0.0899^{*}				0.0640	0.0610
		(0.0402)				(0.0393)	(0.0391)
KDisp			-0.8370^{***}			-0.7648^{***}	-1.1247^{***}
			(0.1705)			(0.1674)	(0.1999)
KScope			0.5013^{***}			0.4508^{***}	2.4064^{***}
			(0.1358)			(0.1347)	(0.5366)
$KScope^2$							-2.1431^{***}
							(0.5662)
Spinoff				0.3662^{***}		0.3484^{***}	0.3378^{***}
				(0.0866)		(0.0847)	(0.0839)
FF1					0.0505^{**}	0.0423^{*}	0.0415^{*}
					(0.0184)	(0.0184)	(0.0181)
FF2					0.0394	0.0302	0.0295
					(0.0241)	(0.0248)	(0.0248)
FF3					-0.0761^{**}	-0.0609^{*}	-0.0600^{*}
					(0.0240)	(0.0240)	(0.0239)
FirmAge	0.0295^{*}	0.0258^{*}	0.0262^{*}	0.0300^{*}	0.0328^{**}	$0.0215^{\dagger}$	$0.0203^{\dagger}$
	(0.0122)	(0.0125)	(0.0121)	(0.0122)	(0.0122)	(0.0122)	(0.0121)
logit(R&DInt)	0.0349	0.0458^{*}	0.0476^{*}	0.0505^{*}	0.0487^{*}	0.0300	0.0291
	(0.0224)	(0.0225)	(0.0217)	(0.0222)	(0.0226)	(0.0223)	(0.0221)
(Intercept)	2.1280^{***}	2.0563^{***}	2.6131^{***}	2.3413^{***}	2.3462^{***}	2.2319^{***}	2.1190^{***}
	(0.2335)	(0.2183)	(0.2196)	(0.2049)	(0.2068)	(0.2510)	(0.2502)
Sectoral Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Scale parameter: gamma	0.9701	0.9658	0.9571	0.9577	0.9587	0.9098	0.8964
Scale parameter: SE	0.0550	0.0543	0.0531	0.0529	0.0536	0.0493	0.0478
Correlation parameter: alpha	0.9657	0.9621	0.9626	0.9607	0.9614	0.9639	0.9618
Correlation parameter: SE	0.0180	0.0181	0.0176	0.0183	0.0175	0.0178	0.0184
Num. obs.	4415	4415	4415	4415	4415	4415	4415
Num. clust.	1306	1306	1306	1306	1306	1306	1306

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$

Table 8.10: Model 3.2 Pre-imputation: GEEGLM for all specifications

	OpRev	OpRev	OpRev	OpRev	OpRev	OpRev	OpRev
EmpEdu	-0.2206					-0.1287	-0.1357
	(0.1965)					(0.1978)	(0.1984)
EmpHiEdu	$0.3797^{\dagger}$					0.3500	$0.3717^{\dagger}$
	(0.2251)					(0.2248)	(0.2231)
FoundEdu	0.0375					0.0239	0.0128
	(0.0410)					(0.0411)	(0.0406)
FoundEnt		0.2146^{**}				0.1860^{*}	0.1929^{**}
		(0.0724)				(0.0733)	(0.0731)
FoundUni		-0.3620				-0.4189	-0.4279
		(0.3058)				(0.2996)	(0.2997)
FoundInd		0.0016				0.0003	0.0003
		(0.0040)				(0.0041)	(0.0041)
AgeMax		0.0671				0.0489	0.0485
		(0.0505)				(0.0505)	(0.0505)
KDisp			-0.2955			-0.2919	-0.5639^{*}
			(0.1982)			(0.1958)	(0.2381)
KScope			0.4019^{*}			0.3336^{*}	1.7762^{**}
			(0.1652)			(0.1645)	(0.6525)
KScope^2							-1.5852^{*}
							(0.6877)
Spinoff				0.2539^{*}		0.2177^{*}	0.2103^{*}
				(0.1034)		(0.1031)	(0.1029)
FF1					0.0248	0.0186	0.0176
					(0.0208)	(0.0207)	(0.0206)
FF2					0.0630^{*}	0.0448	0.0450
					(0.0309)	(0.0315)	(0.0316)
FF3					$-0.0643^{\dagger}$	-0.0586	-0.0589
					(0.0362)	(0.0364)	(0.0364)
FirmAge	0.0077	0.0014	0.0046	0.0060	0.0081	0.0044	0.0033
	(0.0156)	(0.0160)	(0.0157)	(0.0157)	(0.0157)	(0.0161)	(0.0161)
logit(R&DInt)	-0.0165	-0.0108	-0.0147	-0.0100	-0.0070	-0.0149	-0.0155
	(0.0264)	(0.0261)	(0.0257)	(0.0256)	(0.0260)	(0.0271)	(0.0271)
lag(log(NumberEmp + 1), 1)	1.0870^{***}	1.0868^{***}	1.0812^{***}	1.0766^{***}	1.0744^{***}	1.0657^{***}	1.0606^{***}
	(0.0466)	(0.0461)	(0.0465)	(0.0467)	(0.0471)	(0.0476)	(0.0477)
(Intercept)	3.7915^{***}	3.6681^{***}	3.9505^{***}	3.9421^{***}	3.9747^{***}	3.7193^{***}	3.6527^{***}
	(0.2895)	(0.2657)	(0.2918)	(0.2681)	(0.2682)	(0.3183)	(0.3193)
Sectoral Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Scale parameter: gamma	1.4244	1.4064	1.4278	1.4295	1.4295	1.4016	1.4004
Scale parameter: SE	0.0947	0.0915	0.0929	0.0940	0.0931	0.0939	0.0934
Correlation parameter: alpha	0.8637	0.8592	0.8638	0.8648	0.8668	0.8639	0.8636
Correlation parameter: SE	0.0212	0.0215	0.0209	0.0209	0.0207	0.0206	0.0206
Num. obs.	4037	4037	4037	4037	4037	4037	4037
Num. clust.	1118	1118	1118	1118	1118	1118	1118

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$

Figure 8.14: 3.1 Multiple imputed effects plots

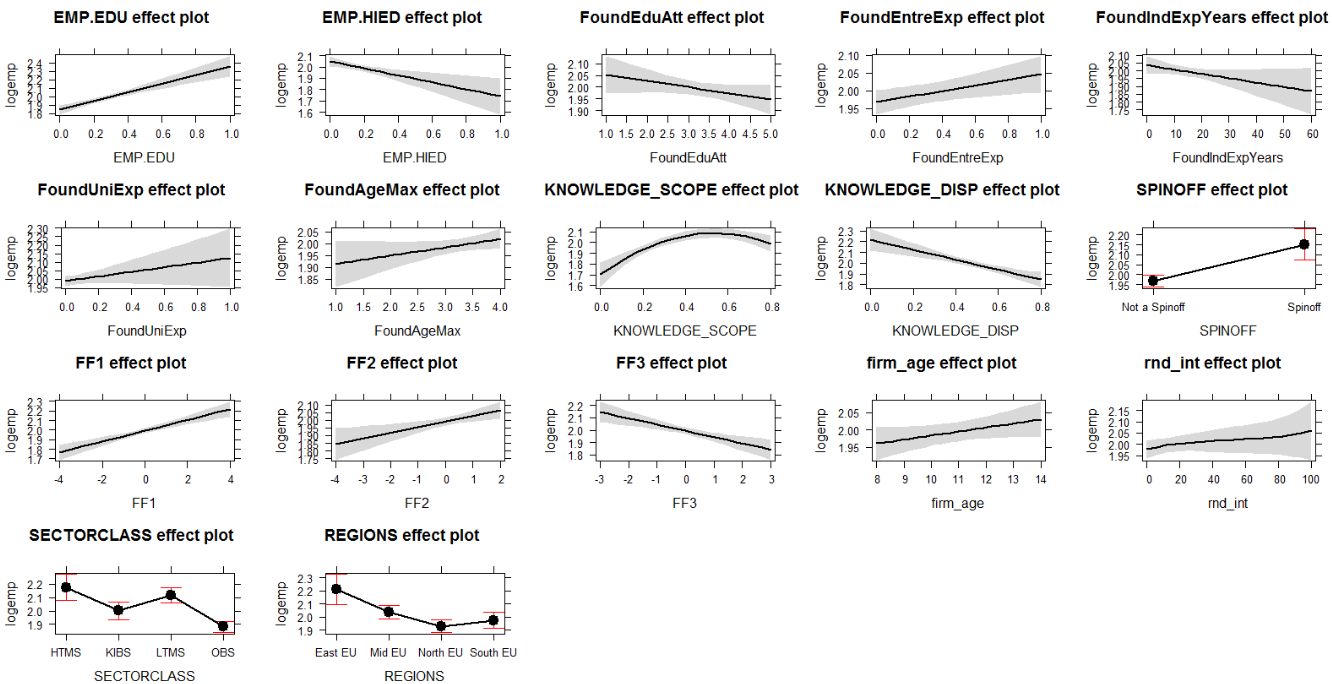


Figure 8.15: Residual plots for 3.1: Working residuals (y-axis) plotted against variables with loess smoother (red line)

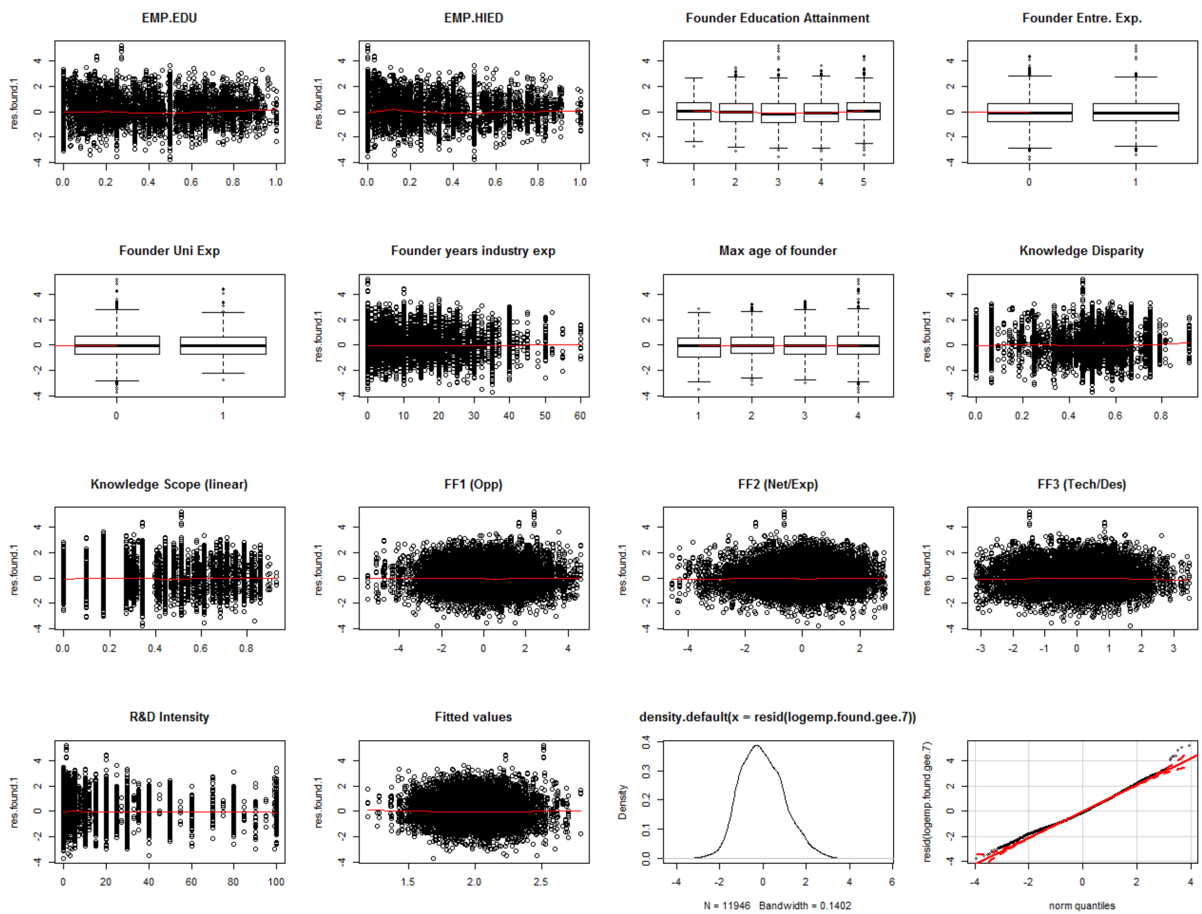


Figure 8.16: 3.2 Multiple imputed effects plots

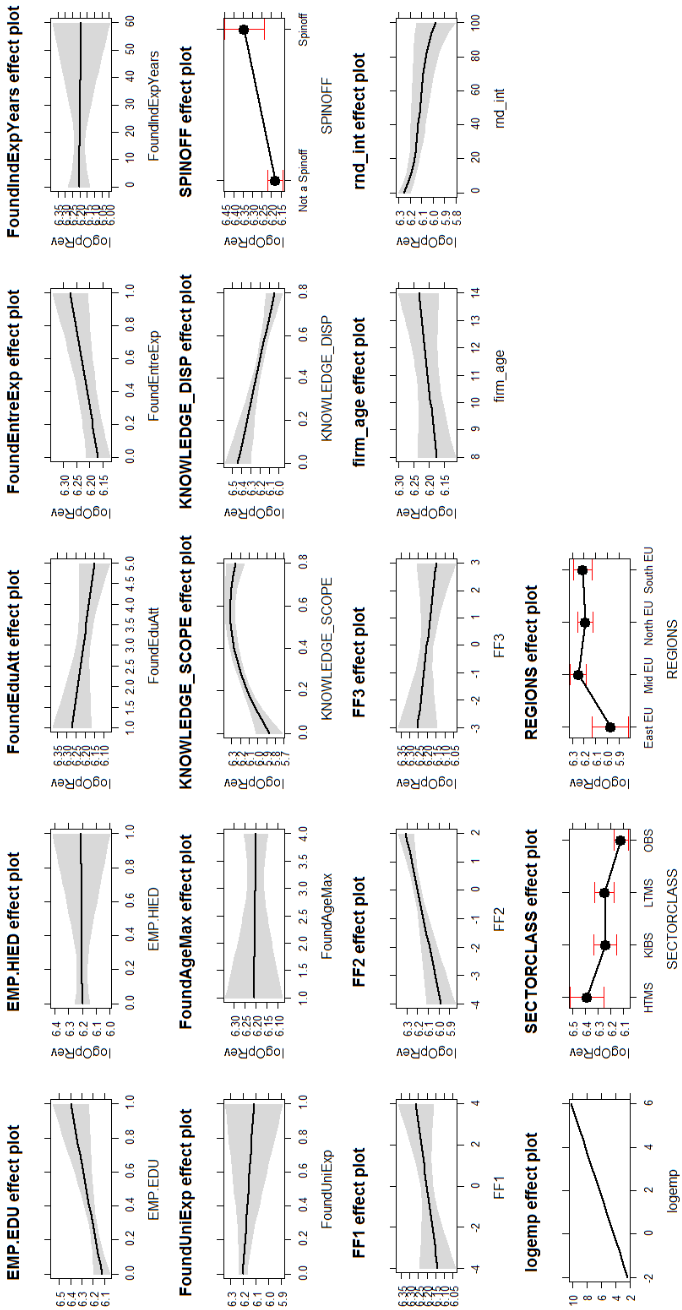


Figure 8.17: Residual plots of 3.2

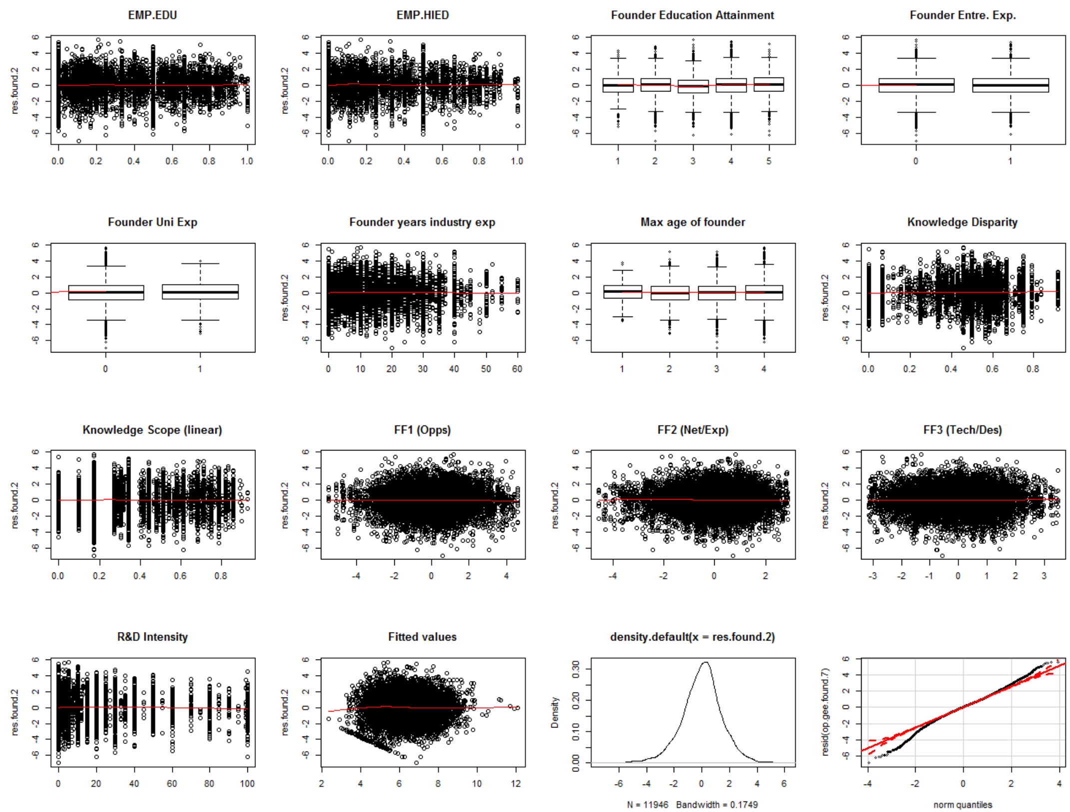


Figure 8.18: Effects plots for Cox PH model 3.3.

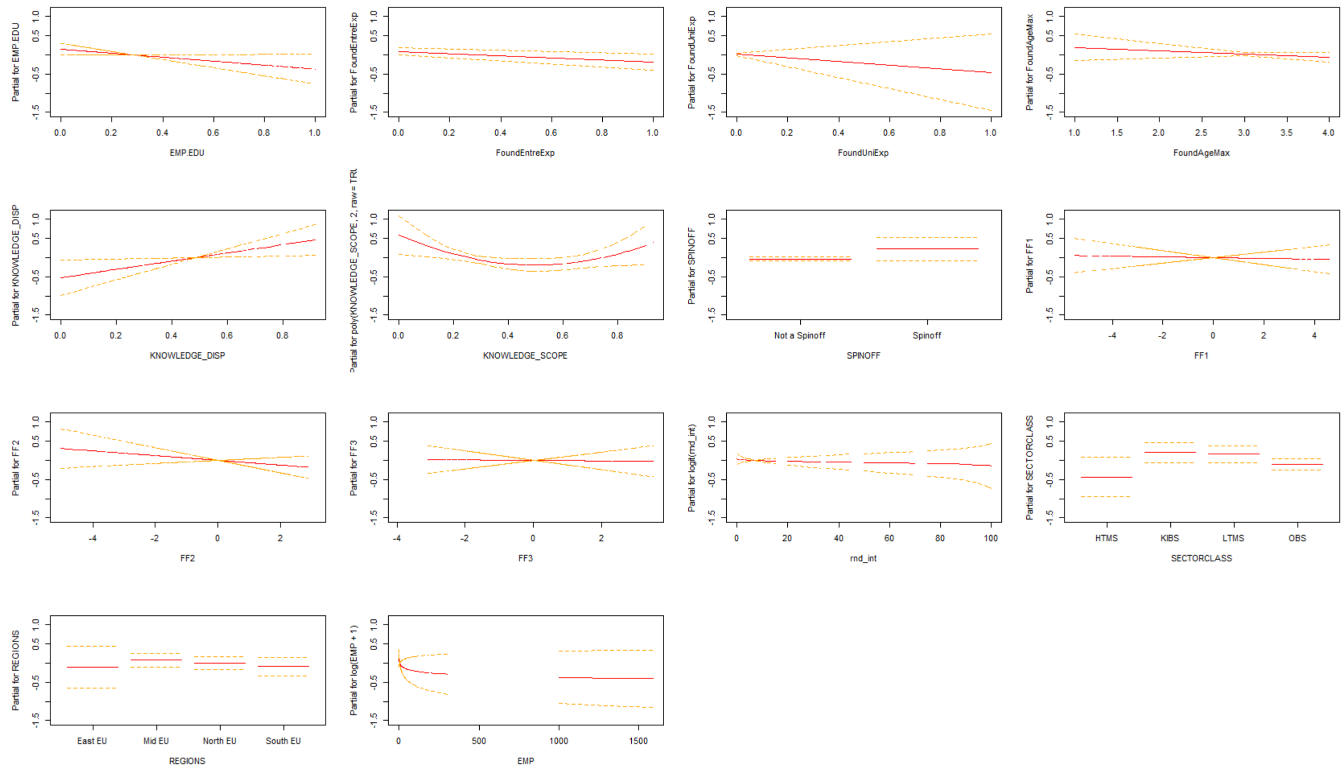


Figure 8.19: Schoenfeld residuals of Model 2.3

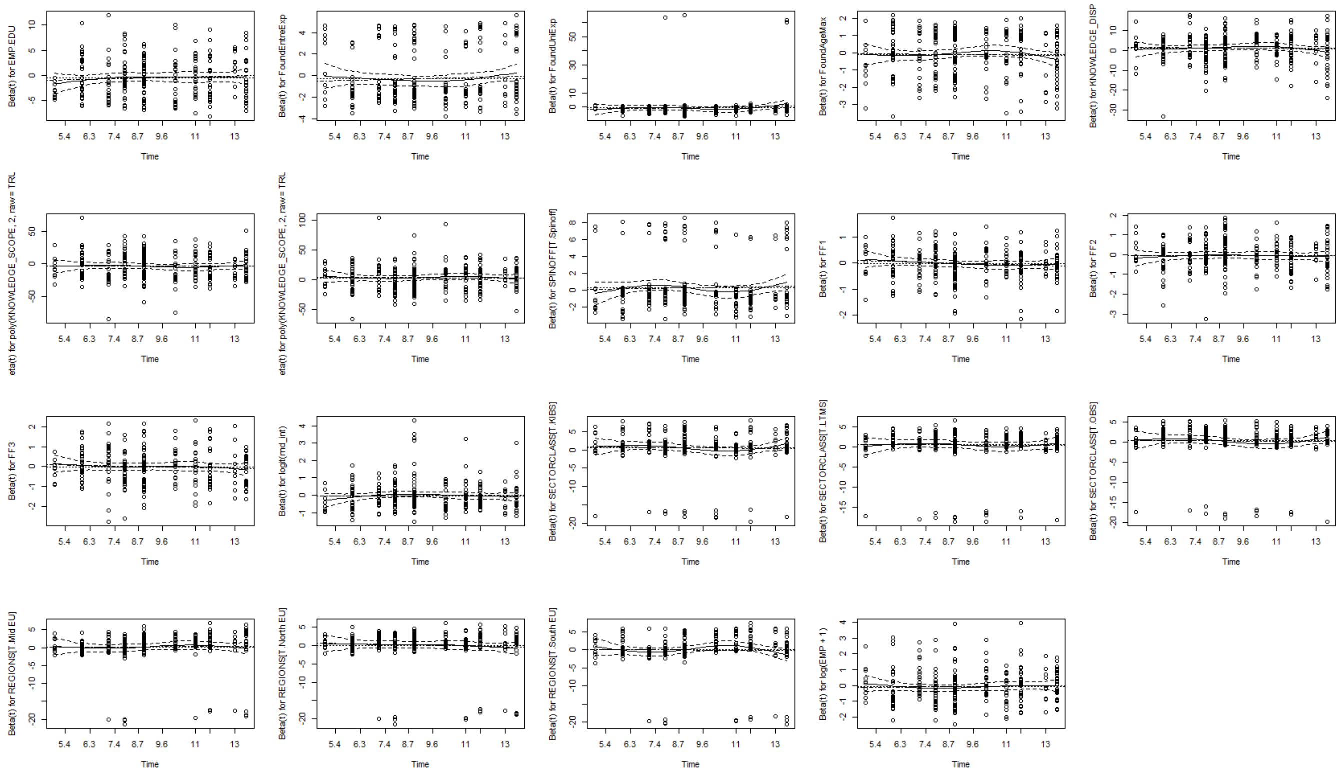


Figure 8.20: 4.1 Multiple imputed effects plots

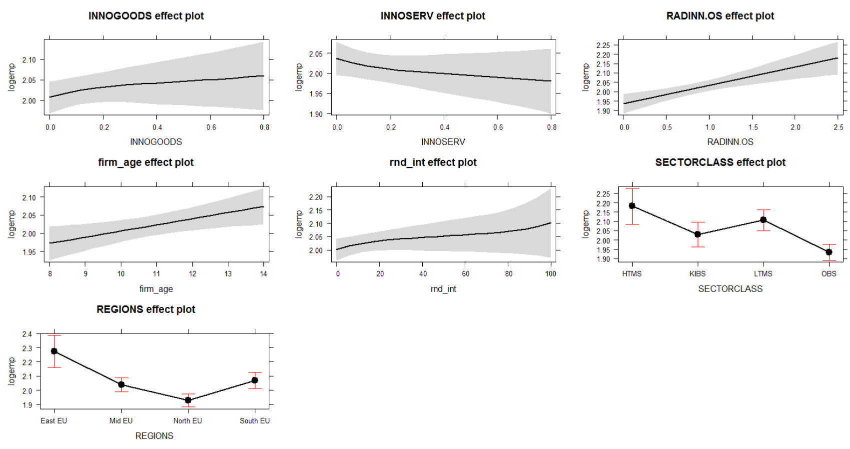


Figure 8.21: Residual plots of 4.1, red line denotes loess smoother line

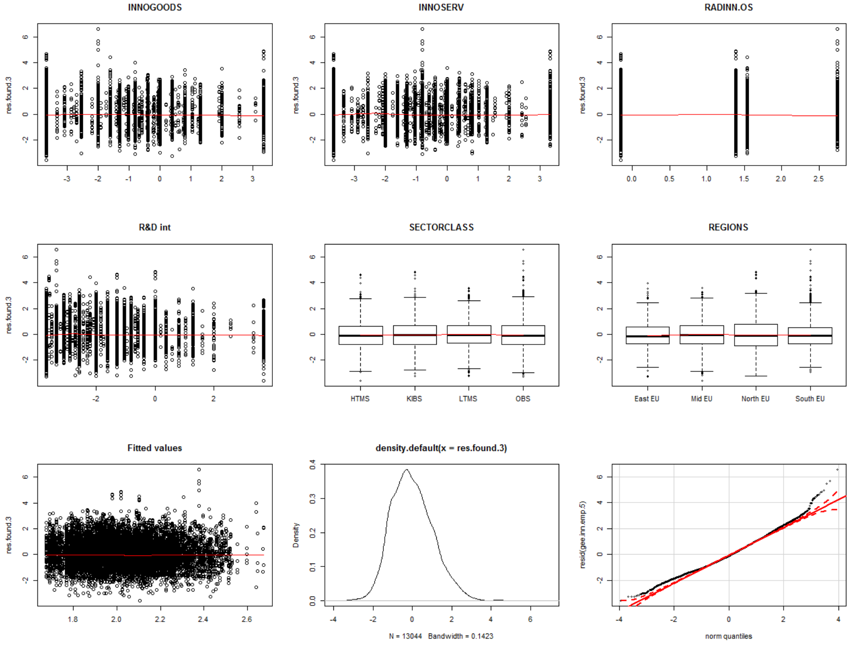


Figure 8.22: 4.2 Multiple imputed effects plots

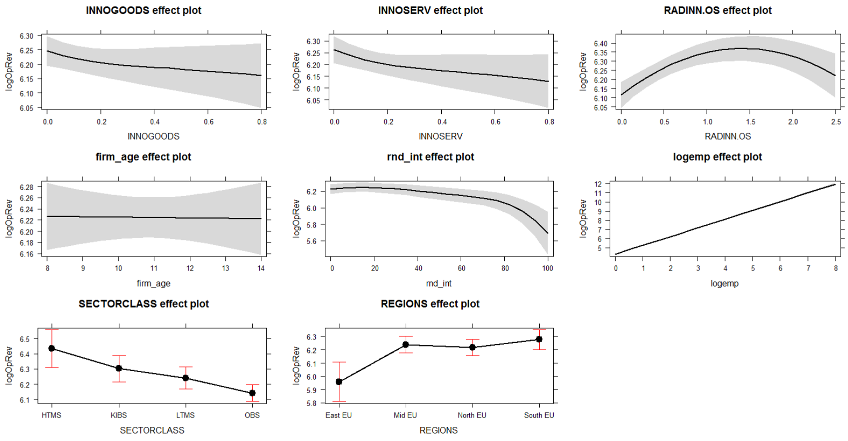


Figure 8.23: Residual plots of 4.2

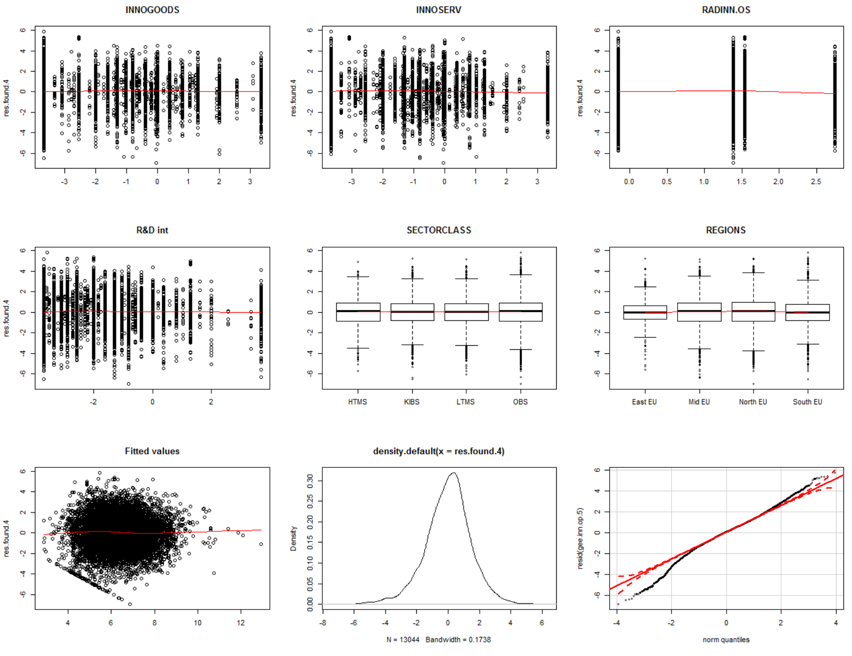


Table 8.11: Model 4.1 and 4.2 GEEs pre-imputation: All specifications

	N.Emp	N.Emp	N.Emp	N.Emp	N.Emp	OpRev	OpRev	OpRev	OpRev	OpRev
logit(InnoGoods)	0.0378^{*}		0.0373^{*}		0.0199	0.0513^{**}		0.0507^{**}		0.0590^{*}
	(0.0163)		(0.0165)		(0.0213)	(0.0191)		(0.0195)		(0.0253)
logit(InnoServ)		0.0085	0.0037		−0.0141		−0.0042	0.0048		0.0133
		(0.0168)	(0.0169)		(0.0225)		(0.0240)	(0.0197)		(0.0244)
RadInnOS				0.0732^{*}	0.0673		0.0530		0.0484	−0.0318
				(0.0305)	(0.0503)		(0.0461)		(0.0373)	(0.0589)
lag(log(NumberEmp + 1), 1)						1.0853^{***}	1.0857^{***}	1.0853^{***}	1.0857^{***}	1.0857^{***}
						(0.0436)	(0.0439)	(0.0436)	(0.0438)	(0.0436)
FirmAge	0.0360^{**}	0.0349^{**}	0.0361^{**}	0.0348^{**}	0.0353^{**}	0.0072	0.0049	0.0072	0.0049	0.0076
	(0.0119)	(0.0119)	(0.0119)	(0.0119)	(0.0118)	(0.0149)	(0.0151)	(0.0149)	(0.0151)	(0.0150)
Spinoff	0.3141^{***}	0.3074^{***}	0.3142^{***}	0.3058^{***}	0.3091^{***}	$0.2010^{\dagger}$	$0.1886^{\dagger}$	0.2012^{*}	$0.1890^{\dagger}$	0.2046^{*}
	(0.0847)	(0.0845)	(0.0847)	(0.0841)	(0.0843)	(0.1026)	(0.1029)	(0.1026)	(0.1030)	(0.1018)
logit(R&DInt)	0.0141	0.0256	0.0133	0.0159	0.0131	−0.0236	−0.0126	−0.0245	−0.0129	−0.0245
	(0.0219)	(0.0226)	(0.0227)	(0.0227)	(0.0228)	(0.0253)	(0.0256)	(0.0256)	(0.0256)	(0.0256)
(Intercept)	2.2836^{***}	2.2941^{***}	2.2903^{***}	2.1640^{***}	2.1503^{***}	3.9549^{***}	3.8574^{***}	3.9644^{***}	3.8726^{***}	4.0299^{***}
	(0.1990)	(0.2010)	(0.2001)	(0.2080)	(0.2328)	(0.2612)	(0.2729)	(0.2636)	(0.2639)	(0.2794)
Sectoral controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Scale parameter: gamma	0.9742	0.9800	0.9744	0.9749	0.9719	1.3674	1.3736	1.3675	1.3736	1.3668
Scale parameter: SE	0.0525	0.0525	0.0525	0.0526	0.0525	0.0827	0.0831	0.0827	0.0831	0.0826
Correlation parameter: alpha	0.9587	0.9593	0.9588	0.9590	0.9584	0.8565	0.8580	0.8566	0.8580	0.8565
Correlation parameter: SE	0.0145	0.0147	0.0146	0.0145	0.0145	0.0196	0.0195	0.0196	0.0196	0.0197
Num. obs.	4758	4758	4758	4758	4758	4337	4337	4337	4337	4337
Num. clust.	1415	1415	1415	1415	1415	1195	1195	1195	1195	1195

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$

Figure 8.24: Schoenfeld residuals plotted against time for 4.3

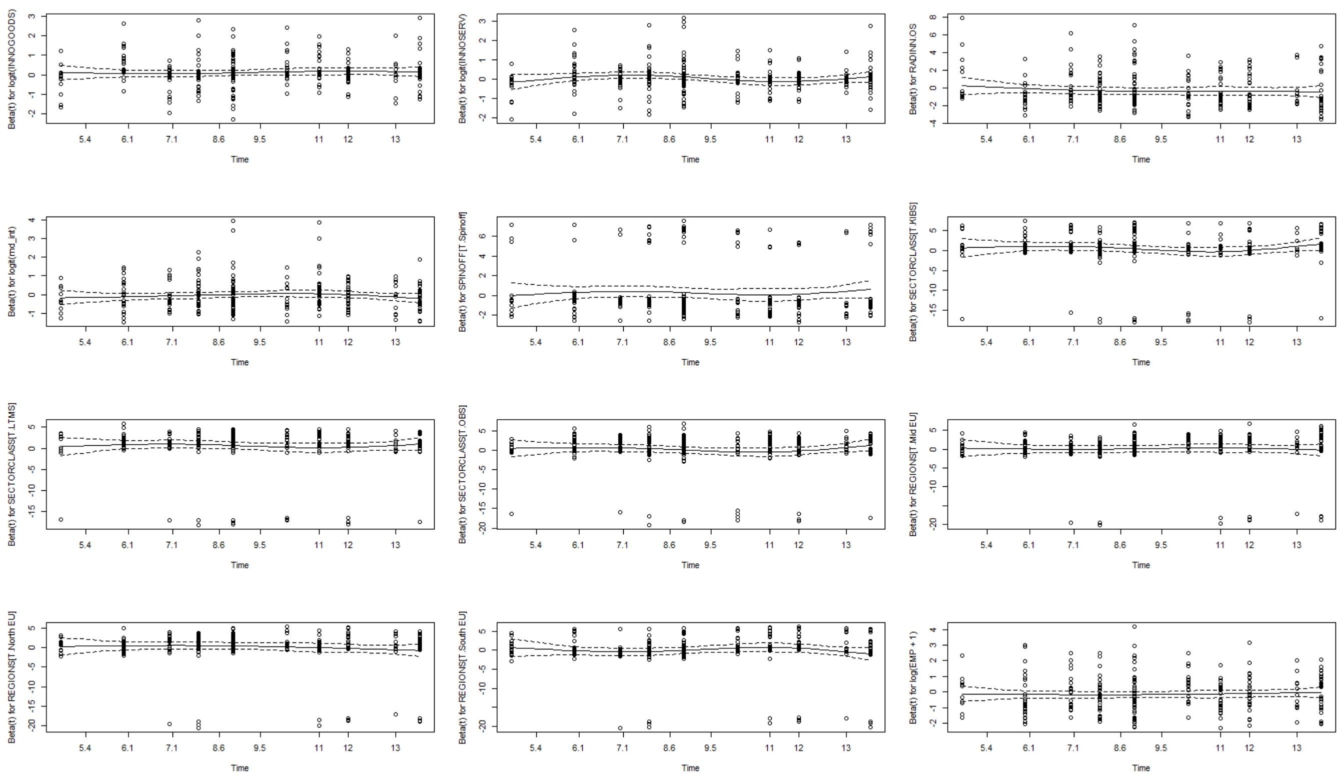


Figure 8.25: Model 4.3 effects plots for Cox PH

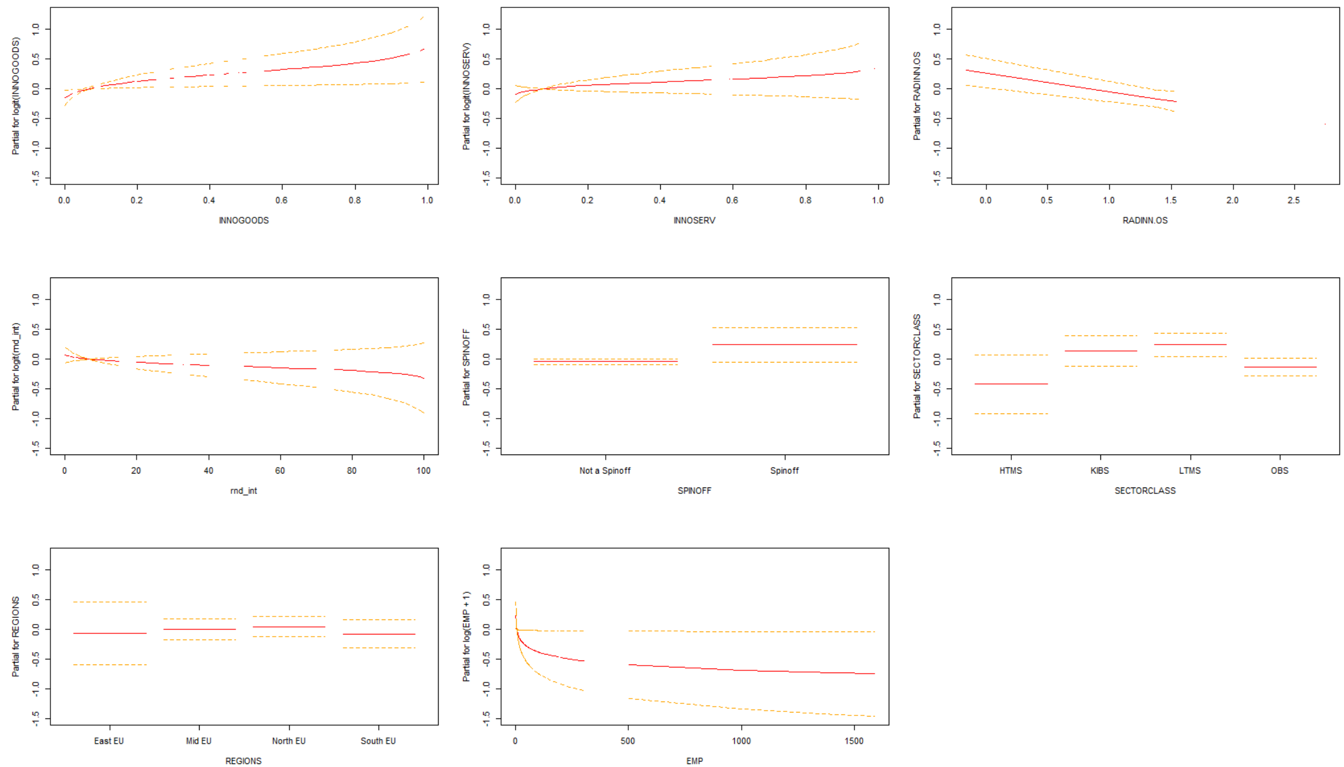


Table 8.12: OLS logged Number of Employee (Size) regressions by year: Pre-imputation robustness checks

	log(Emp) 2010	log(Emp) 2011	log(Emp) 2012	log(Emp) 2013	log(Emp) 2014	log(Emp) 2010	log(Emp) 2011	log(Emp) 2012	log(Emp) 2013	log(Emp) 2014
EmpEdu	0.3945^{*}	0.5907^{***}	0.5522^{**}	0.5669^{**}						
	(0.1631)	(0.1708)	(0.1848)	(0.1982)	(0.2201)					
EmpHiEdu	−0.1046	−0.2530	−0.2777	−0.3195	−0.1184					
	(0.1921)	(0.2002)	(0.2161)	(0.2262)	(0.2648)					
FoundEdu	−0.0164	−0.0218	−0.0298	−0.0132	−0.0154					
	(0.0349)	(0.0363)	(0.0396)	(0.0419)	(0.0473)					
FoundEnt	0.1423^{*}	0.0759	0.0273	0.0568	0.1333					
	(0.0658)	(0.0674)	(0.0734)	(0.0768)	(0.0892)					
FoundUni	0.1432	0.4197^{*}	$0.3731^{\dagger}$	0.2905	−0.0005					
	(0.1778)	(0.1926)	(0.2171)	(0.2215)	(0.2719)					
FoundInd	0.0027	−0.0004	−0.0034	−0.0043	−0.0030					
	(0.0032)	(0.0035)	(0.0037)	(0.0039)	(0.0045)					
AgeMax	0.0683	$0.0850^{\dagger}$	0.1081^{*}	$0.1004^{\dagger}$	0.0579					
	(0.0427)	(0.0455)	(0.0498)	(0.0520)	(0.0598)					
KDisp	-0.7545^{***}	-0.8501^{***}	-0.7739^{***}	-0.7972^{***}	-0.6424^{*}					
	(0.1785)	(0.1853)	(0.2030)	(0.2149)	(0.2523)					
KScope	0.5235^{***}	0.4445^{**}	0.5027^{**}	0.4299^{*}	0.4685^{*}					
	(0.1535)	(0.1626)	(0.1783)	(0.1848)	(0.2198)					
Spinoff	0.4473^{***}	0.3679^{***}	0.3413^{***}	0.3900^{***}	0.3573^{**}	0.4595^{***}	0.4041^{***}	0.3351^{***}	0.3795^{***}	0.3170^{**}
	(0.0841)	(0.0837)	(0.0917)	(0.0953)	(0.1114)	(0.0838)	(0.0841)	(0.0898)	(0.0921)	(0.1080)
FF1	0.0618^{**}	0.0636^{**}	0.0501^{*}	0.0466^{*}	$0.0477^{\dagger}$					
	(0.0191)	(0.0202)	(0.0223)	(0.0237)	(0.0272)					
FF2	0.0235	0.0280	0.0278	0.0277	0.0250					
	(0.0260)	(0.0278)	(0.0305)	(0.0315)	(0.0369)					
FF3	-0.0591^{*}	$-0.0554^{\dagger}$	-0.0938^{**}	-0.0826^{*}	-0.0776^{*}					
	(0.0283)	(0.0288)	(0.0323)	(0.0337)	(0.0390)					
logit(InnoGoods)						0.0005	0.0172	0.0187	0.0394	0.0405
						(0.0219)	(0.0222)	(0.0237)	(0.0248)	(0.0284)
logit(InnoServ)						−0.0319	−0.0357	$-0.0432^{\dagger}$	−0.0301	−0.0216
						(0.0219)	(0.0229)	(0.0244)	(0.0253)	(0.0283)
RadInnOS						$0.0969^{\dagger}$	0.1038^{*}	0.1121^{*}	$0.0972^{\dagger}$	0.0390
						(0.0498)	(0.0516)	(0.0549)	(0.0576)	(0.0641)
logit(R&DInt)						0.0014	0.0003	0.0126	0.0083	$0.0494^{\dagger}$
						(0.0227)	(0.0235)	(0.0249)	(0.0262)	(0.0299)
FirmAge	0.0150	0.0118	0.0311^{*}	0.0228	0.0296	0.0274^{*}	0.0286^{*}	0.0483^{**}	0.0372^{*}	0.0385^{*}
	(0.0139)	(0.0143)	(0.0154)	(0.0163)	(0.0186)	(0.0135)	(0.0139)	(0.0148)	(0.0156)	(0.0178)
R&DInt	−0.0005	−0.0002	0.0023	0.0011	0.0031					
	(0.0017)	(0.0017)	(0.0019)	(0.0020)	(0.0024)					
(Intercept)	2.2367^{***}	2.2779^{***}	2.1573^{***}	2.3463^{***}	2.2557^{***}	2.1913^{***}	2.1303^{***}	2.0590^{***}	2.2688^{***}	2.4615^{***}
	(0.2618)	(0.2686)	(0.3013)	(0.3153)	(0.3637)	(0.2328)	(0.2452)	(0.2615)	(0.2706)	(0.3089)
Sectoral controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	964	979	898	869	668	1054	1067	982	948	731
R^2	0.2760	0.2330	0.2041	0.1677	0.1938	0.2274	0.1894	0.1716	0.1403	0.1606
adj. R^2	0.2599	0.2162	0.1851	0.1471	0.1676	0.2185	0.1802	0.1613	0.1293	0.1466
Resid. sd	0.8805	0.9317	0.9733	1.0062	1.0146	0.9246	0.9768	1.0082	1.0378	1.0361

Standard errors in parentheses

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 8.13: OLS logged Operating Revenue (size) regressions by year: Pre-imputation robustness checks

	log(OR) 2010	log(OR) 2011	log(OR) 2012	log(OR) 2013	log(OR) 2014	log(OR) 2010	log(OR) 2011	log(OR) 2012	log(OR) 2013	log(OR) 2014
EmpEdu	−0.2154	−0.3385	−0.1724	−0.1942	−0.2840					
	(0.2189)	(0.2384)	(0.2427)	(0.2665)	(0.3259)					
EmpHiEdu	0.1930	0.2897	0.1603	0.0310	0.2350					
	(0.2399)	(0.2610)	(0.2644)	(0.2881)	(0.3581)					
FoundEdu	0.0300	−0.0132	0.0664	0.0577	0.0512					
	(0.0408)	(0.0447)	(0.0449)	(0.0494)	(0.0636)					
FoundEnt	0.2083**	0.1982*	0.1523†	0.2213*	0.3241**					
	(0.0801)	(0.0872)	(0.0885)	(0.0967)	(0.1240)					
FoundUni	−0.1646	0.0046	−0.1270	−0.4688†	−0.9469**					
	(0.2120)	(0.2410)	(0.2390)	(0.2584)	(0.3228)					
FoundInd	0.0012	0.0002	−0.0015	−0.0022	0.0023					
	(0.0039)	(0.0043)	(0.0043)	(0.0048)	(0.0060)					
AgeMax	0.0460	−0.0281	0.0289	0.0084	−0.0555					
	(0.0513)	(0.0563)	(0.0568)	(0.0628)	(0.0805)					
KDisp	−0.4063†	−0.3792	−0.1737	−0.1923	−0.0953					
	(0.2129)	(0.2356)	(0.2389)	(0.2624)	(0.3328)					
KScope	0.0978	0.0877	−0.0778	−0.1481	0.0367					
	(0.1795)	(0.1952)	(0.1984)	(0.2177)	(0.2823)					
Spinoff	0.2614*	0.2218*	0.1921†	0.2174†	0.4149**	0.2825**	0.2968**	0.2246*	0.2110†	0.4712**
	(0.1018)	(0.1110)	(0.1124)	(0.1244)	(0.1588)	(0.0985)	(0.1059)	(0.1077)	(0.1187)	(0.1516)
FF1	0.0269	0.0132	0.0101	0.0347	0.0160					
	(0.0228)	(0.0250)	(0.0255)	(0.0280)	(0.0351)					
FF2	0.0832**	0.0920**	0.0752*	0.0337	0.0702					
	(0.0311)	(0.0335)	(0.0336)	(0.0373)	(0.0465)					
FF3	−0.0166	−0.0456	−0.0752*	−0.0840*	−0.0661					
	(0.0329)	(0.0358)	(0.0367)	(0.0401)	(0.0510)					
logit(InnoGoods)						0.0588*	0.0675*	0.0189	−0.0017	0.0093
						(0.0263)	(0.0280)	(0.0285)	(0.0313)	(0.0395)
logit(InnoServ)						0.0150	0.0210	−0.0155	0.0060	0.0428
						(0.0261)	(0.0281)	(0.0285)	(0.0315)	(0.0397)
RadInnOS						−0.0714	−0.0751	0.0432	0.0396	−0.0024
						(0.0629)	(0.0675)	(0.0684)	(0.0746)	(0.0920)
logit(R&DInt)						−0.0372	−0.0625*	−0.0414	−0.1123***	−0.1154**
						(0.0268)	(0.0285)	(0.0286)	(0.0315)	(0.0382)
log(Emp)	0.9880***	1.0663***	1.0995***	1.1497***	1.1475***	0.9907***	1.0354***	1.0905***	1.1317***	1.0987***
	(0.0361)	(0.0386)	(0.0389)	(0.0424)	(0.0549)	(0.0333)	(0.0349)	(0.0350)	(0.0385)	(0.0497)
FirmAge	0.0144	0.0250	0.0406*	0.0266	0.0340	0.0243	0.0302†	0.0370*	0.0280	0.0326
	(0.0164)	(0.0179)	(0.0181)	(0.0199)	(0.0248)	(0.0159)	(0.0169)	(0.0171)	(0.0189)	(0.0237)
R&DInt	−0.0030	−0.0045*	−0.0037†	−0.0082***	−0.0067*					
	(0.0020)	(0.0022)	(0.0022)	(0.0024)	(0.0030)					
(Intercept)	3.8804***	3.9665***	3.1285***	3.4778***	3.2290***	4.0063***	3.8213***	3.2813***	3.2932***	3.1537***
	(0.3337)	(0.3640)	(0.3661)	(0.3988)	(0.5119)	(0.2991)	(0.3194)	(0.3185)	(0.3497)	(0.4484)
Sectoral controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Regional controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	1474	1435	1383	1277	929	1575	1530	1474	1364	984
R^2	0.4296	0.4219	0.4499	0.4434	0.4089	0.4212	0.4219	0.4560	0.4467	0.3984
adj. R^2	0.4209	0.4129	0.4410	0.4336	0.3946	0.4164	0.4170	0.4511	0.4414	0.3903
Resid. sd	1.3197	1.4195	1.4106	1.4860	1.6071	1.3485	1.4179	1.4069	1.4890	1.6028

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 8.14: Model 3.1b - Summary growth rates of logged Number of Employees (slope) regressed on internal knowledge intensity - OLS pre-imputation

	I 2010- 2014	II 2010- 2014	III 2010- 2014	IV 2010 2014	V 2010- 2014	VI 2010- 2014
EmpEdu	−0.1360					−0.1419
	(0.1066)					(0.1069)
EmpHiEdu	0.0118					0.0557
	(0.1311)					(0.1305)
FoundEdu	0.0159					0.0311
	(0.0234)					(0.0241)
FoundEnt		0.0424				0.0575
		(0.0425)				(0.0448)
FoundUni		−0.0286				−0.0587
		(0.1203)				(0.1270)
FoundInd		−0.0017				−0.0020
		(0.0020)				(0.0022)
AgeMax		−0.0126				−0.0260
		(0.0280)				(0.0297)
KDisp			−0.0697			−0.0002
			(0.1128)			(0.1247)
KScope			−0.0480			−0.1293
			(0.1009)			(0.1122)
Spinoff				−0.0175		0.0032
				(0.0565)		(0.0573)
FF1					0.0001	0.0045
					(0.0119)	(0.0133)
FF2					−0.0126	0.0008
					(0.0161)	(0.0185)
FF3					−0.0061	−0.0014
					(0.0172)	(0.0196)
FirmAge	−0.0026	−0.0013	−0.0030	−0.0060	−0.0004	0.0029
	(0.0092)	(0.0088)	(0.0087)	(0.0090)	(0.0084)	(0.0094)
R&DInt	0.0008	0.0002	0.0004	0.0007	0.0003	0.0008
	(0.0011)	(0.0011)	(0.0010)	(0.0011)	(0.0011)	(0.0012)
(Intercept)	−0.0192	0.0547	0.0749	0.0942	0.0316	0.0026
	(0.1588)	(0.1514)	(0.1445)	(0.1385)	(0.1304)	(0.1825)
Sectoral Controls	Yes	Yes	Yes	Yes	Yes	Yes
Regional Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	564	607	619	630	607	536
R^2	0.0086	0.0088	0.0068	0.0050	0.0067	0.0215
adj. R^2	−0.0111	−0.0112	−0.0095	−0.0094	−0.0117	−0.0185
Resid. sd	0.4649	0.4609	0.4608	0.4863	0.4457	0.4499

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$
Response variables take the form $log(X_{it}) - log(X_{i,t-1})$

Table 8.15: Model 3.2b - Summary growth rates of logged Operating Revenue (slope) regressed on internal knowledge intensity - OLS pre-imputation

	I 2010-2014	II 2010-2014	III 2010-2014	IV 2010 2014	V 2010-2014	VI 2010-2014
EmpEdu	0.0663					0.0826
	(0.2396)					(0.2522)
EmpHiEdu	0.1160					0.1379
	(0.2639)					(0.2735)
FoundEdu	−0.0119					0.0108
	(0.0450)					(0.0490)
FoundEnt		0.0507				0.0360
		(0.0867)				(0.0946)
FoundUni		$-0.4253^{\dagger}$				-0.5625^{*}
		(0.2228)				(0.2440)
FoundInd		−0.0017				−0.0013
		(0.0040)				(0.0046)
AgeMax		−0.0325				−0.0356
		(0.0564)				(0.0620)
KDisp			0.2077			0.2276
			(0.2292)			(0.2556)
KScope			−0.0307			−0.0643
			(0.1902)			(0.2190)
Spinoff				0.0768		0.1139
				(0.1086)		(0.1223)
FF1					0.0254	0.0324
					(0.0244)	(0.0268)
FF2					0.0371	0.0306
					(0.0320)	(0.0361)
FF3					−0.0029	−0.0138
					(0.0362)	(0.0395)
FirmAge	−0.0019	−0.0081	−0.0066	−0.0068	−0.0097	−0.0035
	(0.0180)	(0.0176)	(0.0172)	(0.0169)	(0.0175)	(0.0191)
R&DInt	−0.0026	−0.0028	$-0.0034^{\dagger}$	−0.0032	$-0.0036^{\dagger}$	−0.0024
	(0.0022)	(0.0021)	(0.0020)	(0.0020)	(0.0021)	(0.0024)
log(Emp)	$0.0777^{\dagger}$	0.0966^{*}	0.0885^{*}	0.0834^{*}	0.0785^{*}	$0.0770^{\dagger}$
	(0.0413)	(0.0377)	(0.0372)	(0.0366)	(0.0380)	(0.0441)
(Intercept)	−0.4778	−0.3067	−0.5212	−0.3949	−0.3494	−0.4922
	(0.3363)	(0.3272)	(0.3178)	(0.2909)	(0.3013)	(0.3975)
Sectoral Controls	Yes	Yes	Yes	Yes	Yes	Yes
Regional Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	933	975	995	1007	960	878
R^2	0.0221	0.0298	0.0249	0.0237	0.0256	0.0354
adj. R^2	0.0093	0.0167	0.0140	0.0139	0.0133	0.0106
Resid. sd	1.1872	1.1717	1.1686	1.1623	1.1775	1.2025

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$
Response variables take the form $log(X_{it}) - log(X_{i,t-1})$

Table 8.16: Model 4.1b - Summary growth rates of logged Number of Employees (slope) regressed on innovative performance - OLS pre-imputation

	I 2010- 2014	II 2010- 2014	III 2010- 2014	IV 2010- 2014	V 2010- 2014
logit(InnoGoods)	$0.0207^{\dagger}$		0.0166		0.0060
	(0.0119)		(0.0123)		(0.0153)
logit(InnoServ)		0.0375^{**}	0.0321^{*}		0.0222
		(0.0118)	(0.0124)		(0.0151)
RadInnOS				0.0719^{**}	0.0395
				(0.0218)	(0.0339)
FirmAge	−0.0006	−0.0047	−0.0010	−0.0072	−0.0024
	(0.0094)	(0.0092)	(0.0095)	(0.0089)	(0.0096)
Spinoff	−0.0123	−0.0222	−0.0158	−0.0223	−0.0206
	(0.0580)	(0.0573)	(0.0585)	(0.0560)	(0.0587)
logit(R&DInt)	0.0022	−0.0028	−0.0058	−0.0014	−0.0069
	(0.0152)	(0.0151)	(0.0157)	(0.0149)	(0.0157)
(Intercept)	0.0637	0.1691	0.1286	0.0228	0.0545
	(0.1448)	(0.1468)	(0.1494)	(0.1439)	(0.1623)
Sectoral Controls	Yes	Yes	Yes	Yes	Yes
Regional Controls	Yes	Yes	Yes	Yes	Yes
N	593	602	582	630	582
R^2	0.0113	0.0221	0.0227	0.0225	0.0250
adj. R^2	−0.0057	0.0055	0.0038	0.0067	0.0044
Resid. sd	0.4866	0.4867	0.4884	0.4825	0.4883

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$
Response variables take the form $log(X_{it}) - log(X_{i,t-1})$

Table 8.17: Model 4.2b - Summary growth rates of logged Operating Revenue (slope) regressed on innovative performance - OLS pre-imputation

	I 2010- 2014	II 2010- 2014	III 2010- 2014	IV 2010 2014	V 2010- 2014
logit(InnoGoods)	−0.0261		−0.0283		-0.0677^{*}
	(0.0227)		(0.0235)		(0.0296)
logit(InnoServ)		$0.0395^{\dagger}$	0.0530^{*}		0.0133
		(0.0225)	(0.0236)		(0.0297)
RadInnOS				$0.0813^{\dagger}$	0.1514^{*}
				(0.0422)	(0.0693)
FirmAge	−0.0034	−0.0055	−0.0037	−0.0055	−0.0060
	(0.0176)	(0.0172)	(0.0176)	(0.0169)	(0.0176)
Spinoff	0.1143	0.0886	0.1061	0.0711	0.0987
	(0.1130)	(0.1111)	(0.1129)	(0.1083)	(0.1127)
logit(R&DInt)	$-0.0525^{\dagger}$	-0.0599^{*}	$-0.0536^{\dagger}$	-0.0711^{*}	$-0.0561^{\dagger}$
	(0.0286)	(0.0283)	(0.0292)	(0.0278)	(0.0292)
log(Emp)	0.0829^{*}	0.0879^{*}	0.0841^{*}	0.0780^{*}	0.0762^{*}
	(0.0382)	(0.0376)	(0.0383)	(0.0368)	(0.0384)
(Intercept)	-0.6097^{*}	−0.5020	−0.5102	-0.6934^{*}	-0.8072^{*}
	(0.3084)	(0.3077)	(0.3115)	(0.3041)	(0.3392)
Sectoral Controls	Yes	Yes	Yes	Yes	Yes
Regional Controls	Yes	Yes	Yes	Yes	Yes
N	953	959	934	1007	1007
R^2	0.0271	0.0323	0.0336	0.0294	0.0294
adj. R^2	0.0157	0.0211	0.0210	0.0186	0.0186
Resid. sd	1.1713	1.1593	1.1640	1.1596	1.1596

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$
Response variables take the form $log(X_{it}) - log(X_{i,t-1})$

8.4 AEGIS Survey Questionnaire

The following pages show the complete questionnaire used in the AEGIS survey, and is taken from AEGIS Research Project (2013).

APPENDIX

AEGIS QUESTIONNAIRE

Introduction

We are contacting you in the context of a survey funded by the European Commission that is carried out in 10 countries across Europe.

The aim of the survey is to investigate new companies that incorporate knowledge and have significant innovative activity. The results of this survey, which will be communicated to your firm, will provide valuable recommendations for shaping EU policies in this field.

I would like to speak with one of the founders and ask some general questions about your company. This interview will take around 15 minutes.

Please note that the information you provide will not be used at an individual level nor handed over by name to the European Commission or any other third party. The information will only be used for aggregate analysis.

Screener Questions

S1

S1. We are looking for new firms that were established during the period 2001-2007, is it correct that your firm is established under the current legal status in <Start year >?

Items

Code	Description
1	Yes
2	No

S2

Condition S1 = 2

S2. In which year is your firm established?

S3

S3. Was this a new establishment or just a change in the legal status?

Code	Description
1	Yes, new establishment
2	No, change of legal status

S4a. Was the company in its current legal form established as a spin off of an established company with more than 25% ownership?

Code	Description
1	Yes
2	No

S4b. A subsidiary of another company?

Code	Description
1	Yes
2	No

S4c. A merger, acquisition or joint venture?

Code	Description
1	Yes
2	No

S5. Please indicate what are the firm's most important activities?

SECTION 1: General information about the firm

Q1. What is the total number of ...

Code	Description
1	Full time employees in your company
2	Part time employees in your company

Q2. What is the total number of employees in your firm with a University Degree?

Code	Description
1	Employees with an University Degree

Q2B

Q2b. How many of them hold a:

Code	Description
1	Postgraduate degree

2	PhD

SECTION 2a: General information about the founder or the founding team

Q3. How many people founded your firm?

Code	Description
1	Number of people

Q4. Who founded your firm?

Q4A

Items

Code	Description
1	Founder 1
2	Founder 2
3	Founder 3
4	Founder 4

Labels

Code	Description
1	Mr
2	Mrs
3	Ms

Q5. What is/are the highest educational attainment of the founder(s)?

Items

Code	Description
1	Founder 1:
2	Founder 2:
3	Founder 3:
4	Founder 4:

Labels

Code	Description
1	Elementary education
2	Secondary education
3	Bachelor degree
4	Postgraduate degree
5	PhD
6	Don't know

Q6. What was the last occupation of the founder(s) before the establishment of this company?

Items

Code	Description

	Code	Description
	1	Founder 1:
	2	Founder 2:
	3	Founder 3:
	4	Founder 4:
Labels	Code	Description
	1	Owner of a firm still in existence
	2	Owner of a firm that has ceased operations
	3	Employee of a firm in the same industry
	4	Employee of a firm in a different industry
	5	Self-employed
	6	University or research institute employee
	7	Government employee
	8	Unemployed
	9	None of the above - this is his/her first job
	10	Don't know

Q7. Approximately how many years of professional experience did the founder(s) have in the current sector your company is active before the establishment of this company?

Items	Code	Description
	1	Founder 1:
	2	Founder 2:
	3	Founder 3:
	4	Founder 4:

Q8. What are the main areas of expertise of the founder(s) that are relevant for the operation of this company?

Items	Code	Description
	1	Founder 1:
	2	Founder 2:
	3	Founder 3:
	4	Founder 4:
Labels	Code	Description
	1	Technical and engineering knowledge
	2	General management
	3	Product design
	4	Marketing
	5	Finance
	6	None of these / Don't know

Q9. What is the age of the founder(s)?

Items		
	1	Founder 1:
	2	Founder 2:

	Code	Description
	3	Founder 3:
	4	Founder 4:
Labels	Code	Description
	1	18-29
	2	30-39
	3	40-49
	4	>50
	5	Don't know

Q10. Were/was the founder(s) born in this country?

	Code	Description
Items		
	1	Founder 1:
	2	Founder 2:
	3	Founder 3:
	4	Founder 4:
Labels	Code	Description
	1	Yes
	2	No
	3	Don't know

SECTION 2b:Formation process

Q11. Did the company come out of another pre-existing organization?

	Code	Description
Items		
	1	Yes
	2	No

Q12a

Condition Q11 = 1

Q12a. What is the parent organization?

	Code	Description
Items		
	1	University
	2	Company
	3	Other, specify

Q12b

Condition (Q11 = 1) and (Q12a = 2)

Q12b. Is this company still related to the firm as a:

Items Code Description

1	Partner
2	Competitor
3	Customer
4	Supplier
5	None of these

Q13. Please indicate the importance of the following factors for the formation of the company on a 5 point scale, were 1 is not important and 5 is extremely important.

Items

Code	Description
1	Work experience in the current activity field
2	Technical/engineering knowledge in the field
3	Design knowledge
4	Knowledge of the market
5	Networks built during previous career
6	Availability of finance
7	Opportunities in a public procurement initiative
8	Existence of a large enough customer
9	Opportunity deriving from technological change
10	Opportunity deriving from a new market need
11	Opportunity deriving from new regulations or institutional requirements

---

Q14. Please, estimate the percentage of funding coming from the following sources for setting up your company.

Items

Code	Description
1	Own financial resources (savings)
	| %
2	Funding from family member
	| %
3	Funding from previous employer (corporate venture capital, university incubator technology transfer)
	| %
4	Venture capital
	| %
5	Funding from a bank
	| %
6	Public fund from national government or local authorities (programs supporting entrepreneurship, etc)
	| %

7	European Union funds (programs supporting SMEs etc) | %
8	Other sources (please specify) | %

SECTION 3: Market environment

Competitive and institutional environment

Q15. Right now, are there other businesses offering the same products and/or services to your potential? customers?

Items	Code	Description
	1	Yes, many business competitors
	2	Only a few business competitors
	3	No other business competitors

Q16. During the last three years (2007-2009) what was the % of your firm's sales in :

Items	Code	Description
	1	The local/regional market
	2	The national market
	3	The international market

Q17. Please identify the most important type of customer of the company.

Items	Code	Description
	1	Large firms
	2	Small and medium sized firms
	3	Final consumers (e.g. private households, private consumption)
	4	Public sector
	5	Other (please specify)

Q18. Please indicate to what extent you agree or disagree with the following statements characterizing your business environment. On a 5 point scale, were 1 is completely disagree and 5 is completely agree.

In the principal industry in which our firm operates...

Items	Code	Description
	1	... the life cycle of products is typically short
	2	... customers regularly ask for new products and/or services
	3	... the speed of technological changes is high
	4	... the activities of our major competitors are unpredictable and

	competition is very intense
5	... a company only succeeds if it is able to launch new products/services continuously
6	... price competition is prevalent
7	... quality competition is prevailing

Success factors

Q19. Please, indicate the contribution of the following factors in creating and sustaining the competitive advantage of this company. On a 5 point scale, were 1 is no impact and 5 huge impact.

Code	Description
1	Capability to offer novel products/services
2	Capacity to adapt the products/services to the specific needs of different customers/market niches
3	Capability to offer expected products/services at low cost
4	R&D activities
5	Establishment of alliances/partnerships with other firms
6	Capability to offer high quality product/services at a premium price
7	Networking with scientific research organizations (universities, institutes, etc.)
8	Marketing and promotion activities

Obstacles

Q20. Please indicate to what extent the following factors have been obstacles to the firm growth and expansion of business activities. On a 5 point scale, were 1 is not at all and 5 is to a great extent.

Items	Code	Description
	1	Technology risk / uncertainty
	2	Market risk /uncertainty
	3	Large initial investment
	4	Difficulty in finding the necessary funding for growth investments
	5	Difficulty in finding business partners
	6	Difficulties in recruiting highly-skilled employees
	7	Lack of technological know-how

Q21. Please indicate how serious the following barriers have been to the firm growth and expansion of business activities. On a 5 point scale, were 1 is no barriers and 5 is very serious barriers.

Items	Code	Description
	1	Continuously changing taxation regulations
	2	High tax rates
	3	Time consuming regulatory requirements for issuing permits and licenses

	4	Poorly enforced competition law to curb monopolistic practices
	5	Poorly enforced property rights, copyright and patent protection
	6	Strict property, copyright and patent protection
	7	Government officials favor well connected individuals
	8	Bankruptcy legislation makes immense the cost of failure
	9	Rigid labor market legislation

SECTION 4: Strategy

Identification and utilization of technical and market opportunities

Q22. What is the main strategy of the company?

Items	Code	Description
	1	Offer standardized products and services at low cost (cost leadership strategy)
	2	Offer unique products and services (differentiation strategy).
	3	Exploit opportunities in new market niches (focus strategy).
	4	Other, specify

Q23. Please indicate to what extent you agree or disagree with the following statements regarding the sensing and seizing of opportunities within your firm. On a 5 point scale, were 1 is strongly disagree and 5 is strongly agree.

Items	Code	Description
	1	Our firm actively observes and adopts the best practices in our sector
	2	Our firm responds rapidly to competitive moves
	3	We change our practices based on customer feedback
	4	Our firm regularly considers the consequences of changing market demand in terms of new products and services
	5	Our firm is quick to recognize shifts in our market (e.g. competition, regulation, demography)
	6	We quickly understand new opportunities to better serve our customers
	7	There is a formal R&D department in our firm
	8	There is a formal engineering and technical studies department in our firm
	9	Design activity is important in introducing new products/services to the market
	10	We implement systematic internal and external personnel training
	11	Employees share practical experiences on a frequent basis

---

Sources of knowledge

Q24. Please evaluate the importance of the following sources of knowledge for exploring new business opportunities on a 5 point scale, were 1 is not important and 5 is extremely important.

Items	Code	Description
	1	Clients or customers
	2	Suppliers
	3	Competitors

4	Public research institutes
5	Universities
6	External commercial labs/R&D firms/technical institutes
7	In-house (know how, R&D laboratories in your firm)
8	Trade fairs, conferences and exhibitions
9	Scientific journals and other trade or technical publications
10	Participation in nationally funded research programmes
11	Participation in EU funded research programmes (Framework Programmes)

Networking

Q25. To what extent do the networks your firm participates in have contributed to the following operations of the company? On a 5 point scale, were 1 is not important and 5 is extremely important.

Items

Code	Description
1	Contacting customers/clients
2	Selecting suppliers
3	Recruiting skilled labor
4	Collecting information about competitors
5	Accessing distribution channels
6	Assistance in obtaining business loans/attracting funds
7	Advertising and promotion
8	Developing new products/services
9	Managing production and operations
10	Assistance in arranging taxation or other legal issues
11	Exploring export opportunities

Q26. Please indicate to what extent your company has participated in the following types of agreements? On a 5 point scale, were 1 is not at all and 5 is very often.

Items

Code	Description
1	Strategic alliance
2	R&D agreement
3	Technical cooperation agreement
4	Licensing agreement
5	Subcontracting
6	Marketing/export promotion
7	Research contract-out
8	Other (please specify)

SECTION 5: Innovation and business models

Q27a. Did this company introduce new or significantly improved goods or services during the past three years? (Exclude the simple resale of new products purchased from other enterprises and changes of solely aesthetic nature).

Items	Code	Description
	1	Yes
	2	No

Q27b. Please estimate:

Items	Code	Description
	1	The share of new or significantly improved goods to total sales
	2	The share of new or significantly improved services to total sales

Multiple response

Q28. The new or significantly improved goods or services were ...

Items	Code	Description
	1	New to the firm
	2	New to the market
	3	New to the world

Q29. Please indicate to what extent has the firm introduced new or significantly improved goods or services as a result of participation in a publicly supported or subsidised activity? On a 5 point scale, where 1 is not at all and 5 to a great extent, is that ...

Labels	Code	Description
	1	Not at all
	2	2
	3	3
	4	4
	5	To a great extent
	6	Don't know

Q30. Please indicate which of the following methods were used by your firm to protect its intellectual property during the last three years.

Items	Code	Description
	1	Patents
	2	Trademarks
	3	Copyrights
	4	Confidentiality agreements
	5	Secrecy
	6	Lead-time advantages on competitiors
	7	Complexity of design

Labels	Code	Description
	1	Yes

2	No
3	Don't know

Q31 During the last three years the company has introduced new or significantly improved ...

Items

Code	Description
1	Methods of manufacturing
2	Logistics, supply chain, delivery or distribution methods for its inputs, goods or services
3	Supporting activities for your processes, such as maintenance systems or operations for purchasing, accounting, or computing
4	Improved knowledge management systems
5	Changes in the managing structure

Labels

Code	Description
1	Yes
2	No
3	Don't know

Q32 On average, which percentage of your sales has been spent on R&D during the last three years?

Items

Code	Description
1	which percentage of sales is spend on R&D? ….. %

SECTION 6: Firm performance and the effect of economic crisis

Q33. Please estimate the average increase/decrease of ...

Q33A

2007 - 2009

Items

Code	Description
1	sales | %
2	employment | %
3	exports | %
4	R&D to sales ratio | %

Q33B

End of 2010

Items	Code	Description
	1	sales
		| %
	2	employment
		| %
	3	exports
		| %
	4	R&D to sales ratio

Q34. Could you please indicate the impact (if any) of the current economic crisis on your firm in terms of the following elements (please consider the effect on the activity of 2009 compared to the activity of 2008).

Items	Code	Description
	1	Sales
	2	Exports
	3	Employment
	4	Profits
	5	Investments

Labels	Code	Description
	1	Significant increase (>5%)
	2	No significant changes (+/- 5%)
	3	Slight Decrease (-5%to -10%)
	4	Significant Decrease (-10% to -20%)
	5	Very Significant Decrease (>-20%)
	6	Don't know

Q35. How do you think your firm /sector will be affected in terms of financing and creation of new opportunities in the post crisis period?

Items	Code	Description
	1	Liquidity will be significantly restricted in my sector
	2	Borrowing costs will significantly increase
	3	A lot of my customers / suppliers will face significant liquidity problems which may cause problems to my firm
	4	Bankruptcies and restructuring in my sector might create new opportunities for my firm

Labels	Code	Description
	1	Yes
	2	No
	3	Don't know

Q36. Please indicate the average turnover of your firm during the last three years (2007-2009)

Items

Code	Description
1	upto 200 thousand pounds
2	200-400 thousand pounds
3	401-1700 thousand pounds
4	1701-4000 thousand pounds
5	4001-8500 thousand pounds
6	8501-40000 thousand pounds
7	More than 40000 thousand pounds
8	Don't know
9	Refused

Q37. Please indicate average profits of your firm during the last three years (2007-2009)

Items

Code	Description
1	Losses
2	upto 40 thousand pounds
3	41 to 130 thousand pounds
4	131 to 170 thousand pounds
5	171 to 450 thousand pounds
6	451 to 850 thousand pounds
7	850 thousand to 4 Million pounds
8	More than 4 Million pounds
9	Don't know
10	Refused

Thank you very much for your time and cooperation.