

Early-Stage Entrepreneurship and Regional Economic Growth in the European Union

Mateo Ivan Radman Student ID: 6855180 UTRECHT UNIVERSITY

Early-Stage Entrepreneurship and Regional Economic Growth in the European Union

Abstract

Entrepreneurship is often viewed as a driver of the global economy. However, sometimes the previous research on the relationship between entrepreneurship and economic growth shows contradictory results depending on the research settings. Therefore, this thesis is set to investigate how the early-stage entrepreneurship – including only enterprises that are less than 3 and a half years old – affects regional economic growth in the European Union. The panel sample included 273 NUTS 2 regions between 2008 and 2017. The analysis included three methods: bivariate correlation, fixed effects regression with region and time fixed effects, and spatial fixed effects regression. The results support the hypothesis of this research and show that the early-stage entrepreneurship positively affects the economic growth of European regions. However, the potential bidirectional nature of this relationship obliterates the ability to comment on the causality of this link.

Keywords: Entrepreneurship, Economic Growth, EU NUTS2 Regions, Economic Geography, Spatial Analysis.

Table of Contents

ABSTRACT	2
1. INTRODUCTION	4
2. THEORETICAL FRAMEWORK	7
 2.1. THE DEVELOPMENT OF THE ECONOMIC LITERATURE ON ENTREPRENEURSHIP 2.2. WHAT DETERMINES A LEVEL OF REGIONAL ENTREPRENEURSHIP? 2.3. THE INTERPLAY BETWEEN ENTREPRENEURSHIP AND ECONOMIC GROWTH 2.4. PHASES OF ENTREPRENEURSHIP AND EARLY-STAGE ENTREPRENEURSHIP 2.5. HOW DOES EARLY-STAGE ENTREPRENEURSHIP AFFECT REGIONAL ECONOMIC GROWTH? 2.6. CONCEPTUAL MODEL 	7 11 13 15 16 18
3. METHODOLOGY AND DATA	20
3.1. METHODOLOGY3.2. DATA AND VARIABLES3.3. SPACE DIMENSION	20 24 29
4. EMPIRICAL RESULTS	36
5. DISCUSSION AND RESEARCH LIMITATIONS	42
5.1. DISCUSSION OF THE RESULTS	42 45
6. CONCLUSION	46
BIBLIOGRAPHY	47
APPENDIX A	56
APPENDIX B	57
APPENDIX C	58

1. Introduction

Fifty years ago, very few economic policymakers and think-tanks recognised entrepreneurship as an accelerator of the economy. Today, entrepreneurship and small and medium-sized enterprises (SMEs) suddenly stand at the forefront of new economic policies around the globe. For instance, in the aftermath of the global economic crisis of 2008, the European Union experienced one of the worst economic performances in its history. The economy was on a brink to collapse which severely wounded SMEs and left 25 million people without a job. Unfortunately, a large number of those enterprises had to be shut, and many of those that managed to survive still attempt to reach the pre-crisis levels. Therefore, the European Commission decided to amend its economic policy and 'reignite the entrepreneurial spirit in Europe' by introducing a Small Business Act review and the Action Plan 2020 which are expected to support entrepreneurs around the EU. The European Commission regards entrepreneurship as a powerful driver of growth and employment which makes the European economy more competitive and resilient to external economic shocks. According to Eurostat (2019), SMEs create four million jobs annually in the European Union and account for the biggest employer. By placing such importance on promoting entrepreneurial activity and the production of knowledge in the past three decades politicians, policymakers, and scientists contributed to the creation of a large body of research on this topic available today. Generally, entrepreneurship acquired a great image and it is viewed as an accelerator of the economy (Acs et al., 2012; Acs et al., 2008; Aparicio et al., 2016). According to Joseph Stiglitz, a Nobel Prize winner in economics, the popularity of entrepreneurship in the modern politics and academia can be attributed to the likes of Apple and Microsoft in the United States which have unexpectedly grown from 'basement' projects to multi-billion-dollar companies within a decade. Similarly, Bresnahan and Gambardella (2004) claim that every region in the world wants to create the next Silicon Valley which is a role model of entrepreneurial success and a conduit of economic growth for the United States.

However, many explanations of why entrepreneurship leads to economic growth do not hold because they fail to take into account that the entrepreneurial activity does not come in isolation. Instead, the context greatly influences the level of entrepreneurship, type of firms, and their potential outcome. This effect is much more evident when we zoom on the regional level as a unit of observation since countries are very large and heterogenous. Hence, this thesis is going to investigate the relationship between entrepreneurship and economic growth on the NUTS 2 level of the European Union taking into account differences in the regional context.

Since entrepreneurship is a multi-disciplinary term studied from many different fields such as economics, entrepreneurship, economic geography, politics, management and many others, it is important to outline the definition of entrepreneurship that this thesis assumes. Business dictionary (2020) defines entrepreneurship as the "capacity and willingness to develop, organize and manage a business venture along with any of its risks to make a profit." The people who form these new enterprises are called entrepreneurs. Generally, entrepreneurs are considered to be entities which are endowed with 'special' skills and willingness to take a risk. Audretsch et al. (2002) define those skills as the ability "to recognize the commercial

potential of the invention and organize the capital, talent, and other resources that turn an invention into a commercially viable innovation."

The purpose of this study is to explore the relationship between entrepreneurship and economic growth on the regional level in Europe. The ultimate goal is to find evidence to support the hypothesis that the early-stage entrepreneurship (enterprises younger than 3.5 years of age) causes regional economic growth in the European Union. The analysis is based on the sample of 273 NUTS 2 regions between 2008 and 2017. In order to investigate whether the early-stage entrepreneurship causes any change in the levels of economic growth across NUTS 2 regions in the EU, this thesis is structured as follows. The following two sections will outline scientific and societal relevance. Chapter 2 contains a literature review, theoretical framework, conceptual model, hypothesis and research questions of this thesis. Chapter 3 discusses methodology, data collection strategy and the results of the exploratory spatial data analysis. Furthermore, Chapter 4 presents the results from three different methods of statistical analysis. Next, chapter 5 consists of two sections. The first section discusses the results presented in chapter 4 in greater depth, and I attempt to answer the research guestions of this thesis. Then, the second section of chapter 5 deals with the possible limitations of this study where I outline some problems that I have encountered during data collection and analysis, and I offer some solutions to those problems and finally, in section 7 I conclude the thesis.

Scientific Relevance

Today, one can find numerous studies which investigate the effect of entrepreneurship on the economic performance of neighbourhoods, regions, nations, and even continents (Wennekers and Thurik, 1999; Szerb et al., 2013; Mitra, 2020). While there is a welldocumented relationship between entrepreneurship and economic growth (OECD SME and Entrepreneurship Outlook, 2019), the magnitude of this effect differs from country to country and from region to region. For instance, the United States has more than 65% of the total value added by enterprises that employ over 250 people. On the contrary, almost 70% of the total value in Italy is added by enterprises that employ under 250 people (OECD, 2019). Therefore, although this established consensus shows that entrepreneurship, in most cases, leads to economic growth, its extent differs based on many factors such as the regional framework, free market, infrastructure, existing technology and so on. Moreover, the magnitude of this relationship also changes based on the type of enterprise you take into account in your research. Hence, David Birch established an animal analogy for the three types of enterprises based on their size and the speed of their growth. Elephants are companies that are very large and do not grow quickly (for example: Coca Cola). Mice are small companies that mostly stay small and rarely get a chance to grow (i.e. local butcher shop). Lastly, gazelles which are companies that grow from a mouse to an elephant at a rapid rate and usually disrupt the market (example: Facebook). The point is that if you study gazelles, you are likely to find extremely good results of the effect of entrepreneurship on economic growth. However, if you repeat your study using only mice enterprises in your sample, you are unlikely to get similar results. Similar goes for the maturity of the enterprise you investigate. The research which includes all enterprises disregarding the number of years since their establishment can produce biased results. According to Fritsch and Mueller (2004), time is a very important dimension in the magnitude of the effect that enterprises have on

the economy. Driven by the aforementioned methodological issues, I decided to investigate this relationship addressing them and avoiding biased results. Bearing that in mind, this research intends to fill the gap in the current literature in several ways listed below:

- 1. A substantial portion of academics that investigate the interplay between entrepreneurship and economic growth attempt to do so on a national level. If they decide to zoom in and inquire about the effect on the regional economy, researchers tend to pick regions in only one or a few selected countries. However, this thesis is going to be different, and contribute to the small body of literature that investigates entrepreneurship in NUTS 2 regions in the European Union.
- 2. Unlike former work on this topic, I managed to avoid issues from the last paragraph using a measure for the early-stage entrepreneurship which controls both for the size and the stage (maturity) of the enterprise. Hence, the results of this thesis are expected to be more robust and accurate.
- 3. Frequently, scholars that investigated this topic on the European Union level, do not control for NUTS 2 regions that contain large cities such as Paris and London within their borders. Therefore, I decided to control for that and introduce a city versus countryside effect which is going to reveal whether densely populated areas affect the magnitude of the relationship between the entrepreneurship and regional economic growth.

Societal Relevance

The importance of entrepreneurship for society is a widely popular topic in modern economic literature. According to the Global Entrepreneurship Monitor (2020), entrepreneurship offers a shortcut to health and wealth of the society, and it is a performance indicator of every modern economy. Furthermore, entrepreneurship has proven to be a great productivity promoter, employment creation tool, and it is environmentally friendly (UN Sustainable Development Goals, 2016). Entrepreneurship has a very important function where entrepreneurs serve as innovators and provide new tools or services for the market. Data from GEM shows that 5% of adults in the UAE, Canada, Colombia and Chile are involved in innovative entrepreneurship, meaning that they provide goods that were not available at the local level. Therefore, it is possible to conclude that innovation started to emerge through entrepreneurship more and more around the globe. The efforts made by the governments, NGOs, and international organizations in promoting entrepreneurship testify to its socioeconomic importance. The main point that this thesis seeks to explain is whether early-stage entrepreneurship leads to economic growth in European regions. Regional development and economic growth of regions significantly improve inhabitants' quality of life. By lifting the regional GDP, more people are employed which in turn results in a wealthier region and wealthier citizens. Wealthier and more competitive regions attract more inhabitants, spark productivity, attract companies and thus further improve the economic and social environment. Regional economic growth also leads to a greater tax income for the government which can choose to raise wages for the employees, finance large projects, or fill the gaps in the national budget if there is a deficit. Overall, the societal benefit of economic growth is, perhaps, the most well-recorded and established phenomenon of all time.

2. Theoretical Framework

This chapter is structured as follows: the first section 2.1. briefly covers the history of the development of entrepreneurship in economic literature through three different periods in the former and current century. Furthermore, section 2.2. explains what is likely to induce entrepreneurship in the region and how the regional environment attracts entrepreneurs to the region. Section 2.3. explains several causal mechanisms proposed in the literature of how an increase in entrepreneurship affects the economy. Section 2.4. presents different phases of entrepreneurship and defines the early-stage entrepreneurship. Then, section 2.5. outlines the most important theoretical basis of this research and offers answers to some questions that may have arisen throughout the previous several sections. Lastly, the theoretical framework is concluded with section 2.6. which presents the conceptual model, research questions, and hypothesis of this thesis.

2.1. The development of the economic literature on entrepreneurship

Very few researchers looked into the importance of entrepreneurship in the first half of the twentieth century. The first several decades of the past century were dominated by Karl Marx's (1867) seminal work, which claims that the physical capital creates a shortcut to economic growth and development. The world was dominated by large corporations and enterprises which employed thousands of people. The central thought of that time was that both the development of society and the economy are shaped by the physical capital. The Second World War is considered to be a turning point of entrepreneurial thought, and it is possible to identify three periods in which this thought developed (Wennekers and Thurik, 1999; Audretsch, 2018):

- 1. The early post-WWII era also known as the capital economy (the 1940s to 1970s).
- 2. The late post-WWII era also known as the knowledge economy (the 1970s to 1990s).
- 3. The entrepreneurial economy which lasts from the 1990s until the present.

Capital-based Economy

The importance of small enterprises and individual entrepreneurship rarely caught any publicity in the immediate post-World War II period. The emphasis was put on large incumbent companies and the well-being of the working class who worked for those companies. Small enterprises were ranked quite low on the social and political agenda. The literature in this period was dominated by Nobel Laureate Robert Solow's work (1956, 1957) which considered physical capital and unskilled labour as main drivers of the economy. As Nelson (1981) pointed out in his extensive literature survey, Solow's neoclassical production function became a focal point of the entire body of econometrical research on economic growth. Several subsequent generations relied on Solow's model while trying to explain what drives the economy. Physical capital and unskilled labour remained the most influential factors in virtually every research, and the only thing that varied is the unexplained residual in growth rates which was attributed to the fluctuations in technological advancement across countries and over time (Audretsch, 2018). The economic policy of the Western democracies,

intentionally or not, definitely corresponded to Solow's prescription. The macroeconomic organisation of countries focused on capital and labour on firm, industry, and market levels.

Small firms were regarded as less efficient than large-scale mass-production companies (Weiss, 1976). Moreover, Brown and Medoff (1989) found that remuneration of workers in small companies in this period was significantly lower than those that workers received in larger counterparts. Scherer (1970) asserts that the companies in this first period between the 1940s and 1970s who were involved in any kind of innovation were those that could afford R&D investments. In other words, smaller companies had no access to funds which they could invest in innovation. Lastly, Scherer (1970) also noted that there was a lot of effort and inclination by the people and governments across the Western world towards keeping larger companies afloat rather than investing in SMEs and entrepreneurs. All in all, the first postwar period was very collectively organised, and there was very little room for entrepreneurs and the small firms to manoeuvre. However, the situation gradually improves, and we see the rise in the popularity of entrepreneurship and SMEs in the knowledge-based economy.

Knowledge-based Economy

The second post-war period which lasted from the 1970s to 1990s is referred to as the knowledge-based economy. The most influential scholar of this era is yet another Nobel Laureate, Paul Romer. In his seminal work entitled "Increasing returns and long-term growth" in 1986, he described his view of how economic growth works. The emphasis that needs to be put here is that Romer did not criticise Solow's (1956) work. He thought that Solow's neoclassical function worked perfectly for capital and labour, but he considered that Solow overlooked one important factor: knowledge. Two years after Romer's (1986) seminal paper, Lucas (1988) confirmed his idea and argued that knowledge is indeed a crucial factor for the production function. Romer's and Lucas's argument was that externalities, spillovers and the context of the economy were crucial for long-term growth. Paul Romer is considered to be one of the vital contributors to the endogenous growth theory. The theory claims that the so-called technological change within the company is not a random noise like Solow proposed. Instead, Romer (1986) stated that it is a product of intentional individual action, like, innovations within the research and development department.

A considerable amount of quantitative studies support knowledge-based production function and endogenous growth theory. For instance, Cohen and Klepper (1992) found that new economic knowledge as a product of investment in R&D yields high innovation. Moreover, Acs and Audretsch (1990) found that there is a strong relationship between R&D and the number of patents, and they found a very strong positive correlation between investments in R&D and innovative output based on SIC industrial classification. However, despite large support in favour of this new production function, entrepreneurship was still not compatible with Romer's theory. Small companies were not able to invest enough in the production of knowledge and their innovation output was quite low. However, globalisation brought technological change, especially in ICT and other high-tech industries, which had several breakthroughs during this period. Innovation output in the ICT sector in the United States increased and the innovation was compensated with very high wages. Kortum and Lerner (1997) state that the number of patents from 1985 until 1990 rose from only a few thousand per year to about 40 to 80 thousand per year. Suddenly, by placing so much importance on investment in innovation and human capital, endogenous growth theorists unravelled a chain reaction which would later heavily influence macroeconomic policies in Western democracies. Many countries shifted away from the previous models dominated by large companies and unskilled labour, towards more open, innovative and competitive markets. This shift in policy enabled entrepreneurs and small companies, which were young and innovative, to grow beyond anyone's imagination.

Entrepreneurial economy – a product of a knowledge spillover effect?

Throughout the 20th century, there was a little room for entrepreneurship and experimenting with new and small enterprises, especially in post-war Europe. Conventional economic models, such as Solow's or Romer's were hostile or incompatible with entrepreneurship. On a similar note, economic policies at the time were consumed by this thought that bigger companies are better for most of the post-war period. The evidence to this is that only 20 percent of manufacturing sales in the United States in 1976 was accounted for by small firms (Brock and Evans, 1989). Nevertheless, rapid expansion and the strong influence of globalisation brought good news for entrepreneurs, start-ups and small enterprises. Therefore, just ten years later in 1986, a stake in manufacturing sales by small firms had risen to more than 25% (Brock and Evans, 1989).

Interestingly, Storey and Johnson (1987) reported that small companies create most of the jobs in the United Kingdom. Similar evidence about high job creation, and generally low gross job destruction, started flooding in the world of academia. At the turn of the century, a lot of researchers from Western countries presented their results about the positive impact of entrepreneurship on the economy and labour market (Konings, 1995; Hohti, 2000; Heshmati, 2001). As scholars presented more and more results, researchers were looking for the potential causal mechanisms that could explain this major shift to 'smallness' and entrepreneurship. Brock and Evans (1989) put forward five causal mechanisms. Firstly, foreign rivalry and globalisation created uncertainty and volatility in the global markets. Secondly, technological change influenced some radical alterations in the scale of manufacturing and the labour required for it. Thirdly, the rise in privatisation and the lack of government's regulation of the market created opportunities for entrepreneurs. Fourthly, the work flexibility and good remuneration that is associated with entrepreneurship became more attractive to young people, immigrants, and families with children compared to a standard nine-to-five job. Lastly, consumers became pickier and tended to consume nicheproduced rather than mass-produced goods.

The loss of large enterprises' competitiveness on the market forced many companies to reform their businesses. Many, faced with a foreign competition which produced the same good for a much cheaper cost, had to scale down their production and restructure the workforce. Some of them had to take even more severe measures such as relocating to low-cost areas or shutting the company down entirely. Therefore, the globalisation forced the market in Europe and the United States to shift towards a knowledge-based economy. In these circumstances, virtually every small enterprise thrived because they were mostly knowledge-intensive businesses which could not be affected by the strong foreign competition and cheap labour. Technological firms are the most extreme example of this advantage where start-ups such as Microsoft, Apple, IBM, Intel and many others dominate

the market share. During this shift towards a knowledge-based economy, many scholars and politicians had to adapt to the new situation and start researching how to shape the economic policy for the future. Conventional theories throughout the post-war period considered innovation a product of a hefty investment in R&D, mostly done by large incumbent firms which could afford it (Acs and Audretsch, 1990). Suddenly, the same large incumbents that were at the forefront of innovation, found themselves in big trouble facing foreign competition and restructuring their business.

The potential breakthrough in research about the importance of entrepreneurship occurred when Cohen and Levinthal (1989) found that small firms are very likely to commercialise the unpursued existing ideas which had been produced by large incumbent firms' R&D departments. This idea formed a new academic theory in the literature usually referred to as the knowledge spillover theory of entrepreneurship. Several strands within this theory appeared but the main point is that the knowledge was already in the proximity, and it was not commercialised yet. Smaller firms picked up on that knowledge and formed their business around it (Audretsch, 1995). While globalisation was thriving, researchers such as Leamer (1997) predicted that the role of geographical proximity would significantly reduce. With the rise of the internet and the possibility of working and acquiring knowledge remotely, people disregarded the importance of social proximity. However, history has proven them wrong. More and more companies started locating in a similar region, city, or even neighbourhood, thus forming clusters such as Ruhr Valley near Dortmund. Therefore, academics started realising that the proximity is more important than previously thought. In his seminal book 'The Triumph of the City', Edward Glaeser (2011) concluded that the globalisation brought the speed of transferring information, and not transferring the knowledge. Therefore, tacit knowledge remains one of the best explanations of why geography plays a crucial role. Audretsch and Feldman (1996) claim that the marginal cost of the knowledge transfer increases as the distance between the source and the recipient increases. Additionally, Klepper (2007) shows that new small businesses tend to be located near the incumbent companies or sources of knowledge such as universities and science parks. The best example of this relationship is Silicon Valley, which is extremely close to Stanford University. The proximity to a great source of young and talented students perhaps created the biggest and the most innovative high-tech cluster in the world.

All in all, the modern economy has shifted from managed to entrepreneurial since the 1990s. With this shift, the central actors of the contemporary growth are no longer labourers nor physical capital. Today individuals and their knowledge play a central role in economic growth (Fritsch and Storey, 2014). Small and medium enterprises became a focal point of policies because every regional economy wants to facilitate high growth start-ups and become the next Silicon Valley. According to the US Census Bureau (2012), almost half of the entire workforce is employed by small businesses. The importance of the small enterprises became evident with companies such as Alphabet or Walmart and many others that were very small only a few decades ago. Today, those companies are global conglomerates. Flexibility, availability of venture capital, availability of human capital from the adjacent universities, and open immigration policies all contributed to the growth in the popularity of entrepreneurship. Acs and Audretsch (1990) emphasised that the shift from large company culture to 'smallness' increases innovation. Similarly, Cohen and Klepper (1992) argue that the main difference between large and small firms is the level of entrepreneurship within them. Smaller

companies tend to be led by one or a small group of individuals – entrepreneurs - who make all decisions about the company's future. With such power vested in them, entrepreneurs have the main role of implementing innovation and thus creating knowledge spillover which leads to economic growth (Audretsch et al., 2006; Audretsch, 2007). Moreover, Kirzner (1997) explains that innovation and entrepreneurship go hand in hand and help economic growth by revealing the economic viability of new business models which have not been used since. This creates an opportunity for even more companies to enter these niche markets. An example of such an innovative business model is a social network such as Facebook or Twitter which monetize their service with paid advertising.

While the stream of literature on the relationship between entrepreneurship and economic growth has, more or less, reached the consensus that entrepreneurship helps economic growth, very little is known about the mechanisms of how entrepreneurship affects economic growth, taking into account regional differences. Therefore, according to Capello and Lenzi (2016), the impact of entrepreneurship on regional growth should be studied to greater detail. The next chapter will focus on entrepreneurship in the region.

2.2. What determines a level of regional entrepreneurship?

Entrepreneurship is seen by academia as a very important driver of a regional economy, development and employment opportunities (see Praag and Versloot, 2007). Furthermore, it increases productivity, the value of the real estate and has many other positive externalities. However, the level of entrepreneurship varies from region to region and from country to country. Scholars such as Michael Porter (1990) emphasised that regional entrepreneurship varies because of context. Others, such as Praag and Versloot (2007) in their extensive literature survey, argue that regional differences can be attributed to entrepreneurship at the individual level. As with every topic, there is some disagreement in the literature, but two general levels can be identified that explain the variation in the regional entrepreneurship: regional and individual level. The regional level refers to institutions, policies, financial market, and the overall context (example in Figure 1). The individual-level refers to the 'entrepreneurial spirit' or the ability and willingness to take risks for profit (Batchelor and Burch, 2012).

On a micro level, studies have shown that individuals who seek to own their business have certain characteristics (Lee et al., 2004). They either have higher educational attainment, are of a particular ethnicity or just have an 'entrepreneurial' personality, which refers to divergent thinking, or the ability to think 'out of the box' is one of the main characteristics of an entrepreneur (Runco et al., 2011). Chell et al. (1991) explain that individual entrepreneurs have a vision and are proactive. Furthermore, they favour individualism and dislike herd behaviour. Roberts and Wainer (1971) claim that entrepreneurship attracts individuals that want to make more money. It is a completely valid claim since entrepreneurship is proven to be a great driver of personal wealth. All in all, the body of research that studies micro-level entrepreneurship is the largest of all. However, regardless of the individual traits required to start a business, entrepreneurs require the right policy, services and the so-called 'ease of doing business'. That is why regional framework needs to be taken into account.

It is possible to draw some conclusions from the literature on why entrepreneurship varies from region to region. While individual traits are given a lot of attention (example Global Entrepreneurship Monitor), much of the literature on what causes a growth in entrepreneurial activity neglects the importance of the regional environment (Fischer and Nijkamp, 2019). On a macro level, Malecki (1997) points out that entrepreneurial environment, defined as socio-economic, political, infrastructural, and market environment, is crucial for entrepreneurship. Moreover, cultural characteristics play a significant role because they reflect the way the business is done. Regions with high social trust levels are more likely to have a higher level of entrepreneurship (Roberts, 1991). On the contrary, environments where businesspeople and governments cannot be trusted is very unlikely to have high levels of entrepreneurship. Unfortunately, this field of social trust and entrepreneurship is very unexplored by academia. The most important factor that makes a great entrepreneurial environment is a network of people who engage in entrepreneurship. In a thriving network, it is possible to get access to capital, information and support. Therefore, Glaeser (2011) infers that metropolitan regions are the most optimal form of cohabitation, where innovativeness, human capital, and great entrepreneurial climate have virtually no boundary.



Figure 1. Entrepreneurial Conditions Composite Index by EOCIC (2019)

0.1700 1.0000

Note: Higher score denotes better entrepreneurial conditions in the European Union. The composite index is an aggregate score of several measurements outlined in EOCIC metadata (2019). Source: European Cluster Collaboration Platform (2019)

As shown in Figure 1, regions have macro-level characteristics which foster or hinder the level of entrepreneurship. According to the map above, large European economies such as the UK and Germany have more favourable conditions for entrepreneurs. However, they also vary from region to region. Berlin and London are much more favourable than some other regions in the same country. Unsurprisingly, Scandinavian countries, the Netherlands and Switzerland have great conditions for entrepreneurship as well. The country which scores surprisingly high is Estonia. Zooming into Estonia after seeing this report, one might notice that it became an entrepreneurship hotspot for many tech companies in the past decade because of the favourable policies by the local and national government and great digitalised governmental services. One instance is a very popular multi-billion fintech company TransferWise.

2.3. The interplay between entrepreneurship and economic growth

The fact that entrepreneurship is positively related to the economic growth of regions is deeply rooted in the literature on this topic. However, the causal mechanism has not been established properly. A lot of the studies suffer from the hen-egg (or egg-hen) problem where they are unable to detect whether higher levels of entrepreneurship cause regional economic growth or vice versa. The most researched explanation of why entrepreneurship drives regional growth is the high-tech sector clustering. Researchers in this strand of the literature claim that high technology start-up firms significantly improve the economic growth of the regions (Hart and Acs, 2011). Their theoretical framework lies on the assumption that human capital and entrepreneurial spirit increase innovativeness which results in the growth of new high-tech firms. Then, those new firms form relationships with the supply chains, labour market, institutions, educational facilities and so on. The success of high-tech firms in one region attracts even more human capital and people move into the same region to acquire knowledge from the people already there, and potentially start their own business (Porter and Stern, 2001). This behaviour helps in the formation of high-tech clusters such as in Bangalore, India. According to Crescenczi et al. (2007), the formation of high-tech clusters is an example of an endogenous growth which can be attributed to the knowledge spillovers, and the proximity to the high-skilled labour. All in all, the high-tech sector has been researched extensively and a lot of studies point out that this particular sector is the biggest contributor to regional growth. However, a lot of high-tech start-up companies become profitable only after several years of existence. Therefore, due to their nature, such companies do not contribute significantly to the regional economy for the first few years (Fritsch and Mueller, 2004).

Another explanation in the literature suggests that the positive correlation between entrepreneurship and regional growth can be attributed to the increased competition between incumbent companies and new start-ups (Fritsch, 2013). The entry of new businesses which challenge the incumbents in the same region affects the regional growth positively because they offer new jobs and generate value for the region. However, Fritsch (2013) distinguishes between high-quality start-ups and poor-quality start-ups. He states that while high-quality start-ups generate value and offer new employment opportunities, poor-quality ones fall trapped into what is known as 'revolving door development'. This term was coined by Audretsch and Fritsch in 2002 and represents industries or regions where new businesses quickly leave the market and fail to contribute to the development of the region.

In a 'revolving door' regions, start-ups tend to be less or not innovative at all and supply already available products using the same technology as incumbent companies. Surprisingly, Audretsch and Fritsch (2002) found that the high concentration of high-quality start-ups in a region has a long-term impact on the regional economy, even if the number of start-ups significantly reduces over time. In other words, just a few great start-up companies can set the region on a long-term growth path. In short, they suggest that quality is more important than quantity when it comes to the new firm formation.

The third causal mechanism in the literature is that the entrepreneurial culture affects the level of entrepreneurship in the region which then affects the regional growth. As in the previous chapter, entrepreneurial culture refers to the quality of facilities for start-up companies, willingness to take a risk, an existing network of entrepreneurs, et cetera (Beugelsdijk and Noorderhaven, 2004; Sternberg, 2009). Studies from a number of Western European countries and the United States show that the locational advantage due to the culture is clearly visible for at least 10 to 20 years after the peak. According to Bosma et al. (2012), regional start-ups appear to be motivated by regional norms and values which induce regional economic growth. Therefore, it is in line with Figure 1 from the previous section which displays this locational advantage across NUTS 2 regions in the EU.

Lastly, Reynolds et al. (1994) claim that population growth and density can positively affect the number of entrepreneurs and economic growth in regions. The rationale behind this study, and many others who found a similar effect, is that a dense area 'forces' people to interact and share knowledge. Moreover, regions with high population growth and density present large market and open up opportunities for any type of SMEs. The economies of scale in the large market makes the business cheaper and more financially sustainable.

In summary, there are numerous causal mechanisms presented in the preceding literature on how entrepreneurship causes (regional) economic growth. This literature review only covers a portion of them. Nevertheless, despite a number of studies showing the positive relationship between growth in entrepreneurship and regional economic growth, some researchers present mixed results. For instance, Acs and Storey (2004) show that a lot of the studies they reviewed show inconclusive evidence about the effect of an increase in the number of new firms on regional growth. Moreover, Stearns et al. (1995) and Hart and O'Reilly (2005) argue that the level of entrepreneurship differs from region to region but the relationship between the location and a decision to pursue entrepreneurship remains blurred. Captivatingly, Fritsch and Mueller (2004) investigated the firm formation in Germany and discovered that the effect of a new firm formation on employment is likely time-lagged. Fritsch and Mueller (2004) then argue, there are also time lags between the effect of a new firm on its environment (region). In their discussion of the results, they argue that the effect of the new firm formation can last for up to 9 or 10 years. This 'echo' could have significantly disrupted many results of research done before and after their study that did not address this issue. All in all, there are many more critiques of the previous research that point out the deficiency in the methodology, framework or data that could have been improved.

2.4. Phases of entrepreneurship and early-stage entrepreneurship

According to the Global Entrepreneurship Monitor, there are several phases of entrepreneurship depicted in Figure 2. The process starts from a potential entrepreneurship stage. The potential entrepreneur requires certain skills and knowledge to start their own company. Apart from individual skills, regional context (environment) also affects the likelihood of an individual becoming an entrepreneur. In case an entrepreneur decides to move forward and try to start a new business venture, they enter the early-stage entrepreneurship stage. It lasts up to 3.5 years and involves creating initial employment opportunities, seeking funding, cooperating with the industry, acquiring knowledge, et cetera. Should they succeed for the first 3.5 years, entrepreneurs enter the last phase of entrepreneurship, which is an established business phase. By the time they enter the last phase, companies are more than 3.5 years old and are likely profitable, or investors bet on their profitability in the near future.

This thesis primarily concerns Total Early-Stage Entrepreneurship (Figure 2). The term stands for new businesses which are less than 3.5 years old. Instead of the number of people, I will use the number of firms that are in the 'first' phase of entrepreneurship in the region to capture the early-stage entrepreneurship across NUTS 2 regions (see Figure 2).



Figure 2. Definition of Total Early-Stage Entrepreneurial Activity

Source: Global Entrepreneurship Monitor (2013)

2.5. How does early-stage entrepreneurship affect regional economic growth?

The effect of entrepreneurship on regional development is a very broad and interdisciplinary field that can be studied from multiple perspectives. The foundation of all theories so far is Schumpeter's (1934) seminal study which pointed out multiple positive effects of entrepreneurship on the local economy. Additionally, the gradual change throughout history from managed to the entrepreneurial economy allowed the entrepreneurship to find its place in many economic policies around the world. The growth in popularity and the availability of data since the 1990s enabled scholars to identify small and medium enterprises as the main drivers of the labour market dynamics and job creation (Birch, 1981; Armington and Odle, 1982; Audretsch, 2002). Moreover, the availability of venture capital, the global mobility, and the proliferation of business models that never existed before, all contributed to the 'explosion' in the level of entrepreneurship being done today (Lockett and Wright, 2005). In recent times, significant developments have taken place around the globe which opened new opportunities for entrepreneurs. Namely, with the rise of the internet, the world suddenly became very small. People work remotely, offer paid services online, collaborate, develop new business models and ideas, and create markets where no business has operated before. However, what impact did this explosion in new and small-scale entrepreneurship have on the regional economic growth?

The extent of the contribution of entrepreneurship to regional growth varies from place to place. It was not until the endogenous growth theory by Romer in 1986, that individualism, knowledge, and innovation became formally included in the macroeconomic models of economic growth. However, the endogenous growth theory does not reveal how knowledge and innovation might be translated into economic growth. Several theories in the literature attempt to explain the effect of entrepreneurship on regional growth (Carree and Thurik, 2010). Due to the space constraint, I will only consider one which I find the most relevant for the theoretical framework. It is the so-called knowledge spillover theory of entrepreneurship which highlights the role of start-up companies that foster radical innovation. It is believed that the combination of talent and knowledge produces radical innovation.

According to Acs et al. (2000), the knowledge that is required for a decision to engage in entrepreneurship often comes from a nearby educational institution, such as a university or an institute. Apart from higher education, employees in incumbent companies recognise entrepreneurial ideas and decide to pursue them. Those ideas often come from R&D departments and are unpursued by the management of the company (Klepper, 2009). Since some skills and knowledge are not easily marketed or transferred, the person who possesses the knowledge is very likely to be the founder of the new innovative start-up. Sometimes, start-ups which are products of knowledge spillover from educational institutions or incumbent companies are referred to as the spinoff companies. In the process of this spillover, Boschma (2005) reports that the knowledge remains tacit to the region. It does not move freely across space as it was believed in the past. It is quite contradictory to the globalisation and the rise of the internet we mentioned before. Interestingly, Dahl and Sorenson (2009) claim that the founders of these spinoffs tend to locate very close to the place they used to study or work. Therefore, it is possible to identify clusters of radically innovative companies close to HEI and incumbent large companies. The spatial concentration

of many new innovative companies in the region has proven to significantly correlate with the economic growth of that region.

The theoretical basis of this thesis is that the number of new business formed, namely the ones that are less than 3.5 years old, explains the difference in the economic growth across regions in the European Union between 2008 and 2017. The newly formed firms pose a threat to the incumbent companies and results in a competition between them. Therefore, only a fraction of start-ups manage to survive the early-stage entrepreneurship, and many of them experience a market exit in this period. In case they survive, incumbent companies can also experience market exit and 'be thrown out' of the competition. The competition between new start-ups and incumbents is based on natural selection. Firms that have high gross productivity will keep their spot, while those with relatively low productivity will have to adjust their productivity or lose the competition and leave the market. This constant competition and a necessity to increase productivity results in regional productivity to increase as well. The increase in regional productivity translates into fewer resources and labour to produce the same good. One might infer that the increase in productivity results in the negative labour market effect where people lose jobs. However, this is not always the case, because the labour market dynamics often compensate for the negative externalities of the increase in efficiency.

Although competition between start-ups and incumbents always exists, it is not necessarily the elimination of one or the other. Generally, competition increases the supply side in the economy which is considered as positive for regional competitiveness and growth (Fritsch, 2013). More specifically, new entrants offer the following to the regional economy:

1. Create new markets with novel solutions and products.

Saviotti and Pyka (2004) report that diversification in the regional industry and knowledge base results in a sturdier and more resilient regional economy. Moreover, such entrants offer innovative products and services to consumers and solve problems which were unsolved before they entered the regional market.

2. Enhance innovation output

Klepper and Sleeper (2005) report that incumbent firms often decide not to pursue the same product as entrants even if they planned to because the profits are not satisfactory for their scale of operation. Instead, incumbents decide to leave the idea 'hanging in the air' and pursue another plan. Additionally, incumbents are very unlikely to create a product which is likely to compete with their current products on offer. However, inventors are very likely to start their own business and commercialise the idea on a smaller scale (Klepper, 2009).

3. Firm dynamics

The complete structure of the market changes. Some new firms enter the markets and incumbents leave the market at the same time. The correspondence can be found in Marshall's (1920) influential thesis when old trees in the forest have to fall to make room for new ones.

4. Change in productivity and efficiency

As aforementioned, the new entrants or even potential entrants force incumbent companies to restructure and make their production more efficient in order to adapt to the new situation. Else, they need to leave the market.

However, the change in the structure of the market and the supply-side effect has no spatial boundary. It can vary from region to region because new start-ups in one region can disrupt markets in other regions and force incumbents to retreat from the market. It is also cross-industry, because seemingly unrelated industries may disrupt other companies through the supply side. Audretsch et al. (2006) explain that the cross-industry and cross-region competitiveness drives regional growth of one region and in the process might hinder the growth of another. The dynamic changes in the market can offer new employment opportunities, attract demand and create other positive externalities.

All in all, early-stage enterprises are seen as a positive influence on the economic growth of the regions. However, the current proposed theory is rather naïve and simple. Economic growth can be boosted by entrepreneurial activity and firm creation, but it is impossible to attribute the growth entirely to the entrepreneurship because it is heavily influenced by external factors. The level of entrepreneurship is determined by the regional cultural norms and values, national policy, foreign direct investments and many more crucial elements that foster or hinder, entrepreneurship in the region. The conceptual model outlined in the next chapter will present more about the external effects that create a breeding ground for regional entrepreneurship, which then can influence regional economic growth.

2.6. Conceptual model

Building upon the literature review which presented the historical development of the entrepreneurial thought and an increase in the importance of entrepreneurship in the modern economy, we now move on to the conceptual framework of this thesis. Figure 3 represents the conceptual model of this thesis. The model assumes that everything starts from the social, cultural, and political context. This context can be divided into two frameworks. One is a regional framework which is exogenous to the level of entrepreneurship. The regional framework can include an existing labour market, special regulations, rule of law, infrastructure, existing technology, financial market opportunities, large incumbent companies, existing SMEs and many more factors that can influence the decision of an entrepreneur to start or not to start a business in a certain region. One example would be a region of 'Catalunya' in Spain which has a lot more favourable conditions for starting a business compared to 'Jadranska Hrvatska' (Adriatic Croatia) region of Croatia. Therefore, the conceptual model assumes that the regional framework exogenously affects the entrepreneurial framework. This framework consists of the opportunities to access finance, access existing technology, go through special programmes for start-ups such as incubation, access to the free market, access to the infrastructure and so on. This entrepreneurial framework creates a perception of opportunity for entrepreneurs who start new enterprises. These new enterprises are believed to positively affect the regional economy. The rationale in this study is motivated by the literature in the past, especially the Global Entrepreneurship Monitor's Report in 2020.





Note: The model is inspired by the conceptual model outlined in the Global Entrepreneurship Monitor's 2015/16 Report (GEM, 2016).

This thesis seeks to answer the main research question:

To what extent can the early-stage entrepreneurship explain for regional economic growth across all NUTS-2 regions in the European Union between 2008 and 2017?

And a sub-question:

Is entrepreneurship's contribution to the regional growth more pronounced in the places with higher population density such as large cities compared to their town and countryside counterparts, as suggested in the previous literature (Audretsch, Belitski and Desai, 2015)?

Lastly, the hypothesis of this research is that the early-stage entrepreneurship causes higher economic growth across NUTS2 regions in the European Union between 2008 and 2017. The rationale behind this is that a higher concentration of new enterprises located in the same area creates many benefits for the economy and the inhabitants of that region.

3. Methodology and Data

Following the theoretical framework and the conceptual model, section 3.1. presents the methodology and the estimation models that are used to investigate the relationship between early-stage entrepreneurship and regional growth. Then, section 3.2. presents variables in the study, data collection strategy, and the further process of cleaning the data. Lastly, the section ends with several choropleth maps which serve as regional descriptive statistics.

3.1. Methodology

Despite entrepreneurship can be traced to ancient history, it is a quite young academic field (Low, 2001). Regardless of its only recent gain in popularity, entrepreneurship has been researched in virtually every possible way. For instance, academics used both quantitative and qualitative methods such as surveys, case studies, interviews, secondary data analysis etc. (Gartner and Birley, 2002). While entrepreneurship is complex and cannot be captured using only one type of analysis, sometimes the unfeasibility of doing another method simply constrains the researcher. In this case, qualitative analysis is not possible due to the high number of units of observation investigated. The cost and time required to do it is a large obstacle. Therefore, I decided to do a quantitative analysis based on the already collected data by Eurostat.

In order to investigate the relationship between the early-stage entrepreneurship and regional economic growth, three methods will be employed: bivariate correlation analysis, fixed effects regression with individual and time effects including the time-lag structure of the independent variable (Figure 7), and spatial econometric analysis including both temporal and spatial lag of independent variable X, based on the suggestion from Elhorst (2014).

The analysis starts with a bivariate correlation, where variables are not treated as dependent or independent. They are simply paired to check the strength of correlation between them using Pearson's coefficient. The coefficient ranges from -1, which stands for a perfect negative relationship between two variables, and 1, which stands for a perfect positive relationship. The results from the bivariate correlation are used to check the correlation between each of the variables and their connection to the natural log of GDP per capita. Besides, the results from this analysis help to check for multicollinearity between variables and create models based on that inference. This way, the bias of explaining away the effect of the early-stage entrepreneurship on regional economic growth is eliminated because too significant variables will be found and removed promptly.

Apart from bivariate correlation, tests of normality, linearity, and homogeneity of variance were conducted to proceed with linear regressions. According to Herwartz (2006), the Breusch-Pagan Lagrange Multiplier test should be conducted which checks whether I can use the OLS model or fixed/random effects model. The test resulted as statistically significant at 1% level which suggests the use of fixed or random effects for this panel. After the Breusch-

Pagan LM test, I conducted another test called the Hausman test to check whether fixed or random effects yield more precise results. The test's result was statistically significant at 1% level which suggests the use of fixed effects, in this case, is better and more efficient. After deciding to model fixed effects regressions, I had to check whether to include time effects, individual (regional) effects or both. This can be checked with the Lagrange Multiplier test two-ways effects (Honda) for balanced panels, where significant results recommend the use of both individual and time effects. Indeed, the test was statistically significant at 1% level, and hence I decided to use a fixed-effects model with both regional and time effects.



Figure 4. Time-lag structure of the economic effect of enterprises

Source: Fritsch and Mueller (2004)

The time-lag structure (Figure 4) is used because changes in the economy caused by the business formation can be a long-term process. Fritsch and Mueller (2004) explained that there are three periods of how new enterprises affect economic growth (Figure 4). In the first phase, potential entrepreneurs first have to educate themselves and get proper guidance on how to enter the market. Then, entrepreneurs need to start their company and obtain resources for operating a business. After the business is up-and-running, there is an immediate effect of employment growth because new businesses are very likely to hire people at the time of their establishment. The first phase lasts for approximately 1 to 1.5 years and it is expected to have a generally positive effect on the regional growth. The second stage lasts from 3 to 4 years and refers to the negative effect where infant enterprises are the most vulnerable and most of them likely to exit the market. Nevertheless, not only new firms exit the market, but also the incumbent companies. Sometimes, newly formed enterprises force incumbents to reduce the manpower or completely move out of the region. In that case, jobs are lost and the impact on economic growth is negative.

Lastly, the third phase corresponds to the increased market efficiency, competition, innovation, and a variety of products available that satisfy various consumers' preferences. The new enterprises then become incumbents with more than 6 years of experience on the market. Additionally, Fritsch and Mueller (2004) explain that the third year of existence

corresponds to the peak negative, and the sixth year corresponds to the peak positive effect of new enterprises on the regional economic growth. However, Fritsch (2008) warns that the time-lag structure may only apply to some regions and some sectors. Carree and Thurik (2008) report that they tested for the three periods in the time-lag structure and found evidence that the lags influenced their results on three different economic measures: GDP, labour productivity, and employment growth. All in all, I am going to test the lag-structure in my thesis and inspect the results. <u>The analysis is going to include third and sixth lag of the number</u> <u>of three-year-old enterprises</u>. If I observe a change in results, the effect can be attributed to the time-lag structure as described in Fritsch and Mueller (2004).

In addition to the time-lag model, it is important to control for spatial effect among European regions. Ever since Marshall's (1980) seminal work, economists became aware of the spatial externalities that can influence production. Large localised labour pools, tacit knowledge and high population density resulted in lower production costs (Guerrero et al., 2014). However, the emergence of new economic geography as an academic field started to investigate spatial effects in economic theory and include them directly in mathematics (Krugman, 1991). According to Varga (2006), geography is a very important factor when looking at technological change and economic growth for three reasons. Firstly, the "proximity effect", which refers to the spatial proximity of knowledge that spurs innovation. Secondly, the "agglomeration effect", and lastly, the "cumulative causation effect", which is described as the initiation of the cumulative growth that agglomeration economies can experience (Varga, 2006, p. 1172).

Spatial patterns of knowledge transfer became a focal point of many economic geographers. Capello and Lenzi (2016) report that a significant portion of knowledge remains immobile. While some knowledge can be transferred through communication networks such as the internet, some knowledge such as tacit or immobile strictly requires personal interactions to transfer from one to another. Therefore, it is plausible to claim that new businesses which manage to find market niches are more likely to locate in regions that have existing tacit knowledge they can build upon. Additionally, agglomeration impacts the spatial concentration of businesses and the number of related firms and supply chains of knowledge and products (Fujita and Thisse, 2002). Therefore, Varga (2006) concludes, the proximity of knowledge and agglomeration effect significantly reduce the cost of innovation and incentivises businesses to relocate into the region. All in all, it is suggested that spatial spillovers across regions need to be taken into consideration. Therefore, a spatial econometric model is introduced to empirically assess the strength of this spillover (Vega and Elhorst, 2015). Moreover, Vega and Elhorst (2015) suggest that the spatial lag of X (SLX) model should be used as 'a point of departure'. This method is also used in Capello and Lenzi (2016), who investigated European NUTS 2 regions in a similar setting. According to Vega and Elhorst (2015), SLX model is the best fit because of its simplicity and the Hausman test for endogeneity which can identify the probability of endogenous variable in the model. The rationale behind these tests is explained further in Hill et al. (2018). I now move on to the econometric model.

The econometric foundation of this thesis' model can be traced to Solow's (1956) production function which considers human capital and physical capital as main drivers of the process of regional convergence and economic growth. Several decades after the establishment of the neoclassical production function, Barro (1991) came up with regressions that include

exogenous drivers of growth, better known as "Barro regressions". More recently, Audretsch and Keilbach (2004), Mueller (2007) and other academics adapted and extended the traditional Barro's (1991) regression models on the field that investigates the relationship between entrepreneurship and economic growth. Correspondingly, Lopez-Baso et al. (2004), and Ertur and Koch (2007) provided a theoretical foundation of spatial spillovers, arguing that technological interdependency between regions can be captured with spatial models. Although many spatial econometric models can be applied here, Vega and Elhorst (2015) (also in Capello and Lenzi, 2016) argue that the spatial lag of independent X (SLX) is suggested because of its simplicity and efficiency. Additionally, this model includes the temporal lag of the independent variable. Therefore, the final model is a combination of the econometric specification of the Fixed Effects SLX model with individual and time effects, per the suggestion in Vega and Elhorst (2015, p. 343); Lam and Qian (2017), and is the following:

(1)

$$\begin{split} \ln (GDPpc_{it}) &= \alpha + \beta_1 \ln(Enterprises_{it}) + \beta_2 \ln (Enterprises_{it-3}) \\ &+ \beta_3 \ln (Enterprises_{it-6}) + \beta_4 \ln(Tertiary Education_{it}) + \beta_5 \ln (HRST_{it}) \\ &+ \beta_6 \ln(Capital_{it}) + \beta_7 \ln (Economic \ activity_{it}) \\ &+ \beta_8 \ln (Population \ density_{it}) + \beta_9 \ln (R\&D \ Expenditure_{it}) \\ &+ \ln (WEnterprises)\theta + \mu_i + \lambda_t + \varepsilon_{it} \end{split}$$

$$i = 1, 2, \dots, 273.$$

 $t = 1, 2, \dots, 10.$

Where i stands for any NUTS 2 region in the EU, and t stands for the period (year) between 2008 and 2017. Furthermore, $\ln (GDPpc_{it})$ represents the natural log of GDP per capita for any given individual region i at any given year t. Beta coefficients represent parameters to be estimated in the regression. $Enterprises_{it}$ is the independent variable which is a proxy for the early-stage entrepreneurship. It is already by definition forward lagged by three years (see Table 1). Enterprises_{it-3} and Enterprises_{it-6} are additional three-year and six-year time-lagged independent variables, as per the time-lag structure suggests in Figure 4. The following list of variables in the equation are the control variables and WEnterprises is the spatially lagged independent variable. θ represents the spatial effect associated with SLX (LeSage and Pace, 2011, p. 22). μ_i is a region-specific fixed effect and λ_t is a time-specific fixed effect. Region and time-specific effects are included because of the Lagrange Multiplier Test - two-ways effects (Honda) for balanced panels was conducted. Additionally, the rationale behind this is that for instance, the European Union is not operating like one large state. Every country or even every region has a right to amend its economic policy at any given moment. Hence, some regions put more emphasis on entrepreneurship as a driver of growth than others. Besides, region-specific and time-specific fixed effects are included to avoid estimation bias (Lee and Yu, 2010). Likewise, Elhorst (2014) argues in his results that studies which investigate space but decide to ignore time-specific fixed effects tend to underperform. Lastly, ε_{it} is an error term.

3.2. Data and Variables

Data were obtained solely from Eurostat to avoid a potential bias in the amendment of the NUTS 2 classification across the years. The data source is official, credible and widely used in the previous research on this topic. NUTS 2 refers to the second level of the Nomenclature of Territorial Units for Statistics which was drawn up by Eurostat to create comparable regional statistics across the European Union. Apart from data collection and comparison between the regions, NUTS is also used for policy interventions and access to the European Structural and Investment Funds which are created to help the lagging regions (Eurostat, 2020). This research uses the most recent NUTS 2 classification from 2016. The sample includes 273 NUTS 2 regions in the European Union in the period between 2008 and 2017, excluding 8 overseas and very distant territories (Região Autónoma dos Açores (PT), Região Autónoma da Madeira (PT), Martinique, Guadeloupe, Canarias, Mayotte, La Réunion and Guyane). Importantly, the sample includes the United Kingdom since it was still a member of the European Union in the period of interest. The years from 2008 to 2017 and the NUTS 2 level of European regional classification are both chosen exclusively because of the availability of data for variables of interest in this thesis. The emphasis must be put on the fact that if data were available for NUTS 3 level (the smallest regional division within the Eurostat framework), they would have been the best fit for this type of research since the literature on economic benefits of entrepreneurship suggests smaller administrative units yield more precise results. (Guerrero et al., 2015; Colombelli and Quatraro, 2013)

The panel collected from Eurostat was not complete. Roughly 8% of the panel was missing due to the reasons that are not disclosed to the public. Based on Eurostat's metadata, it is possible to notice that some variables are collected based on a gentlemen's agreement, and there is no legal obligation for regional or national authorities to produce statistics for those variables annually. Since missing data can be misleading and point the research into a different direction, they were imputed using the 'mice' package in R. 'Mice' uses a very complex algorithm which performs multiple imputations using Fully Conditional Specification (FCS) (van Buuren and Groothuis-Oudshorn, 2011). The algorithm creates an imputation model for each variable and imputes data based on other variables. After the algorithm has passed through data once and identified the patterns, it creates multiple possible imputations (default = 5) and uses the average of those 5 as the final guess. After careful statistical inspection of the imputations, it is visible that the imputation was performed very well. The example of a distribution of the original versus imputed variable is presented in Figure 12 in Appendix. Therefore, the panel is complete with 2730 observations for each variable. Then, the entire panel is transformed using a natural logarithm since it helps to standardize skewed data.

The dependent variable in this research is regional economic growth. Regional economic growth can be measured in many different ways, but it refers to the growth in the overall production in the regional economy. Economic growth tends to have direct positive effects on the region and its inhabitants' quality of life. Regional economic growth can lead to an increase in wages, locational advantages, growth in real estate prices, and many more socio-economic benefits. Therefore, it is important to study what drives that growth to a greater detail. The measure of the regional economic growth in this study is the natural logarithm of the regional gross domestic product per inhabitant. According to the OECD (2014), gross

domestic product (GDP) refers to the market value of all goods and services produced expressed in monetary terms in any given period. The regional GDP per inhabitant is simply calculated by dividing the gross domestic product of a region by the number of inhabitants.

This study seeks to answer whether the independent variable early-stage entrepreneurship significantly impacts regional economic growth. While many studies before this investigated the impact of entrepreneurship on the economy of a certain unit of observation, this thesis only looks at the very first stage of entrepreneurship. The underlying assumption is that who makes it through the first stage is often going to succeed in the long run. Therefore, the rise in the number of early-stage enterprises is expected to positively affect the regional economy. Linan and Fernandez-Serrano (2014) investigated the impact of the early-stage entrepreneurship on the national economic growth using cross-sectional data. According to them, the benefit of using this variable as a measure of a starting phase in entrepreneurship is because it is comparable across units of observation. Therefore, it is safe to assume that this variable is a reliable proxy for the starting phase of the entrepreneurship. Reynolds et al. (2005) describe that the first stage entrepreneurs significantly impact the economy by the swift expansion, thus creating employment opportunities. In general, scholars have reported that this overall early-stage entrepreneurship has a positive relationship with economic growth (Reynolds et al., 2001; Reynolds et al., 2002).

The early-stage entrepreneurship, which is a term coined by the Global Entrepreneurship Monitor refers to all enterprises which are less than 3.5 years old. Although GEM and many other scholars investigate the early-stage entrepreneurship as a measure for a number of people who are engaged in it, this study looks at the number of enterprises in the early-stage phase of entrepreneurship. The exact measure is the number of enterprises born in t-3 having survived to t where the unit of observation is an enterprise. The target population is the private sector (Eurostat). According to Eurostat's metadata, there is no size threshold but, generally, the enterprise is considered active as long as at least one employee is working at any time. All in all, the early-stage entrepreneurship is expected to be positively affecting regional economic growth.

I include several control variables since they are expected to influence the economic growth in the real world. Furthermore, they are chosen subject to the conceptual model outlined above and the availability of data on Eurostat. Therefore, not every factor mentioned in Figure 3 is included in the analysis. Control variables in this study are:

A) Human capital which is an important factor for regional economic growth. Countries and regions attracting more human capital are much more likely to grow their competitiveness and the economy (Glaeser, 2011). Therefore, human capital is expected to be positively related to the regional GDP. Human capital in this study is measured using two variables. Firstly, as a percentage of the regional population holding a tertiary degree. And secondly, the stock of people in Human Resources in Science and Technology (HRST) in high-tech. HRST is measured as the percentage of people aged between 15 and 74 who fulfil the condition of a definition of HRST (Eurostat). The data on HRST is obtained from the EU Labour Force Surveys based on the sectoral NACE 2-digit level.

B) Physical capital is probably one of the most influential factors which determine economic growth. Looking back hundred and seventy years, Karl Marx's Communist Manifesto celebrated the capitalist society and the importance of the physical capital for the socioeconomic wellbeing. Since then, the evidence on the importance of capital on economic growth has been a mixed bag. Some countries such as the United Kingdom and the Netherlands benefited largely from capitalism. However, Sub-Saharan African nations, for instance, experienced lower per capita outputs with the introduction of capitalism. Nevertheless, physical capital in economic theory is considered input of the production function and is regarded as a major influence on the level of productivity and economic growth. It is measured by Eurostat as Regional Gross Fixed Capital Formation (GFCF) in millions of EUR. According to OECD's Economic Outlook (2020), Gross fixed capital formation, also referred to as 'investment' captures expenditure on all produced assets that are going to be used in the production of other assets for at least one year. All in all, GFCF is expected to be positively related to the regional GDP per capita.

C) Economic activity rate refers to the number of people who are eligible to work and supply labour regardless of their employment status. and is expressed in percentage relative to the population (Eurostat's Manual on Employment Labour Survey, 2019). Since one of the top priorities for all governments is stimulating employment and creating economic activity (European Union, 2020), it is quite obvious that labour force is highly important for the production and regional development. Thus, employment is expected to be positively related to economic growth.

D) Population density which represents a number of people living in a certain area. It is calculated by division of the number of people and the regional area and it is measured in square kilometres. Population density is an important factor because densely populated regions with large cities are believed to be more efficient and have higher economic growth (Glaeser, 2011). Moreover, Heady and Hodge (2009) report that population growth is likely to result in positive economic growth. All in all, population density is expected to be positively related to the regional GDP per capita.

E) Innovation, which is measured as Intramural R&D Expenditure in billions of Euros. According to the Frascati Manual (OECD, 2015), research and experimental development (R&D) refers to the creative work to increase the stock of knowledge and create an application of that knowledge in the everyday life. Intramural expenditure refers to the current and gross fixed expenditure for R&D in any given year, disregarding the source of funds (OECD, 2015). Intramural R&D Expenditure is likely to be positively related to the regional GDP per capita since investment in innovation is generally thought to positively influence the economy.

Since it is difficult to measure culture, political environment, and the quality of institutions, such variables are excluded even though they are considered very important for the influence on entrepreneurs, and ultimately, on economic growth in the regions as well (see Figure 3).

	Variable name	Eurostat code
Dependent variable	Regional GDP per capita [logGDPpc] (log)	nama_10r_2gdp
Independent variable	Number of regional enterprises founded in t-3 and still active in t (number of three-year-old enterprises) [enterprises] (log)	bd_hgnace2_r3 (indicator: V11943)
Control variables	Percentage of people with tertiary education [tertiary_educ] (log)	edat_lfse_04
	The stock of people in Human Resources in Science and Technology (HRST) in high-tech (percentage of the active population) [hrst] (log)	hrst_st_rcat
	Gross fixed capital formation (in millions of EUR) [capital] (log)	nama_10r_2gfcf
	Economic activity rate relative to the population (in percentages) [economic_activity] (log)	lfst_r_lfp2actrt
	Population density (people per square kilometre) [popdensity] (log)	demo_r_d3dens
	Intramural R&D Expenditure (in billions of EUR) [RD_expenditure] (log)	rd_e_gerdreg

Table 1. Variables in the study

Note: Where applicable, measurement units are reported in round brackets. Variable abbreviations that are used throughout the thesis are reported in square brackets.

Table 2 presents descriptive statistics for all variables in the thesis. The number of observations is 2730 since there are 273 NUTS 2 region in the EU multiplied by 10 years time period. Looking at the table, it is possible to notice that data deviates a lot because of the divergence of European regions. For instance, Ciudad Autónoma de Melilla (ES), which is a very small region in Spain had only €216.4 million in gross fixed capital formation in 2013. Meanwhile, in the same year, Ile de France (FR) region has more than €154 billion in gross fixed capital formation. Moreover, variables such as population density, the number of enterprises founded three years ago and still active, and intramural R&D expenditure have standard deviation larger than the mean. Overall, the data suggest that the sample is remarkably diverse.

Table 2. Descriptive statistics							
Statistic	Ν	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Natural log of GDP per capita	2730	10.02	0.60	8.04	9.70	10.42	12.33
The number of three-year- old enterprises	2730	7446	7650	139	2979	8960	67618
Percentage of people with tertiary education	2730	27.60	9.65	6.80	20.20	33.50	74.70
Human Resources in Science and Technology	2730	39.75	9.79	12.90	32.70	45.60	81.80
Gross fixed capital formation	2730	9539	11091	216.4	3449	11950	154285
Intramural R&D expenditure	2730	929	1576	63.30	113.02	1134	18664
Economic activity rate	2730	63.83	5.69	41.30	60.43	68.10	76.80
Population density	2730	464.36	1208	2.70	74.23	322.77	11357

Note: descriptive statistics of all variables in the study for 273 NUTS 2 region in the European Union between 2008 and 2017.

28

3.3. Space dimension

This section is devoted to the exploratory spatial data analysis which will check the existence of spatial clusters and spatial heterogeneity.



Figure 5. Natural logarithm of GDP per capita across EU NUTS 2 regions in 2017

Figure 5 shows the diversity in regional GDP per capita across NUTS 2 regions in the European Union. It is possible to notice that Germany, the Netherlands, Southern UK, and Scandinavian countries have the highest GDP per capita. Meanwhile, the Eastern European countries are the least developed with the lowest GDP per capita in the European Union. Interestingly, Italian region Nord-Est is an outlier being surrounded by regions in that part of Europe which generally score very well.

Figure 6 displays a map of population density across the same 273 regions in the EU. It is possible to draw a connection between population density and GDP per capita. More urban and dense areas are likely to have more business activity, higher efficiency and greater production.



Figure 6. Population density across NUTS 2 regions in the EU in 2017

Percentile of popdensity

0,36% 100,00%

Note: Percentiles were used to plot the data since it contains a lot of outliers. It presents a problem because outliers manipulate colours in the legend.

Figure 7 presents an independent variable number of enterprises in t-3 and still active in t across European regions. The largest cities are the hotspots and have the highest number of enterprises that are in the beginning phase of entrepreneurship (younger than 3.5 years). This is in line with academics who argue that cities offer the best opportunity for business success (Glaeser, 2011).



Figure 7. Number of three-year-old enterprises across NUTS 2 regions in the EU in 2017

Percentile of enterprises

0,0% 100,0%

Note: Percentiles were used to plot the data since it contains a lot of outliers. It presents a problem because outliers manipulate colours in the legend.

Since it is possible to identify patterns on all of these maps plotted, it is plausible to claim that there is a spatial heterogeneity present among NUTS 2 regions in the European Union (also visible from descriptive statistics in Table 2). We can also observe Waldo Tobler's First Law of Geography (1970) is here, which states that "Everything is related to everything else, but near things are more related than distant ones." Some regions which are close to other regions that score high, also score high. Distance matters and the spatial effect should lose strength across larger distances (Tobler, 1970).

We now move on to create a quantification of the spatial structure of NUTS 2 regions in the EU. In other words, a row-standardised spatial weights matrix W which is based on spatial connections between regions is created. Since there are many different types of weight matrices, Abreu et al. (2004) pointed out that one should consider every possible type of the

spatial matrix because they have different definitions of what is considered a neighbouring region. According to Smit et al. (2015, pp. 11-12), Gabriel matrix is a good choice for European NUTS 2 regions since it does not allow a region to have no neighbours. Gabriel matrix has the power of capturing remote territories and islands which are also EU NUTS 2 regions. Had I used some other contiguity matrix; those regions would have been left out since the sample includes 273 NUTS 2 region including some islands and distant territories.





Note: Spatial weight matrix links among NUTS 2 regions using the Gabriel Method. The figure only includes EU-28 members.

Figure 8 shows the results from creating a spatial weight matrix using the Gabriel Method. The sample includes 273 NUTS 2 (2016) regions from 28 EU countries excluding 8 overseas and very distant territories because they are true outliers which are not very connected to mainland Europe. The results from creating a spatial weight matrix show that there is no region without a link and 2 regions have only 1 link. The maximum number of links is 7 and there is an average of 4.13 links between every region.

Based on the suggestion from Smit (2017), in addition to the complete Gabriel matrix, I created an additional two weight matrices using a Gabriel method which separate within-

country and cross-border neighbours to check whether border effects significantly change the results. Within-country matrix is calculated using a Hadamard product of the original Gabriel matrix and an equal-dimension binary matrix containing 1s if regions belong to the same country and 0s if they do not. Cross-border matrix is then calculated as a Hadamard product of the original Gabriel matrix and the within-country matrix. Summary statistics of all three matrices are given in Table 3 (below).

	Complete Gabriel	Within-country	Cross-border Matrix
	Matrix	Matrix	
Matrix dimensions	273 × 273	273 × 273	273 × 273
Number of links	1138	673	276
Minimum links	1	1	1
Maximum links	7	7	5
Average links	4.17	2.97	0.33

Table 3. Summary statistics of weight matrices

Figure 9. Global Moran plot of the natural log of GDP per capita for 2016.



Note: Standard deviations from the mean of the natural logarithm of GDP per capita are plotted on the x-axis and the spatially lagged standard deviations from the mean of the natural logarithm of GDP per capita on the y-axis. The data is a subset of the original dataset for the year 2016 only.

After the creation of Gabriel spatial weight matrix and two sub-matrices to control for the border effect, Figure 9 presents a Moran plot of the natural log of GDP per capita on the x-axis and its spatially lagged values on the y-axis in 2016. The plot reveals that all four quadrants contain points. When looking at this graph, one should take into consideration that

the bottom-left and upper-right quadrants are the ones that represent positive spatial autocorrelation, and the opposite quadrants represent negative autocorrelation. The results from the scatterplot suggest that most of the points are located on the right side of the scatterplot, in upper- and lower-right quadrants. Unfortunately, Global Moran's I is statistically insignificant with the p-value of 0.39, which suggests that there is no spatial autocorrelation or spatial effect in this sub-sample.





Note: LISA map was constructed using R. Low-high and high-low combinations were not found.

Figure 11. LISA map for the natural logarithm of GDP per capita in the EU, by NUTS2 region for 2017.



Note: LISA map was constructed using R. Low-high and high-low combinations were not found.

After unsatisfactory results from Global Moran's I, we move on to the so-called LISA (local indicator of spatial association) plots presented in Figures 10 and 11. The cluster maps show a spatial association among EU regions based on the natural logarithm of GDP per capita in 2008 (starting year of my sample) and 2017. The plots are derived from the calculation of Local Moran's I. LISA is presented since it takes into account the significance of the local spatial clusters (Anselin, 1995). The results are presented in a coloured map and clusters are identified according to the similarity of the value with their neighbouring regions. The results suggest that there are several statistically significant high-high and low-low clusters across the EU. In 2008, low-low clusters are dispersed around the continent. Namely, Outer London – South in the UK, Umbria in Italy, Sud-Montenia, Sud-Vest Oltenia, and Centru in Romania, Epirus, Western Greece, Peloponnese, and Central Greece in Greece. Compared to 2008, 2017 contains the same low-low clusters are located in Germany, Czechia and the United Kingdom. In 2008, Jihovýchod (CZ), Berkshire, Buckinghamshire and Oxfordshire (UK), and Merseyside (UK) are the only high-high clusters in the entire EU. In 2017, Jihovýchod (CZ)

becomes insignificant and it is replaced by Hannover in Germany. Furthermore, Cheshire and Outer London - West and North West regions in the UK arise as the new high-high clusters replacing Merseyside. All in all, it is possible to notice a change in the clustering of high-high and low-low values according to the economic trends between Southern and Northern EU. It is also possible to identify that low-low values remain unchanged in countries such as Italy, Greece, Romania and Spain which have been hit by the economic crisis in 2008 and had a hard time in the process of recovery since then.

Overall, the exploratory spatial analysis shows mixed results on the existence of spatial autocorrelation and spatial heterogeneity among NUTS 2 regions. On the one hand, we have Global Moran's I which is insignificant. On the other hand, Local Moran's I and Local indicator of spatial association (LISA) show that there is a spatial effect present in the data. Nevertheless, the decision is to model and test for spatial effects in spatial regressions which are going to be presented in the next section.

4. Empirical results

This chapter will present and discuss the results of the tested econometric specification outlined in section 3.1. The first part deals with the correlation between variables and creating models based on the presence of multicollinearity between them. The second part deals with the statistical results from regression analyses.

	Log GDPpc	3-year-old enterprises	tertiary_ educ	hrst	capital	economic _activity	popdensity
logGDPpc	1						
3-year-old enterprises	0.08***	1					
tertiary_educ	0.45***	0.11***	1				
hrst	0.70***	0.14***	0.72***	1			
capital	0.37***	0.41***	0.18***	0.39***	1		
economic_ activity	0.54***	0.08***	0.49***	0.56***	0.22***	1	
popdensity	0.29***	0.00	0.27***	0.38***	0.13***	0.14***	1
R&D_ expenditure	0.13***	-0.03	0.11***	0.11***	-0.03	0.14***	0.09***

Table 4. Correlation matrix

Note: p < .001 '***', p < .01 '**', p < .05 '*'. Correlation table including time-lagged variables is reported in Appendix B.

The correlation matrix reveals that the vast majority of bivariate correlations are statistically significant. The correlation coefficient between entrepreneurship and the natural log of GDP per capita is rather small and equals to 0.08. The stock of people in human resources in

science and technology (HRST) in the high-tech sector has the highest correlation coefficient with the natural log of GDP per capita and equals 0.7. Other variables such as fixed gross capital formation and the percentage of people with tertiary education have also moderately strong correlation with GDP per capita. These results are expected and in line with the outlooks from the data section and the previous literature (Su and Liu, 2016; Lee and Hong, 2012; Blumberga et al., 2014). Independent and control variables have mild strength of correlation among each other, ranging from 0.09 to 0.72. Therefore, based on the results in Table 3, there is no apparent multicollinearity between explanatory variables.

Therefore, the first part of the analysis is going to be a panel fixed effects regression with regional and time effects estimating three. The first model is going to include all variables from the correlation matrix in Table 4. The second model is going to introduce a time-lag structure consisting of two lagged independent variables enterprises t-3 and t-6 to the first model (as described in chapter 3). Lastly, the third model is going to introduce an interaction variable between the number of enterprises in t-3 and still active in t and the population density. This way, it is possible to control for the effect of densely populated areas such as large cities on the impact of the early-stage entrepreneurship on the regional economy (Audretsch et al., 2015). The results are presented in Table 5 below.

	Dependent variable:					
	Regional Gross Domestic Product per capit (log)					
	(1)	(2)	(3)			
The number of three-year-old enterprises (log)	0.015 ^{***} (0.007)	0.017 ^{***} (0.007)	0.051 ^{***} (0.019)			
Third lag of the number of three-year-old enterprises (three-year-old enterprises three years ago) (log)		0.012 (0.007)	0.009 (0.004)			
Sixth lag of the number of three-year-old enterprises (six-year-old enterprises six years ago) (log)		0.016 (0.019)	-0.003 (0.004)			
Human Resources in Science and Technology (log)	0.138 ^{***} (0.040)	0.202 ^{***} (0.051)	0.199 ^{***} (0.052)			
Percentage of people with tertiary education (log)	-0.046 [*] (0.028)	-0.075 ^{**} (0.030)	-0.070 ^{**} (0.033)			
Gross fixed capital formation (log)	0.283 ^{***} (0.013)	0.304 ^{**} (0.019)	0.303 ^{***} (0.014)			
Percentage of economically active population (log)	0.073 (0.123)	-0.069 (0.179)	-0.068 (0.174)			
Population density (log)	-0.010	-0.018	0.080**			

Table 5. Fixed Effects 'Two-ways Within' Regressions

	(0.009)	(0.012)	(0.034)
Intramural R&D Expenditure (log)	0.465 (0.638)	-1.032 (0.547)	-0.889 (0.940)
The number of three-year-old enterprises (log) * Population density (log)			0.008 ^{**} (0.003)
Observations	2730	2724	2724
Adjusted R ²	0.108	0.187	0.191
Time-fixed effects	Yes	Yes	Yes
Region-fixed effects	Yes	Yes	Yes
LM Spatial Lag test (using the entire Gabriel matrix)	10.32***	11.79***	17.35***
LM Spatial error test (using the entire Gabriel matrix)	4.35	5.01*	6.43

Note: Significance *p<0.1, **p <0.05, ***p<0.01. Standard errors are reported in parentheses.

Table 5 presents results from the fixed effects regression with individual and time effects on the three models outlined above. The overall fit of the model which is measured by Adjusted R^2 is extremely low and close to 0.1 in all three cases. In other words, it seems that the model does not explain much of the variance in the dependent variable. Moreover, the spatial error LM test yields non-significant results, except in model 2 where it is significant at 10% level. Interestingly, LM Spatial Lag test shows statistical significance at 1% level in all three models, suggesting that I should introduce a spatial lag, which I did in the next table of results below.

Table 5 suggests that the natural log of the total number of three-year-old enterprises is positively correlated with the natural log of regional GDP per capita. Therefore, these three models suggest that the early-stage entrepreneurship positively affects regional economic growth at 1% significance level. To estimate the actual magnitude of this effect in log-log models, every percent change in an explanatory variable (x) is associated with a β percent change in the dependent variable (y). Therefore, we can say that every percent increase in the total number of three-year-old enterprises is associated with 0.015, 0.017, and 0.051 percent change in the regional GDP per capita, respectively. To translate this into a real example, Île de France had 45356 enterprises that qualify as the early-stage in 2008. However, in only 10 years this number had risen to 67618. The difference between those two numbers is 22262 and the percentage change is equal to 49.08%. Therefore, based on the results in Table 5, the regional GDP per capita of Île de France is expected to rise in these ten years by 0.74, 0.83, and 2.5 percent, respectively of the model in Table 5. Moreover, this first model shows that HRST and Gross fixed capital formation are statistically significant and have a positive effect on regional economic growth. Surprisingly, the percentage of people with tertiary education has a significant negative effect on regional growth. This is contrary to the Nobel Laureate Paul Romer's belief that knowledge serves as a driver of the economy.

Model 2 shows that the introduction of a time-lag structure does not change the results at all. The variables are insignificant, and the rest is practically unchanged compared to the previous model. In Model 3, however, the introduction of the interaction variable between

the total number of three-year-old enterprises and the population density is statistically significant at 5% level. The results in model 3 are slightly more difficult to interpret compared to the previous two models. The effect of the total early-stage entrepreneurship on the regional economic growth is conditioned by the change in population density's regression coefficient and the regression coefficient by the interaction variable. As we can see, adding the interaction variable significantly improved the explanatory power of our model to R^2 = 0.191. It is still quite small but greater than both of the previous models. The regression coefficient of the interaction variable which equals 0.008 translates to the potential of an increase in the magnitude of the effect of the entrepreneurship on economic growth conditioned by the growth in the population density. In other words, enterprises founded in densely populated areas such as cities and metropolitan areas tend to have a larger effect on the regional economy. This can be attributed to the fact that most of the large companies and their spin-offs tend to be founded in the proximity to large European megacities such as Milan, Paris, London, Stockholm, Amsterdam, Barcelona and others. What surprises me in model 3 is that the regression coefficient became statistically significant and positive which is contrary to the previous two models. Therefore, it is obvious that the interaction variable influenced the entire regression model.

All in all, results from a fixed effects regression with regional and time effects show that the early-stage entrepreneurship is positively associated with a growth in regional GDP per capita. However, exploratory spatial data analysis has shown us that there might be spatial autocorrelation and heterogeneity present in the data, and we must control for it in our regressions. Therefore, we need to include the spatial weights based on the three matrices described in section 3.3 in the regressions to control for the potential bias. The next part presented in Table 6 in the analysis introduces a spatially lagged independent variable in addition to the temporal lag.

	Dependent variable:				
-	Regional Gross Domestic Product per capita (log)				
	(1)	(2)	(3)		
The number of three-year-old enterprises (log)	0.014 ^{***} (0.004)	0.057 ^{***} (0.019)	0.015 ^{**} (0.005)		
Third lag of the number of three-year- old enterprises (three-year-old enterprises three years ago) (log)		0.011 (0.009)	0.013 (0.011)		
Sixth lag of the number of three-year- old enterprises (six-year-old enterprises six years ago) (log)		0.013 (0.008)	0.014 (0.024)		
Human Resources in Science and Technology (log)	0.231 ^{***} (0.041)	0.193 ^{***} (0.051)	0.199 ^{***} (0.061)		

Table 6. Fixed Effects 'Two-ways Within' Regressions with a Spatial Lag of Independent Variable (FESLX)

Percentage of people with tertiary education (log)	-0.088 ^{***} (0.033)	-0.072 ^{**} (0.034)	-0.120 ^{***} (0.042)
Gross fixed capital formation (log)	0.309 ^{***} (0.013)	0.360*** (0.015)	0.296*** (0.022)
Percentage of economically active population (log)	-0.077 (0.170)	-0.088 (0.170)	-0.232 (0.211)
Population density (log)	-0.010 (0.002)	0.080 ^{°°} (0.030)	-0.020 (0.014)
Intramural R&D Expenditure (log)	-1.001 (1.110)	-0.886 (1.103)	0.773 (1.219)
Spatially lagged number of three- year-old enterprises	0.016 [*] (0.009)	0.016 (0.008)	
The number of three-year-old enterprises (log) * Population density (log)		0.008 ^{**} (0.002)	
Spatially lagged number of three- year-old enterprises (within-border)			0.014 (0.008)
Spatially lagged number of three- year-old enterprises (cross-border)			0.008 (0.006)
Observations	2730	2724	2724
Adjusted R ²	0.118	0.207	0.154
Time-fixed effects	Yes	Yes	Yes
Region-fixed effects	Yes	Yes	Yes
LM Spatial lag test (using the entire Gabriel matrix)	5.03*	5.52	4.22
LM Spatial error test (using the entire Gabriel matrix)	3.9	6.93	3.03
Spatial Matrix	Entire	Entire	Split

Note: Significance *p<0.1, **p <0.05, ***p<0.01. Standard errors are reported in parentheses.

The models remain almost the same as in Table 5, except I now add three spatially lagged number of enterprises variables. Spatially lagged number of three-year-old enterprises variable refers to the spatial lag of enterprises variable using a full Gabriel matrix. Spatially lagged number of three-year-old enterprises (within-border) is a spatial lag of enterprises using a within-country weight matrix outlined in Table 3. Lastly, spatially lagged number of three-year-old enterprises (cross-border) refers to the spatial lag of independent variable enterprises using a cross-border weight matrix (Table 3). Generally, we can see a slight improvement in the explanatory power of our new estimation method, measured in Adjusted R² and insignificant spatial LM tests which suggest that my models successfully control for the spatial variation in the sample.

Compared to the previous estimation technique, Table 6 presents similar results. Model 1 contains the same variables as model 1 in Table 5, with the addition of a spatially lagged independent. The significance and regression coefficients of entrepreneurship, human resources in science and technology, the percentage of people with tertiary education, and the gross fixed capital formation remain almost equal. However, the spatially lagged number of three-year-old enterprises is statistically significant which suggests that there is a spatial effect in the sample. Moreover, it suggests that neighbouring regions tend to positively affect each other. This result is in line with the so-called spillover effect of entrepreneurship (Audretsch and Keilbach, 2008). Many newly formed enterprises tend to positively affect regional growth not only in the region they are founded in but also some neighbouring regions. In the second model, the interaction variable was added to check whether it would influence the results of the spatial regression and it did. The explanatory power of the model measured in R^2 jumped to 0.207, which makes it the highest R^2 overall. The interaction variable is statistically significant with a regression coefficient of 0.008 (same as in Table 5). The results confirm that the effect of the early-stage entrepreneurship on the regional economic growth is conditioned by the population density of the region. Regions with higher population density tend to have a higher effect of the early-stage entrepreneurship on the regional growth. Interestingly, the importance of the spatially lagged independent diminishes in the second model.

The third model includes two spatially lagged independent variables which control for border effects (Smit, 2017). This effect is split to cross-border and within-country to control for regions that interact with other countries, and regions which interact with only the ones on the domestic territory. Both spatially lagged number of three-year-old enterprises (withinborder) and (cross-border) are statistically insignificant which signals that there is no border effect in this sample. R² of the third model fell drastically compared to the second model which signals that the removal of spatially lagged independent based on the entire weights matrix and the removal of the interaction variable reduced the explanatory power. The number of three-year-old enterprises , human resources in science and technology, the percentage of people with tertiary education, and the gross fixed capital formation remain statistically significant variables.

Overall, chapter 4 presented the empirical results from three different estimation methods where I briefly interpreted how can those results be translated and applied to the real world. However, the following chapter is designed to discuss the results in greater depth and further explain their importance.

5. Discussion and research limitations

This chapter is designated to discuss the importance of results presented in chapter 4 and their applicability to the real world. Moreover, in this chapter, I attempt to answer the research questions, outline the limitations of the study, and offer some ideas on how to improve the future research conducted on this topic.

5.1. Discussion of the results

Findings in the previous chapter suggest that the total early-stage entrepreneurship measured by the number of three-year-old enterprises has a mild positive effect on the regional GDP per capita. Therefore, the results support the hypothesis outlined in the second chapter. However, models outlined in Table 5 and 6 have generally low explanatory power since the academic consensus says that a minimum R² needs to be greater than 0.19 to be considered applicable to the real world.

The analysis was conducted in three stages. In the first part, I used a bivariate correlation to check the multicollinearity between independent and control variables. I did not identify any problematic pairs in terms of multicollinearity. In the second part of the analysis, I used fixed effects with regional and time effects in three models. One model with just standard variables, one model with an addition of the two temporal lag variables of the entrepreneurship variable and the third model which includes an interaction variable between the number of enterprises and the population density. The results from the fixed effects regressions show that the early-stage entrepreneurship has a statistically significant and positive effect on the regional GDP. This result was expected and in line with the earlier literature on this topic (Capello and Lenzi, 2016). Moreover, the percentage of people out of the active population employed in human resources in science and technology has a significant and impactful effect on the regional GDP. Calculated by the method outlined in the previous chapter, 10% increase in the share of HRST in active population is expected to result in an approximately 2% increase (depending on the model) in regional GDP per capita. Since HRST variable represents human capital in the region, the evidence is in line with the previous literature which suggests that human capital positively impacts the regional economy (Audretsch and Keilbach, 2005; Faggian et al., 2019). Moreover, models in Table 5 and 6 estimate that every percent increase in capital results in approximately 0.3 percent increase in the regional GDP per capita. To put this in perspective, the Austrian NUTS2 region 'Vienna' had approximately 16.9 billion Euros of gross fixed capital formation in 2010. However, in 2016 this number had risen to €20.3 billion. This presents an increase of 20.12% in only 6 years. Therefore, our models would likely estimate a regional economic growth of 6% based on just this steep increase in the physical capital. This result is in accordance with the research that suggests that the physical capital belongs to the production function and positively affects the economy (Sala-i-Martin, 1996; Audretsch and Keilbach, 2005).

Surprisingly, the percentage of people with tertiary education is also statistically significant at 1% level and harms the regional GDP. Every 10% increase in the share of the population

holding a tertiary degree results in circa 0.7% decrease in the regional GDP. Considering that a percentage of the population with tertiary education is a proxy for the human capital, the results are undeniably opposed to the previous research covering this relationship (Coulombe and Tremblay, 2001; Faggian et al., 2019). Lastly, population density's coefficients in all models have a statistically significant effect only in models where interaction between it and the number of enterprises variable is included. This result is unexpected and differs from the literature on regional development and economic growth (Glaeser, 2011). Given that the literature claims that denser areas such as cities and metropolises have a better economic outlook, this result in the regression seems odd.

Adding a temporal lead variables to the model as part of the time-lag structure from Figure 4 does not change the results at all. The results remain consistent throughout the analysis. Therefore, the results differ from Fritsch and Mueller (2004) who claim that the peak negative effect of the newly founded enterprise on regional growth is three years, and the peak positive effect six years after its inception.

Since the exploratory spatial data analysis (ESDA) in section 3.3. and several tests suggested, there is a potential spatial autocorrelation in the sample. Moreover, ESDA hinted that there might be spatial heterogeneity. Similarly, the LM test for spatial lag dependence has resulted in a p-value with less than 0.01. All in all, the evidence suggested that we must control for the spatial effect and use the spatial lag model. Following the suggestion from Vega and Elhorst (2015) and Capello and Lenzi (2016), I decided to introduce a spatial specification SLX (Spatial lag of independent X) and created three versions of spatially weighted enterprises variable. The first one was created using a full Gabriel matrix, the second one with only within-country matrix and the third one with a cross-border matrix to control for the border effect found in Smit (2017). The results are presented in Table 6 and show that the explanatory power of our models improved slightly compared to the non-spatial regressions in Table 5. Once again, the early-stage entrepreneurship appears to be positively related to the regional GDP per capita in all three models. HRST as a proxy for human capital remains strongly and positively related to the regional economy. Spatial regressions show that the impact of human capital on the economy is even stronger compared to the previous analysis. Interestingly, the percentage of people in the region with tertiary education as another proxy for human capital in this thesis seems to be negatively related to the regional economy once again. It is contrary to the general opinion that education positively affects growth (Barro, 2001; Wolf, 2002; Sterlacchini, 2008). Analogously to the non-spatial models, gross fixed capital formation is statistically significant in all models in the spatial setting. The effect of capital on the regional economy remains almost equal. Interestingly, the economic activity rate is insignificant and does not affect the GDP. Furthermore, population density is not significant in models 1 and 3 but it is significant in model 2 when we added the interaction variable. Lastly, Intramural R&D Expenditure, time-lagged number of enterprises, spatially lagged number of three-year-old enterprises (within-border), and spatially lagged number of three-year-old enterprises (cross-border) all have an insignificant effect in the spatial analysis. To sum up, these results suggest that the time-lag structure does not apply to this case, and the border effect is not present in the sample.

All in all, the research questions outlined in the theoretical framework can be answered by the results presented in the previous chapter. The main research question asks, "to what extent can the early-stage entrepreneurship explain for regional economic growth across all NUTS-2 regions in the European Union between 2008 and 2017?". The answer is that the results of this thesis show a positive relationship between the early-stage entrepreneurship and the regional economic growth across 273 NUTS 2 regions between 2008 and 2017. The extent of this effect is measured by the regression coefficient which ranges from 0.10 to 0.50, depending on the model. In other words, every percentage increase in the number of three-year-old enterprises result in an increase of 0.1 to 0.5 percent in the regional GDP per capita.

As far as the sub-question "To what extent is entrepreneurship's contribution to the regional growth more pronounced in the places with higher population density such as large cities compared to their countryside counterparts, as suggested in the previous literature (Audretsch, Belitski and Desai, 2015)?" is concerned, the results in model 3 in Table 5 and model 2 in Table 6 show that the introduction of the interaction variable between the number of three-year-old enterprises and the regional population density significantly affected the overall regression. The interaction variable shows that the magnitude of the effect of the early-stage entrepreneurship on the regional economic growth is conditioned by the population density. In other words, entrepreneurship in the countryside has a lower impact on the regional economy than in large cities.

5.2. Limitations

This research suffers from several limitations which I will address now. The first problem of this study is the constraint on the availability of regional data in the European Union. More versatile data and more indicators should be collected on NUTS 3 and NUTS 2 level by Eurostat. This way, more precise measures can be used to check the validity of the results presented in this thesis. For instance, gross fixed capital formation is the only measure of physical capital available for European regions. Similarly, the total number of R&D staff and R&D expenditure are the only available measures for innovation. This extremely limited choice of variables that can be used pose a limitation on studies like this one since the individual collection of such data for every region in the EU is impossible. Furthermore, other measures of entrepreneurship such as Regional Entrepreneurship and Development Index (REDI) should be introduced and published on an annual basis which are much more precise for research like this compared to the current business demographics statistics available on Eurostat.

Apart from Eurostat's data availability limitation, 8 overseas and distant NUTS 2 regions were dropped from the sample to make the spatial weight matrix more accurate. Hence, one can argue that the inclusion of the remaining 8 regions might have produced different results. Additionally, the period chosen between 2008 and 2017 potentially creates a bias in the data since panel data over a longer period tends to be more accurate (Hsiao, 2007). Also, this period covers the post-economic crisis of 2008. This is an interesting period which captures the economic crisis and its aftermath which struck the European Union; however, it contains potentially biased data which are not very representative of the general trends in the early-stage entrepreneurship, the regional GDP or some other indicators present in this study. As outlined earlier, this period was chosen exclusively due to the availability of data on Eurostat. The suggestion for the future study is to use a bigger panel with more years and calculate the two- or three-year averages due to some outliers in the data. For instance, île de France which has 60 thousand early-stage enterprises or Inner London-East which has a population density of more than 10 thousand people per square kilometre. The removal of such regions might have produced a completely different result.

Next, the chosen matrix of control variables in this study potentially influences the impact and/or significance of the regional early-stage entrepreneurship on the regional economic growth. Despite all checks, such as the bivariate correlation, were done to prevent multicollinearity, it is impossible to completely control for the interaction between the variables in our models. One example would be the population density in Table 5 and 6 where we introduce an interaction variable. It immediately becomes highly significant and positively related to the dependent variable.

The most important limitation of this study is that we cannot argue anything about the causality in the relationship between the total-early stage entrepreneurship and the regional economic growth based on the results of this study. Just as Mojica-Howell et al. (2012) show, the relationship between entrepreneurship and economic growth is often bidirectional and influenced by exogenous factors which have been omitted in this study. However, we can say that our results show that the early-stage entrepreneurship supposedly causes higher regional economic growth.

6. Conclusion

In this chapter, I present conclusions, implications and recommendations for future research on this topic.

The ultimate goal of this thesis was to examine the relationship between early-stage entrepreneurship and regional economic growth across NUTS 2 regions in the European Union between 2008 and 2017. The analysis consisted of three different types of statistical analysis. The evidence suggests that there is a statistically significant and positive relationship between the number of enterprises younger than 3.5 years and economic growth in the European Union. Using the results from data analysis, I successfully answered the research questions and supported the main hypothesis of this thesis. It is a great achievement to confirm the positive link between entrepreneurship and economic growth once again. Not only that, but the results support the evidence presented by academics from other research domains. For instance, the results show that cities are more likely to have young enterprises. Moreover, young enterprises in densely populated areas have larger economic implications compared to their countryside counterparts. Also, human capital and physical capital have proven to be positively affecting regional economic growth. Lastly, there is a positive spatial spillover effect present in the data. In other words, new enterprises formed in one region are likely to be somewhat related to the economic growth in the neighbouring region. All in all, the results in this thesis satisfy results from many different disciplines such as regional growth theory, innovation theory, endogenous growth theory, knowledge spillover theory of entrepreneurship and perhaps many more.

Due to the complex nature of the relationship between entrepreneurship and regional growth, it would be incorrect to affirm that there is a causality present. However, policymakers can certainly use the results of this research. One potential policy-related problem would be to consider a larger investment in the new enterprise formation where lagging regions in the European Union would be able to offer its citizens an opportunity to start their business and become self-employed. Moreover, the European Union and its member states should work on removing the 'elitism stigma' on entrepreneurship because we see an upward trend in the number of entrepreneurs coming from the similar - highly educated and wealthy - background. Entrepreneurial opportunities should be available to everyone regardless of their education, social status, age, race, and other characteristics.

Lastly, the relationship between the early-stage entrepreneurship and regional economic growth is far from fully explored. Numerous research questions are waiting to be tackled. For instance, there is a great need for research which investigates the relationship between the early-stage entrepreneurship and regional economic growth in the EU based on the East-West perspective. That study would capture the ever-bigger desire of Eastern European citizens to become entrepreneurs. This desire can be attributed to globalization, the trendiness of entrepreneurship and cheap prices of sustaining a business. For instance, an early-stage company in Eastern Europe can sustain itself with 20,000€ throughout the entire year, while that amount in London or Paris would not even be sufficient for a 6-month rent. The study can be conducted using a longitudinal panel dataset from Eurostat and creating two subsamples based on the East-West perspective or the new-old Member state perspective, whichever the researcher prefers.

Bibliography

Brock, W. A., & Evans, D. S. (1989). Small business economics. *Small business economics*, 1(1), 7-20.

Abreu, M., De Groot, H. L., & Florax, R. J. (2004). Space and growth: a survey of empirical evidence and methods.

Acs, Z. J., & Audretsch, D. B. (1990). Innovation and small firms. Mit Press.

Acs, Z., & Storey, D. (2004). Introduction: Entrepreneurship and economic development.

Anselin, L., Varga, A., & Acs, Z. J. (2000). Geographic and sectoral characteristics of academic knowledge externalities. *Papers in regional science*, *79*(4), 435-443.

Acs, Z. J., Desai, S., & Hessels, J. (2008). Entrepreneurship, economic development and institutions. *Small business economics*, *31*(3), 219-234.

Acs, Z. J., Audretsch, D. B., Braunerhjelm, P., & Carlsson, B. (2012). Growth and entrepreneurship. *Small Business Economics*, *39*(2), 289-300.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, 27(2), 93-115.

Aparicio, S., Urbano, D., & Audretsch, D. (2016). Institutional factors, opportunity entrepreneurship and economic growth: Panel data evidence. *Technological forecasting and social change*, *102*, 45-61.

Armington, C., & Odle, M. (1982). Small business: how many jobs?. *The Brookings Review*, 1(2), 14-17.

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American economic review*, *86*(3), 630-640.

Audretsch, D. B., & Fritsch, M. (2002). Growth regimes over time and space. *Regional Studies*, *36*(2), 113-124.

Audretsch, D. B., & Keilbach, M. (2005). Entrepreneurship capital and regional growth. *The Annals of Regional Science*, *39*(3), 457-469.

Audretsch, D. B., & Keilbach, M. (2008). Resolving the knowledge paradox: Knowledge-spillover entrepreneurship and economic growth. *Research Policy*, *37*(10), 1697-1705.

Audretsch, D. B., Keilbach, M. C., & Lehmann, E. E. (2006). *Entrepreneurship and economic growth*. Oxford University Press.

Audretsch, D. B., & Keilbach, M. (2004). Entrepreneurship and regional growth: an evolutionary interpretation. *Journal of Evolutionary Economics*, 14(5), 605-616.

Audretsch, D. B. (2007). Entrepreneurship capital and economic growth. Oxford Review of Economic Policy, 23(1), 63-78.

Audretsch, D. B., Belitski, M., & Desai, S. (2015). Entrepreneurship and economic development in cities. *The Annals of Regional Science*, *55*(1), 33-60.

Audretsch, D. B., Kuratko, D. F., & Link, A. N. (2015). Making sense of the elusive paradigm of entrepreneurship. *Small Business Economics*, *45*(4), 703-712.

Audretsch, D. B. (1995). Innovation, growth and survival. *International journal of industrial organization*, *13*(4), 441-457.

Audretsch, D. B. (2002). The dynamic role of small firms: Evidence from the US. *Small business economics*, *18*(1-3), 13-40.

Audretsch, D. B. (2018). Entrepreneurship, economic growth, and geography. *Oxford Review* of *Economic Policy*, *34*(4), 637-651.

Barro, R. J. (1991). Economic growth in a cross section of countries. *The quarterly journal of economics*, *106*(2), 407-443.

Barro, R. J. (2001). Human capital and growth. American economic review, 91(2), 12-17.

Batchelor, J. H., & Burch, G. F. (2012). Creative individuals and entrepreneurship: Is there an inherent attraction?. *Small Business Institute Journal*, 8(2), 1-11.

Beugelsdijk, S., & Noorderhaven, N. (2004). Entrepreneurial attitude and economic growth: A cross-section of 54 regions. *The Annals of Regional Science*, *38*(2), 199-218.

Birch, D. L. (1981). Who creates jobs?. *The public interest, 65,* 3.

Boschma, R. (2005). Proximity and innovation: a critical assessment. *Regional studies*, *39*(1), 61-74.

Bosma, N., Hessels, J., Schutjens, V., Van Praag, M., & Verheul, I. (2012). Entrepreneurship and role models. *Journal of Economic Psychology*, *33*(2), 410-424.

Bosma, N., Hill, S., Ionescu-Somers, A., Kelley, D., Levie, J., & Tarnawa, A. (2020). Global Entrepreneurship Monitor 2019/2020 Global Report. *Global Entrepreneurship Research Association, London Business School*.

Bresnahan, T., & Gambardella, A. (Eds.). (2004). *Building high-tech clusters: Silicon Valley and beyond*. Cambridge University Press.

Capello, R., & Lenzi, C. (2016). Innovation modes and entrepreneurial behavioral characteristics in regional growth. *Small Business Economics*, *47*(4), 875-893.

Carree, M. A., & Thurik, A. R. (2008). The lag structure of the impact of business ownership on economic performance in OECD countries. *Small Business Economics*, *30*(1), 101-110.

Carree, M. A., & Thurik, A. R. (2010). The impact of entrepreneurship on economic growth. In *Handbook of entrepreneurship research* (pp. 557-594). Springer, New York, NY.

Chandler, D. (2000). Bosnia: faking democracy after Dayton. Pluto Press.

Chell, E., Haworth, J., & Brearley, S. (1991). *The entrepreneurial personality* (Vol. 16). London: Routledge.

Cohen, W. M., & Klepper, S. (1992). The anatomy of industry R&D intensity distributions. *The American Economic Review*, 773-799.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and learning: the two faces of R & D. *The economic journal*, *99*(397), 569-596.

Colombelli, A., & Quatraro, F. (2013). *The properties of local knowledge bases and entrepreneurship: Evidence from Italian NUTS 3 regions* (No. 1303). Utrecht University, Department of Human Geography and Spatial Planning, Group Economic Geography.

Coulombe, S., & Tremblay, J. F. (2001). Human capital and regional convergence in Canada. *Journal of Economic Studies*.

Crescenzi, R., Rodriguez-Pose, A., & Storper, M. (2007). The territorial dynamics of innovation: a Europe–United States comparative analysis. *Journal of Economic Geography*, 7(6), 673-709.

Dahl, M. S., & Sorenson, O. (2009). The embedded entrepreneur. *European management review*, *6*(3), 172-181.

Elhorst, J. P. (2014). Spatial panel data models. In *Spatial econometrics* (pp. 37-93). Springer, Berlin, Heidelberg.

European Commission (EC). (2013). Entrepreneurship 2020 action plan: reigniting the entrepreneurial spirit in Europe. European Commission, Brussels, Belgium.

Ertur, C., & Koch, W. (2007). Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of applied econometrics*, *22*(6), 1033-1062.

Eurostat. (2020). Eurostat database.

Eurostat. (2019). Manual on Employment Labour Force Survey.

Faggian, A., Modrego, F., & McCann, P. (2019). Human capital and regional development. In *Handbook of regional growth and development theories*. Edward Elgar Publishing.

Fischer, M. M., & Nijkamp, P. (2019). The nexus of entrepreneurship and regional development. In *Handbook of Regional Growth and Development Theories*. Edward Elgar Publishing.

Fritsch, M. (2008). How does new business formation affect regional development? Introduction to the special issue. *Small Business Economics*, *30*(1), 1-14.

Fritsch, M., & Mueller, P. (2004). Effects of new business formation on regional development over time. *Regional Studies*, *38*(8), 961-975.

Fritsch, M., & Storey, D. J. (2014). Entrepreneurship in a regional context: Historical roots, recent developments and future challenges.

Fritsch, M. (2013). New business formation and regional development: A survey and assessment of the evidence. *Foundations and trends in entrepreneurship*, 9(3).

Fujita, M., & Thisse, J. F. (2002). Agglomeration and market interaction.

Gartner, W. B., & Birley, S. (2002). Introduction to the special issue on qualitative methods in entrepreneurship research.

Glaeser, E. (2011). Triumph of the City. Pan.

Guerrero, M., Urbano, D., Cunningham, J., & Organ, D. (2014). Entrepreneurial universities in two European regions: A case study comparison. *The journal of technology Transfer*, *39*(3), 415-434.

Guerrero, M., Cunningham, J. A., & Urbano, D. (2015). Economic impact of entrepreneurial universities' activities: An exploratory study of the United Kingdom. *Research Policy*, *44*(3), 748-764.

Hart, D. M., & Acs, Z. J. (2011). High-tech immigrant entrepreneurship in the United States. *Economic Development Quarterly*, 25(2), 116-129.

Hill, R. C., Griffiths, W. E., & Lim, G. C. (2018). *Principles of econometrics*. John Wiley & Sons.

Headey, D. D., & Hodge, A. (2009). The effect of population growth on economic growth: A meta-regression analysis of the macroeconomic literature. *Population and Development Review*, *35*(2), 221-248.

Herwartz, H. (2006). Testing for random effects in panel data under cross sectional error correlation—A bootstrap approach to the Breusch Pagan test. *Computational statistics & data analysis*, *50*(12), 3567-3591.

Heshmati, A. (2001). On the growth of micro and small firms: evidence from Sweden. *Small business economics*, *17*(3), 213-228.

Hohti, S. (2000). Job flows and job quality by establishment size in the Finnish manufacturing sector 1980–94. *Small Business Economics*, *15*(4), 265-281.

Kirzner, I. M. (1997). Entrepreneurial discovery and the competitive market process: An Austrian approach. *Journal of economic Literature*, *35*(1), 60-85

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

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

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

Konings, J. (1995). Job creation and job destruction in the UK manufacturing sector. *Oxford Bulletin of Economics and Statistics*, *57*(1), 5-24.

Kortum, S., & Lerner, J. (1997). *Stronger protection or technological revolution: what is behind the recent surge in patenting?* (No. w6204). National Bureau of Economic Research.

Krugman, P. R. (1991). *Geography and trade*. MIT press.

Lam, C., & Qian, C. (2017). Spatial lag model with time-lagged effects and spatial weight matrix estimation.

Leamer, E. E. (1997). Access to western markets, and eastern effort levels. In *Lessons from the Economic Transition* (pp. 503-526). Springer, Dordrecht.

Lee, L. F., & Yu, J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154(2), 165-185.

Lee, S. Y., Florida, R., & Acs, Z. (2004). Creativity and entrepreneurship: A regional analysis of new firm formation. *Regional studies*, *38*(8), 879-891.

LeSage, J. P., & Pace, R. K. (2011). Pitfalls in higher order model extensions of basic spatial regression methodology. *Review of Regional Studies*, 41(1), 13-26.

Liñán, F., & Fernandez-Serrano, J. (2014). National culture, entrepreneurship and economic development: different patterns across the European Union. *Small Business Economics*, *42*(4), 685-701.

Lockett, A., & Wright, M. (2005). Resources, capabilities, risk capital and the creation of university spin-out companies. *Research policy*, *34*(7), 1043-1057.

Low, M. B. (2001). The adolescence of entrepreneurship research: Specification of purpose. *Entrepreneurship theory and practice*, *25*(4), 17-26.

Lucas, R. E. (1988). On the mechanics of economic development. *NBER Working Paper*, (R1176)

Malecki, E. J. (1997). Entrepreneurs, networks, and economic development: A review of recent research. *Advances in entrepreneurship, firm emergence and growth*, *3*, 57-118.

Marshall, A. (1920). Industry and trade.

Marx, K. (1867). Das Kapital: Volume 1 (Vol. 6). William Milner.

Mitra, J. (2020). On Regarding the Value of Aesthetics for Entrepreneurship. In *Innovation and the Arts: The Value of Humanities Studies for Business*. Emerald Publishing Limited.

Mojica-Howell, M. N., Whittaker, W. L., Gebremedhin, T. G., & Schaeffer, P. V. (2012). Examining the bidirectional relationship between entrepreneurship and economic growth: Is entrepreneurship endogenous?. *Entrepreneurship: Born, made and educated*, 283-316.

Moreno, R., Lopez-Baso, L., & Artis, M. (2004). *Infrastructure investments as a tool for regional development policy: Lessons from the Spanish evidence*. mimeo.

Mueller, P. (2007). Exploiting entrepreneurial opportunities: The impact of entrepreneurship on growth. *Small Business Economics*, *28*(4), 355-362.

Nelson, R. E., & Leach, J. A. (1981). Increasing opportunities for entrepreneurship. *Contemporary challenges for vocational education. Ohio: American Vocational Association*, 19-32.

OECD. (2019). SME and Entrepreneurship Outlook.

OECD. (2014). Gross domestic product (GDP) (indicator).

OECD. (2015). *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*: The Measurement of Scientific, Technological and Innovation Activities. OECD Publishing, Paris.

OECD. (2020). OECD Economic Outlook.

O'Reilly, M., & Hart, M. (2005). Global Entrepreneurship Monitor Northern Ireland, 2004".

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of political economy*, *94*(5), 1002-1037.

Porter, M. E. (1990). The competitive advantage of nations. *Harvard business review*, 68(2), 73-93.

Porter, M. E., & Stern, S. (2001). Innovation: location matters. *MIT Sloan management review*, 42(4), 28.

Reynolds, P. D., Hay, M., & Camp, S. M. (1999). Global entrepreneurship monitor. *Kansas City, Missouri: Kauffman Center for Entrepreneurial Leadership*.

Reynolds, P., Bosma, N., Autio, E., Hunt, S., De Bono, N., Servais, I., ... & Chin, N. (2005). Global entrepreneurship monitor: Data collection design and implementation 1998–2003. *Small business economics*, *24*(3), 205-231.

Reynolds, P., Storey, D. J., & Westhead, P. (1994). Cross-national comparisons of the variation in new firm formation rates. *Regional studies*, *28*(4), 443-456.

Reynolds, P., Camp, S.M., Bygrave, W.D., Autio, E., Hay, M. (2001). Global entrepreneurship monitor: 2001 Executive report. Kauffman center for entrepreneurial leadership.

Reynolds, P., Bygrave, W.D., Autio, E., Cox, L., Hay, M. (2002). Global entrepreneurship monitor: 2002 Executive report.

Roberts, E. B., & Wainer, H. A. (1971). Some characteristics of technical entrepreneurs. *IEEE Transactions on Engineering Management*, (3), 100-109. Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, *70*(1), 65-94.

Runco, M. A., Noble, E. P., Reiter-Palmon, R., Acar, S., Ritchie, T., & Yurkovich, J. M. (2011). The genetic basis of creativity and ideational fluency. *Creativity Research Journal*, *23*(4), 376-380.

Solow, R. M. (1957). Technical change and the aggregate production function. *The review of Economics and Statistics*, 312-320.

Roberts, E. B. (1991). *Entrepreneurs in high technology: Lessons from MIT and beyond*. Oxford University Press.

Sala-i-Martin, X. X. (1996). Regional cohesion: evidence and theories of regional growth and convergence. *European economic review*, *40*(6), 1325-1352.

Saviotti, P. P., & Pyka, A. (2004). Economic development by the creation of new sectors. *Journal of evolutionary economics*, *14*(1), 1-35.

Scherer, F. M. (1970). Industrial market structure and economic performance. Chicago: Rand McNally.

Schumpeter, J. A. (1934). The theory of economic development, translated by Redvers Opie. *Harvard: Economic Studies, 46,* 1600-0404.

Smit, M. J. (2017). Cross-border agglomeration benefits. *Letters in Spatial and Resource Sciences*, 10(3), 375-383.

Smit, M. J., van Leeuwen, E. S., Florax, R. J., & de Groot, H. L. (2015). Rural development funding and agricultural labour productivity: A spatial analysis of the European Union at the NUTS2 level. *Ecological Indicators*, *59*, 6-18.

Stearns, T. M., Carter, N. M., Reynolds, P. D., & Williams, M. L. (1995). New firm survival: industry, strategy, and location. *Journal of business venturing*, *10*(1), 23-42.

Sterlacchini, A. (2008). R&D, higher education and regional growth: Uneven linkages among European regions. *Research Policy*, *37*(6-7), 1096-1107.

Sternberg, R. (2009). *Regional dimensions of entrepreneurship*. Now Publishers Inc.

Storey, D. J., & Johnson, S. (1987). *Job generation and labour market change*. Springer.

Szerb, L. A., Acs, Z., & Autio, E. (2013). Entrepreneurship and policy: The national system of entrepreneurship in the European Union and in its member countries. *Entrepreneurship research journal*, *3*(1), 9-34.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, *46*(sup1), 234-240.

UN High Commissioner for Refugees (UNHCR). (2016). *The Sustainable Development Goals and Addressing Statelessness*.

U.S. Census Bureau (2012). Statistics of U.S. Businesses Employment and Payroll Summary: 2012. Retrieved from [https://www.census.gov/library/publications/2015/econ/g12-susb.html].

van Praag, C. M., & Versloot, P. H. (2007). What is the value of entrepreneurship? A review of recent research. *Small business economics*, *29*(4), 351-382.

van Buuren S and Groothuis-Oudshoorn K (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3), 1-67. https://www.jstatsoft.org/v45/i03/

Varga, A. (2006). The spatial dimension of innovation and growth: Empirical research methodology and policy analysis. *European Planning Studies*, *14*(9), 1171-1186.

Vega, S., & Elhorst, J. P. (2015). The SLX model. Journal of Regional Science, 55(3), 339-363.

Verheul, I., Wennekers, S., Audretsch, D., & Thurik, R. (2002). An eclectic theory of entrepreneurship: policies, institutions and culture. In *Entrepreneurship: Determinants and policy in a European-US comparison* (pp. 11-81). Springer, Boston, MA.

Weiss, L. W. (1976). Optimal plant scale and the extent of suboptimal capacity. *Essays on Industrial Organization in Honor of Joe S. Bain, Cambridge, MA: Ballinger,* 126-134.

Wennekers, S., & Thurik, R. (1999). Linking entrepreneurship and economic growth. *Small business economics*, *13*(1), 27-56.

Wolf, A. (2002). *Does education matter?: Myths about education and economic growth*. Penguin UK.

Appendix A

Figure 12. Comparison of a distribution of the original and imputed independent variable – the number of three-year-old enterprises.



Note: Original variable is in red colour and imputed variable is in blue. Obviously, the distribution has not changed.

Appendix B

	Log GDPpc	3-year-old enterprises	tertiary_ educ	hrst	capital	economic_acti vity	popdensity	R&D expenditure	Third lag of 3- year-old enterpris es
logGDPpc	1								
3-year-old enterprises	0.08***	1							
tertiary_educ	0.45***	0.11***	1						
hrst	0.70***	0.14***	0.72***	1					
capital	0.37***	0.41***	0.18***	0.39***	1				
economic_ activity	0.54***	0.08***	0.49***	0.56***	0.22***	1			
popdensity	0.29***	0.00	0.27***	0.38***	0.13***	0.14***	1		
R&D expenditure	0.13***	-0.03	0.11***	0.11***	-0.03	0.14***	0.09***	1	
Third lag of 3-year- old enterprises	0.012*	0.332***	0.054***	0.190***	0.218***	0.081***	0.001	-0.32***	1
Sixth lag of 3-year- old Enterprises	-0.002	0.164***	-0.05	0.006	0.096***	0.022	-0.032*	-0.44***	0.330***

Table 7. Correlation matrix including time-lagged variables

Appendix C





Figure 14. Scatterplot of the number of three-year-old enterprises [enterprises] and its sixth [enterprises6] lag.

