ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics
Master Thesis Strategy Economics

The relationship between entrepreneurship and innovation in European countries between 2000 and 2020.

Student: Brandligt, M.L.A.J.B.
Student Number: 457413
Supervisor: Hessels, S.J.A.
Second Assessor: Bhaskarabhatla, A.S.
Date final version: 21-12-2022

Abstract

In this research the relationship between entrepreneurship and innovation is assessed for European countries over the 2000-2020 interval. The potential mechanisms that drive this effect are knowledge spillovers and competitiveness. Furthermore, the research tests whether there are existing moderating effects of the motivation of the entrepreneurs and governmental support for entrepreneurs. The proxy for innovation that is used is the number of patent applications by residents, which is retrieved from the World Bank Data. While the variables about entrepreneurship are retrieved from the GEM database. The main results of this paper show that early-stage entrepreneurship has a negative and significant relationship with innovation, whereas established business owners have a positive and significant relationship with innovation.

Table of Contents

1. Introduction

Around the start of 2020 a virus called COVID-19 spread around the world, which had an enormous impact on the day-to-day life, especially for entrepreneurs. In this research entrepreneurship is defined as the act of an individual who sets up a business in the hope of making a profit, while taking on a financial risk. While, the entrepreneurs that were active in the e-commerce and technology market, did not experience an extremely negative impact from the lockdowns, this negative impact was more prevalent for others, such as restaurants and hotels. To predict on a large scale the number of failures governments started spending a lot of money on support for entrepreneurs that were hit hard due to the COVID-19 crisis. But why would a government be so involved in keeping a certain level of entrepreneurial activity? The first thing that comes to mind is a potential relationship between entrepreneurship and employment. However, the effect of entrepreneurship on employment seems to be ambiguous according to Audretsch, Carree and Thurik (2001). Because there are diverse ways in which entrepreneurship and employment interact with each other. First, unemployment can have a positive effect on entrepreneurial activity as individuals need a job. On the other hand, higher levels of entrepreneurial activity can also be induced by higher economic growth levels in a country, which in turn will also lead to lower levels of unemployment. Another potential reason could be the importance of entrepreneurship on the innovative ability of a country and governments are mostly interested in keeping those entrepreneurs on the long-term. The concept of innovation can be described as the introduction of a new product in the case of product innovation, which could be either radical or incremental. Radical innovation means the introduction and commercialization of a completely new product, while incremental innovation is the improvement or diversification of an already existing product. On the other hand, there is also process innovation, which is the improvement of the processes in for example the manufacturing of a product. However, it is likely that the entrepreneurs that needed the governmental support are new and smaller businesses as former literature has found a lower survival probability for such firms (Cefis and Marsili, 2006). On top of that smaller firms also seem to contribute less to innovation on country-level (Wong, Ho and Autio, 2005). These views are confirmed by Hagedoorn (1996), who revised Schumpeter's work who found evidence that larger firms are the main drivers of innovation. Based on these findings it could be interesting to research the relationship between entrepreneurship and innovation, to determine whether the results hold. Schumpeter (1934) suggested that there is a relationship between entrepreneurship and innovation. However, more recent academic research showed ambiguous results. Wennekers, van Stel, Thurik, & Reynolds, (2005) found a negative relationship between entrepreneurship and innovation. While Cefis & Marsili (2006) investigated the role of firm survival and innovation, they found that innovation contributes to

the probability of survival. This could mean that innovation is a common strategy for firms to enhance the chance of survival. Based, on these findings it is reasonable to suggest that there might be a relationship between entrepreneurship and innovation. However, most of these papers are written in the early 2000's, so it might be insightful to investigate whether these effects are still prevalent. Additionally, the importance of entrepreneurs that own small firms, versus entrepreneurs that own large firms, is interesting to explore. Furthermore, a lot of former research is focused on one country or worldwide, while Europe is one of the most developed continents. For this reason, it could be relevant to use Europe for academic research on entrepreneurship. This could give meaningful insights which might have been overlooked in earlier research. Therefore, the following research question is formulated:

Does a larger share of individuals active in any stage of entrepreneurship have a positive effect on innovation in Europe for the time interval, 2000-2020?

To find an answer to this research question multiple hypotheses will be evaluated in this paper. The main hypothesis will analyze the overall relationship between entrepreneurship, and innovation. Thereafter, multiple sub-hypotheses will be assessed. First, the differences between early-stage entrepreneurial activity, hereinafter 'TEA', and established business owners will be tested. Former literature has found evidence that innovation is mainly driven by larger firms (Hagedoorn, 1996). Second, the motivation to get involved in entrepreneurship will be reviewed. This moderating effect has been assessed previously by Mrożewski & Kratzer (2017). They tested it for 96 countries using a standard OLS analysis, whereas this research aims to answer the hypothesis using European countries in a panel analysis. Finally, the effect of governmental interference with entrepreneurs will be tested as former literature has found evidence that this might reduce entrepreneurial activity (Aidis, Estrin, & Mickiewicz, 2012). The proxy for innovation that will be used is the number of patent applications by residents, which is retrieved from the World Data Bank. Moreover, the variables about entrepreneurship are retrieved from the GEM consortium database. The control variables that are used are also acquired through either the World Data Bank, or the GEM consortium database. The data that is used for this research is longitudinal data. As a result of this, it needs to be determined whether to use fixed or random effects in the model. The method of a Hausman-test was used to find that a random effects model would be most appropriate for the data sample used. This could be due to the time invariant nature of the main variables that are included in the regressions. What is found in this paper is that there is a positive and significant relationship between the share of established business owners and the number of patent applications by residents. For TEA entrepreneurs the results were negative and

significant which implies that there is a relationship between TEA and innovation. Besides, the findings show that there is no moderating effect for the motivation index. While the results to determine whether there is a moderating effect of the governmental support for entrepreneurship has significant results.

The findings of this research could be both scientifically and socially relevant. There is already prior research to the relationship between entrepreneurship and innovation. However, the scope of this research is focused on European countries in the 2000's, while most prior academic research was focused on worldwide effect. Conducting this research just on Europe might give new insights, as Europe is known to be one of the most developed continents in the world. These insights could potentially be socially relevant as well. In the first place, this paper will focus on whether entrepreneurship contributes towards innovation. Additionally, the effect of governmental support for entrepreneurship on country-level will be tested in this research. The results that are retrieved from the empirical analysis on governmental support could give some insightful observations for policymakers.

In the next section relevant literature on these topics will be reviewed. Thereafter, the data that is used will be discussed and the transformations that are made to acquire a useable dataset will be featured. Then the methodology will be explained and the reasoning behind the use of panel data. This will be followed with the result section, where the estimates that are retrieved from the empirical analysis are discussed. And finally, concluding remarks will be made, and limitations of this research will be discussed.

2. Literature Review

In former literature they have already found a positive relationship between entrepreneurship and innovation. Hagedoorn (1996) reassessed and compared current literature to the work of Schumpeter, who said that innovation is mainly driven by larger companies. Hagedoorn (1996) found that even in the twentieth century the economic theory of Schumpeter still holds. Additionally, TEA seems to be positively related with innovation, but is determined by the initial motive to get involved in entrepreneurship (Mrożewski & Kratzer, 2017). In their paper they split up entrepreneurship in two distinct types, necessity entrepreneurship and opportunity entrepreneurship, as they believe the outcomes of innovation differs between these two types of entrepreneurships. Opportunity driven entrepreneurs are defined as entrepreneurs who are driven by an opportunity and voluntary choose to get involved in entrepreneurship, while necessity entrepreneurship are those individuals that chose to become an entrepreneur because of the lack of an alternative. The differences that Mrożewski & Kratzer (2017) found

show that a higher level of opportunity entrepreneurs, positively influence the level of innovation, while the opposite result holds for necessity entrepreneurs. This is just one example of how entrepreneurship influence the level of innovation in a country. In this literature review multiple reasons and mechanisms that could help foster innovation through entrepreneurship will be discussed. Thereafter the explicated literature on the subject will be used to form multiple hypotheses, to evaluate whether these findings also hold for European countries in the 21st century.

2.1 Firms Need Innovation to Survive

For entrepreneurs that are active in competitive markets it is assumably necessary to find certain ways to distinct themselves from others to survive. One technique that is cited in former literature is that firms should innovate to create some sort of completive advantage that they could exploit to increase the chances of survival. In the following subsection multiple relevant papers about this topic will be explicated.

The first paper that is consulted, to prove that the statement above is right is the paper of Cefis and Marsili (2006). This paper shows that for Dutch manufacturing firms innovation leads to a higher probability of surviving. Furthermore, they found that this effect is stronger over longer periods of time. Lastly, the effect was most prevalent within TEA, which makes sense since TEA could be considered more entrepreneurial. The definition of entrepreneurship that is described in the introduction is especially prevalent for TEA, as they take the highest risk. The effects of innovation as a tool to increase the chance of survival are extra important for small and young firms. Cefis and Marsili (2006) found evidence that small and young firms can achieve the highest innovation premium, which contributes to a survival probability that is comparable to established firms. However, there are different dimensions of innovation that might contribute to a higher survival rate such as products, processes, markets, and the management strategy of a firm (Klippel, Petter & Antunes, 2008).

It is found that innovation is exploited to enhance the chance of survival. To achieve a more detailed view on why innovation is a key part, it is important to review the characteristics which enables these advantages of innovation. There are multiple firm characteristics that play a role in the effectiveness of innovation for firm survival. Cefis and Marsili (2006) show in their paper that the survival probability of innovating firms compared to non-innovating firms is higher. Additionally, they show that firm age plays a role in the effect of innovation on the probability of survival, and that the effect seems to be more prevalent for younger firms. On the other hand, Kahn and Candi (2021) investigated whether the effect differs among firms of different

sizes, and the types of products or services a firm offers. They have found that these characteristics moderate the effect of innovation on firm survival. As small product firms mainly benefit from an exploitation strategy, however when firms grow an exploration strategy seems to be more appropriate. Whereas, it is found that for large product firms both strategies seem to be beneficial. An exploring strategy could be redefined as a strategy that focusses on radical innovation, while an exploiting strategy is more focused on incremental innovation. The difference that are found by Kahn and Candi (2021) were in some way already proven by Audretsch (1995), who found that there are significant differences among industries and their influence of innovation on the likelihood of survival. What Audretsch (1995) has found is contrary to the findings of Cefis and Marsili (2006), they argue that new entrants only have a short amount of time to prove themselves, as it is a tradeoff between time and costs. In markets where there are more barriers for entry, new firms aim to penetrate the market more rapidly, which induces that the chances of survival are lower. Audretsch (1995), shows that businesses are shaped post-entry and will either be able to offer a proper product and survive the first years, or will be unable to and exit the market. The firms that have survived the first stage of entrepreneurship by successfully implementing a new product will experience opportunities. However, whenever a firm succeeds to survive in such market the relationship between innovation and survival seems to reverse, and the firms that succeeded in the first place are now more likely to survive in the long term. The relationship between entry and innovation is further explored by Turut and Ofek (2012), who investigate whether it is the better choice for a new incumbent to take part in radical or incremental innovation. Radical innovation seems to be existent in every industry and is often accompanied with higher uncertainty. Nevertheless, they also find that the new incumbent has a higher chance of making more profit. Turut and Ofek (2012) have shown is that even when a new entrant sees an opportunity and has the available resources to radically innovate, they do not always exercise this possibility. Therefore, a higher share of TEA does not necessarily have to implicate higher levels of innovation as not all new incumbents choose to involve in innovation. In conclusion, the findings of this paragraph former literature shed a light on the relationship between firm size, age and their innovative capacity. This paper will finally aim to conduct research on the relationship between innovation and both early-stage entrepreneurship and established business ownership on country-level. Even though earlier academic literature argues that innovative start-ups have a higher chance of survival, this does not have to imply that a larger share of individuals active in early-stage entrepreneurship will induce a flourishing innovative environment. Anokhin & Wincent (2012) even shed a light on the fact that there is no evidence of the contribution of TEA on the levels of innovation in a country. Wong et al. (2005) argue that TEA does not contribute to economic growth at all. Because of this, a larger share of early-stage entrepreneurship does not have to imply anything about the innovativeness of the

entrepreneurs. Therefore, it could be interesting for this research to also consider the findings of Mrożewski & Kratzer (2017), who divide entrepreneurship in two distinct types as mentioned before.

2.2 Firm Characteristics

Next, it might be insightful to figure out which mechanisms could drive a potential relationship between entrepreneurship and innovation on a country-level. The main characteristics of entrepreneurship that will be used are (1) whether the share of firms that are new or already established and (2) the share of TEA that is driven by opportunity and necessity.

The first characteristic that will be focused on is the firm age and its effect on innovation performance. However, it is questionable whether firm age itself can be regarded as a characteristic of a firm, but there are some other features that appear at various stages of a firms' life cycle. As is shown in former literature, firm size indeed influences the rate of innovation, which is in line with the expectation as younger firms are often smaller firms. Huergo and Jaumandreu (2004) find that the probability that a firm innovates is related to both the size and age of a firm. They find that small manufacturing firms have a lower probability to innovate in comparison to larger firms. They suggest that this is the case due to the fixed costs of expenditures that must be made to innovate, which might be too expensive for small firms, but become workable if a firm is able to scale. However, entering firms have a seemingly higher propensity to actively innovate, in comparison to firms that stay small and larger firms, which is tested for the manufacturing industry. This contradicts the findings regarding firm size, which could indicate that small firms that enter the market have additional capabilities that compensates the size handicap. They also state that the potential mechanism behind the fact that bigger and older firms tend to be more innovative, could be the fact that the firms that were not innovative enough lacked the capabilities that are necessary for a firm to survive. This mechanism could be explained by the paper of Mrożewski & Kratzer (2017), who said that there are two types of entrepreneurs, which will be discussed in more detail for the second hypothesis. In line with the findings presented above, Hansen (1992) similarly has found that both firm size and age are related to the number of new products that are introduced to the market. Because firms that enter the market seem to have an advantage in innovative capacity when comparing to firms that already have been in the market for a while. It could be interesting to test, whether there is a difference between the share of early-stage entrepreneurs in comparison to established business owners in Europe. However, earlier academic papers mainly focused on the micro level, which gives insight in why a larger share of entrepreneurs could generate higher levels of innovation. The aim of this research is to

investigate whether there is a relationship between entrepreneurship and innovation on country-level. So, to build towards a solidified answer potential mechanisms that could foster innovation on country-level through higher levels of entrepreneurship have to be discussed as well. A possible theory could be the effect of knowledge spillovers. Audretsch and Keilbach (2007) investigated the relationship between knowledge spillovers and entrepreneurship and have found evidence that in regions with higher levels of knowledge, there is more entrepreneurial activity in the high-technology and ICT sector. However, their results also suggest that there is no statistically significant evidence that higher levels of knowledge in a region foster the share of general and low-technology entrepreneurship. Based on these findings and the fact that Europe is one of the most developed continents it could be suggested that higher levels of entrepreneurship have a positive relationship with innovation, as it is likely that a high share of the entrepreneurs are either active in the high-technology or ICT sector. Moreover, former literature has also proven that the recruitment of technicians and highly educated workers can also enhance firm productivity for the hiring firm (Parrotta and Pozzoli, 2012). Another factor that might increase the level of innovation is the competitiveness in a country. Tang (2006) found that the constant arrival of competing products and the quick obsolescence of products is positively related with R&D expenditure, and therefore it can be suggested that in a competitive entrepreneurial environment innovation will foster.

Summarizing the findings of this paragraph it is found on firm level that innovation plays a key role in the chance of survival, and therefore pursuing high levels of innovation could be a strategy that a lot of entrepreneurs want to apply. Although on the country level, factors as knowledge spillovers and competitiveness play a vital role in the levels of innovation. This could mean that a higher share of entrepreneurs could be related to higher levels of innovation. Therefore, the following hypotheses have been formulated:

H1: There is a positive relationship between entrepreneurship and innovation for European countries between 2000 and 2020

H1A: The positive relationship between entrepreneurship and innovation is impaired when there is more early-stage entrepreneurial activity for European countries between 2000 and 2020.

H1B: The positive relationship between entrepreneurship and innovation is enhanced when there are more established businesses for European countries between 2000 and 2020.

The second characteristic that will be discussed is the reason for an individual to get involved in entrepreneurship. As described earlier Mrożewski & Kratzer (2017), who evaluated whether the innovative outcomes differ for firms that are led by either an 'opportunity' or 'necessity' entrepreneur. They suggest that entrepreneurs that are opportunity driven are more likely to contribute to innovation than entrepreneurs that are necessity driven. In their paper they defined the two types of entrepreneurs based on an executive report by Reynolds et al. (2002). Necessity entrepreneurship is defined as individuals which choose to get involved in entrepreneurship without the consideration of any entrepreneurial opportunity. On the other hand, opportunity entrepreneurship is described as individuals which voluntarily get involved in entrepreneurship to exploit an opportunity which they have observed, even if there are other alternatives present. Mrożewski & Kratzer (2017) argue that it is unlikely that necessity entrepreneurship will contribute to innovation as those entrepreneurs are becoming an entrepreneur because they would be unemployed otherwise. Therefore, it could be better described as self-employment. Whereas opportunity entrepreneurs have the desire to improve their status, either economically, socially, or mentally, and they are more often growth oriented. Consequently, as described earlier high-technology entrepreneurship is fostered by regions with higher levels of knowledge, while this is not the case for low-technology entrepreneurship (Audretsch and Keilbach, 2007). Therefore, it becomes even more likely that regions with higher levels of opportunity entrepreneurship are also the regions with more high-technology entrepreneurship. It is also argued that necessity entrepreneurs are likely to be individuals who tend to have lower levels of human capital (Mrożewski & Kratzer, 2017). Based on these findings a larger share of entrepreneurs who are opportunity driven in a country should result in higher levels of innovation. Furthermore, they show in their paper that business creation and technology innovation can be regarded as two distinct phenomena. Therefore, it might be interesting for this research to test whether this is also the case for European countries. Distinguishing between those two types of entrepreneurs might be valuable, as it is reasonable to assume that the entrepreneurs that are 'opportunity' entrepreneurs have a higher probability to contribute to higher levels of innovation. Therefore, the second hypothesis is formulated as shown below:

H2: The positive relationship between entrepreneurship and innovation is enhanced if the motivational index is higher for European countries between 2000 and 2020.

In earlier academic literature the influence of the capabilities of the entrepreneur on the innovative outcomes of a firm are well discussed. First, not every individual that sees certain opportunities is willing to become an entrepreneur. The choice of becoming an entrepreneur is related to certain factors as described earlier. As certain industries are more attractive to

enter in comparison to others. Audretsch (1995) showed that there are significant differences between industries and the probability that a startup will succeed in those industries. Moreover, former literature also has shown that whenever an entrepreneur sees an opportunity, it is dependent on the beliefs of the potential entrepreneur. Shane and Venkataraman (2000) state that overly optimistic entrepreneurs are more likely to exploit their ideas, and maybe overestimate the commercial value of their ideas. On the other hand, individuals can also exploit an opportunity within the firm they are already active in or sell their idea to a third party. These findings can also be of value for the research that will be conducted in this paper.

2.3 Differences of Innovative Capabilities between Countries

This research will focus on country level innovation and therefore it is important to figure out which differences in innovative capabilities are present between countries. For example, universities or technology hubs can have an impact on the probability of the occurrence of innovations in certain regions. Consequently, to get proper results from the empirical analysis it is necessary to include those factors that might influence the level of innovation for different countries. Apart from the potential relevance of entrepreneurship on innovation, it is also widely accepted by policy makers as a key source of economic growth (Sobel and King, 2008). For this reason, policy makers already focus on enhancing entrepreneurship which can induce differences between the entrepreneurial environment for those counties.

Not all countries have the same capabilities regarding innovation. Furman, Porter and Stern (2002) discuss multiple determinants that affect those capabilities. They define the concept of national innovative capacity, as the ability to produce and commercialize innovations over the long term. The determinants that they discuss are the common innovation infrastructure, the industrial cluster innovation infrastructure, and the linkage between those two infrastructures. They suggest that public policy is a key tool to shape a country's innovative capacity. However, these policies are not restricted to the availability of research and development resources but are also focused on encouraging the investment in human capital by investing in universities and create a more competitive environment. The next section will look further into the matter of governmental support for entrepreneurship and a possible moderating effect.

2.4 Applicable Strategies to Foster Entrepreneurship

If the result of this paper shows that there are indeed positive effects of entrepreneurship on the level of innovation, it could also be valuable to discuss different actions that can be taken

to use this positive effect to its full potential. One of the aims of this research is to investigate the moderating relationship of governmental support on innovation through an improved entrepreneurial environment. This paragraph will therefore highlight various actions that governments could take to enhance the level of entrepreneurship. In Europe there is an organization which aims to support entrepreneurship, The European Institute of Innovation & Technology (hereinafter 'EIT'), which is a body of the European Union (Entrepreneurship, n.d.). Their aim is to empower entrepreneurs and innovators by (1) helping them developing their new business, (2) specialized business support and (3) accelerate innovations. However, former literature seems to have questioned the functioning of such organizations as they might push the wrong individuals into entrepreneurship.

First, the size of the government is considered, as policymakers aim to create a better entrepreneurial environment, while it does not seem to be the most effective way. Aidis et al. (2012) found that a larger government seems to be negatively related with entrepreneurial entry. They argue that policy makers should focus on minimizing corruption as this has a negative effect on entrepreneurial activity. However, it is not likely that they can entirely solve corruption by more bureaucratic procedures, as corruption is often more deeply embedded in society. Aidis et al. (2012) also state that to create a healthier environment for entrepreneurship the fair legal justice system needs to be transparent. Furthermore, simple and clear regulations are also favorable for entrepreneurship (Parker, 2007). As many of the countries that are included in his research are members of the European union, corruption might not be an enormous issue. On the other side, bureaucratic hassles might affect the innovativeness of the entrepreneurs.

Another action that is taken by the EIT is investing in the educational attainment of entrepreneurs in the European Union. In a study on schooling, it is found that countries that employ a voucher system do score significantly higher in terms of youth entrepreneurship (Sobel & King, 2008). A voucher system means that a student can freely choose a study or course to their own liking, such as the STAP budget in the Netherlands, which could be the mechanism that motivates students. The programs that enhance the rates of youth entrepreneurship, however, are not focused to push the students into entrepreneurship, but are focused on a business-like environment, and stimulating competitiveness and innovation. Their findings could be an excellent choice to review whenever the findings of this paper show that entrepreneurship contributes to innovation.

On the other hand, Shane (2009) places a sidenote on the urge of policy makers to implement policies which enhance entrepreneurship. He argues that this is not always the best strategy

because the majority of start-ups is not innovative and does not create much employment nor does it contribute to overall wealth. The findings of Shane (2009) could be complementary to the findings of Mrożewski and Kratzer (2017), who found significant differences between 'necessity' and 'opportunity' entrepreneur. Thus, it might be good to test whether government policies did indeed push the wrong individuals into entrepreneurship over the last two decades. Therefore, a third hypothesis will be tested to determine whether these statements are true for European countries over the last two decades. The third hypothesis is formulated as follows:

H3A: A higher level of governmental support on entrepreneurship impairs the positive relationship between early-stage entrepreneurship and innovation for European countries between 2000 and 2020.

H3B: A higher level of governmental support on entrepreneurship impairs the positive relationship between established business ownership and innovation for European countries between 2000 and 2020.

If the results of these hypotheses come out to be positive, it might give relevant insights for governments to adjust governmental policies and support for entrepreneurs. Such a system could strive to distinguish the two types of entrepreneurs and make the 'opportunity' entrepreneurs eligible for support. However, such a system will never be perfect it might at least improve the current situation.

3. Data

The main purpose of this research is to investigate the influence of entrepreneurship on the level of innovation for European countries between 2000 and 2020. This time interval is chosen as a lot of data on more recent years is not available yet. The data that is used to conduct the empirical analysis is acquired from both the World Bank: Data and the Global Entrepreneurship Monitor (GEM). The data from GEM can be divided in two categories: Adult Population Survey (APS) and Entrepreneurial Framework Conditions (EFCs). The APS data is retrieved through questionnaires, which are administered to at least 2000 adults in each GEM country, with the aim to collect information on the entrepreneurial activity in those countries. The EFCs data consists of nine conditions that they have defined that affect new business creation in each country. The data of the world bank holds mostly data that is focused on general country characteristics such as GDP and population, while the GEM data is focused on the entrepreneurial aspects that will be necessary to conduct this research. While the data that is retrieved from the World Bank: Data is balanced and available for almost every

year for each country, the data from GEM is highly unbalanced, which should be considered when conducting the analysis. The different variables that are present in the dataset that is used in the empirical analysis is presented in Appendix 1A.

3.1 Inter- and Extrapolation

The data that is retrieved from the GEM consortium database has some missing years which induces an unbalanced dataset. To overcome this issue inter- and extrapolation is used. This is not an ideal situation; however, it is the best solution to overcome this issue. The problem is that the unbalanced character of the dataset seems to be non-random as mostly the northern European countries seem to include observations for all years, while eastern- and southern European countries have gaps in between the observations. Interpolation is used to fill in the gaps of the missing years by generating measures following a linear trend, while extrapolation can be used to fill in data that is on the end of the observations. However, using this method gave some outcomes that are not likely to occur and that fell outside of the initial range. Therefore, the observations that are generated using extrapolation that fell outside of the initial range are removed from the dataset. This method is used for four distinct variables from the GEM consortium database; Total Early-stage Entrepreneurship, Established Business Ownership, Perceived Opportunities and Motivational Index. Due to this method the database is significantly enlarged. Especially, the motivational index, which had an initial number of observations of 233, and after the inter- and extrapolation has 494 observations.

3.2 Independent variables

Total Entrepreneurship

The main independent variables that are used in this research are the three variables about entrepreneurship. These variables are acquired from the APS section of the GEM database; total early-stage entrepreneurial activity rate (TEA), established business ownership rate and the sum of these two variables, to measure the total rate of entrepreneurship. TEA is defined by GEM as the share of the population that is between 18 and 64 years old who are either involved in nascent entrepreneurship or be an owner-manager of a new business. The average that is measured over the dataset is a share of 7.63% that engages in this stage of entrepreneurship. The second variable, established business ownership rate, varies between 0.5% and 18.34%, with a mean of 6.35%, this measure is also applicable to the population that is aged between 18 and 64 years old. Finally, the sum of those two variables is calculated and shows the total percentage of entrepreneurs, for this variable a mean of 13.98% is found, which seems logical as it is approximately the sum of the two distinct variables. The maximum

of the sum of TEA and established entrepreneurs is 36.54%, which seems plausible as the maxima for the two distinct variables are together approximately 38%. The total number of observations is 494 and are measured per country for the interval 2000-2020. These descriptive statistics are found in Table 2 beneath. Finally, for the robustness check the log of the number of new businesses registered per year will be used as a robustness check to test whether TEA plays a role in the relationship with innovation. This proxy for entrepreneurship is chosen as it differs not too much from TEA, however the variable TEA is the share of individuals that started at least one new business, but there is a possibility that an individual registers multiple businesses, which is not accounted for using the variable TEA, although this share probably is small it can function as a good alternative. The reason it is not chosen as main variable, is the fact that the GEM database counts an individual as an established business owner after 42 months and therefore using the number of new businesses registered per year will create a gap of 30 months. Also, the number of businesses registered is only available for the time interval 2006-2020.

Table 2: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Total Entrepreneurship	494	13.977	4.811	3.64	36.54
TEA	494	7.626	3.133	1.78	20.4
Established entrepreneurs	494	6.35	2.766	.5	18.34
Business registered per year	369	.003	.002	0	.012
Log(Business registered)	369	-6.072	.835	-10.047	-4.405

3.3 Innovation Proxies

The second set of variables that will be discussed are the proxies for innovation. The problem with investigating innovation is that there is not a tangible measure that captures it. To overcome this problem former literature commonly uses proxies for innovation. For this research, the main proxy that will be used is the number of patent applications by residents, which is retrieved from the world bank data. This variable is selected as the main proxy for innovation because a patent on a particular innovation can only be awarded once, so people are willing to applicate for a patent as quick as possible, and therefore it is expected to rise rapidly as a response on an improvement of a nation's innovative capacity (Anokhin & Schulze, 2009). Resident patent applications are defined as the applications that are requested by residents of a certain country to the intellectual property office of that country, as is stated by the World Intellectual Property Organization (WIPO). The second proxy that is used is the expenditure on R&D, which is retrieved from the World Bank Data. This variable is chosen as former literature describes R&D expenditure as one of the factors that has a positive effect on innovation (Sharma, Davcik and Pillai, 2016), so if entrepreneurship has a

positive relationship with the expenditure on R&D this might imply a positive relationship with innovation.

Patent Applications

The number of patent applications by residents is retrieved from the World Data Bank and concern data of European countries for the time interval 2000-2020. The data shows that the total number of patent applications per 100.000 capita has a mean of 13.205 patent applications by residents per 100.00 capita's and that they fall in between the interval of 0.166 to 59.462. Which shows signs of a skewed distribution across the database and therefore the logarithm of the number of patent applications by residents is calculated and will be used for this research, as this will make the database closer to a normal distribution.

R&D expenditure

The second proxy for innovation that is used is the expenditure on R&D per capita. However, the expenditure on R&D is not normally distributed over European countries and therefore the log of R&D expenditure per capita is used. This variable is used for the robustness check on the analysis that is done with the patent application by residents. This variable is chosen as proxy for innovation as former literature has shown evidence that the expenditure on R&D is related with the level of innovation in a country (Sharma et al., 2016). There it might be useful to conduct a robustness check using this variable, as this can help to confirm the potential relationship that is found in the first hypothesis. The mean expenditure on R&D in the dataset is 552 dollars per capita, and a minimum of 0.673 dollars and a maximum of 2685 dollars per capita. A summary of the descriptive statistics of the proxies for innovation can be found in Table 3.

Table 3: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Resident patent applications per capita (100.000)	442	13.205	11.689	.166	59.462
Log(PatRes)	442	2.194	.951	-1.798	4.085
R&D expenditure per capita	460	552.127	578.394	.673	2685.358
Log (R&D expenditure pc)	460	5.477	1.61	-.396	7.896

3.4 Types of Entrepreneurs and Governmental Support

Next to the main independent variable, several other variables that might moderate the effect of entrepreneurship on innovation will be tested. First, the before mentioned moderator effect will be tested to check whether there are differences between the effect of TEA and established business owners on the number of patent applications by residents. This research

will test whether there is a difference between 'opportunity' and 'necessity' entrepreneurs and the relationship with governmental support for entrepreneurs.

Motivational Index

The variable that is used from STATA to analyze if there is a moderating effect of the motivation of those that are active in TEA is the motivational index and is also gathered from the GEM database. This variable is calculated by dividing the share of opportunity driven entrepreneurs by the share of necessity driven entrepreneurs. This is an unusual way of estimating the moderating effect of the motivation of entrepreneurship in comparison to the paper of Mrożewski and Kratzer (2017). In this research the motivational index is used as there are also other motives to involve in entrepreneurship, but those were not openly available at the GEM database and neglecting the other motives would potentially give biased results. The descriptive statistics that are shown in table 4 show that the mean for the motivational index is 3.759, and the observations fall in the interval 0.37 till 19.5. This variable is used because it could be argued that entrepreneurs who start a new enterprise opportunity motivated (e.g., they saw a possibility to improve their economic status), might be more innovative in comparison to entrepreneurs that created a new business out of necessity (e.g., because they could not find a job).

Table 4: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Motivational Index	494	3.759	3.636	.37	19.5

Governmental Support for Entrepreneurs

The second moderating variable that will be used in the analysis is the governmental support for entrepreneurs. This variable is retrieved from the GEM database and is one of the EFCs. Former literature has shown that support of the government through a better entrepreneurial environment or the education of their citizens can affect the level of innovation in a country. The variable that is retrieved is on a scale of 0 to 7 points, and a higher value implies a higher extent of governmental support. For the robustness check the variable ease of doing business is used, which is retrieved from the World Bank Data and stands for the business environment in each country for indicators such as starting up a business and the strength of property rights on a scale of 0-100. However, this variable is not chosen as the main variable because of the lack of observations as it is only measured over the 2015-2019 interval.

Table 5: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Governmental Support for Entrepreneurs	484	4.275	.799	2.5	6.6
Ease of Doing Business	150	76.092	4.437	64.128	84.595

3.5 Control Variables

Besides the empirical research will also include multiple control variables, to overcome the problem of biases due to variables that both influence the dependent and independent variable that are omitted. Although, it may not be feasible to control for all relevant factors (e.g., some may not be observed), including multiple variables can help to overcome the issue of biased results at least to some degree. The control variables that will be included are the following: perceived opportunities, GDP growth, population, foreign direct investments (FDI) inflow per capita and the R&D expenditure per capita. First, the perceived opportunities are included, this variable is defined as the share of the population that is aged between 18 and 64 years and are not active in any stage of entrepreneurial activity, which notices an opportunity to start a firm in the area they live in. Besides, it could possibly influence innovation as apparently there are gaps in the market that might be filled with innovative ideas. They also mention that the perception of opportunities is an appropriate measure to determine whether individuals perceive opportunities, which is important for the discovery and creation view of entrepreneurship, they show that there is an actual moderating effect of the perceived opportunities on innovation, at least for some part of TEA. The descriptive statistics show that that the difference between the minimum (7.58%) and maximum (87.28%) is quite high. The second control variable that is included is GDP growth, a higher GDP growth might imply that the economic status of a country is growing, which will likely be the case for the less developed countries in Europe, while most of the northern European countries will experience a lower GDP growth, which is more stable. Third, the log of the population size is considered as country size might influence the number of patent applications, but not the independent variable, entrepreneurship, as this is measured in percentages. The distribution for population sized is not normally distributed in this dataset and therefore the log of the population size will be used for the empirical analysis. The fourth variable, the log of FDI inflow per capita, that is included shows the inflow of foreign direct investments, thus the investments that are made by foreign third parties. In former literature it is discussed that FDI could potentially play a crucial role in innovation, as a higher level of FDI inflow might affect the economic dynamics and performance (Anokhin & Wincent, 2012). This is also highlighted by Mrożewski & Kratzer (2017), as they argue that a higher level of FDI inflow might be aligned with more openness which in turn might also result in more knowledge spillovers from foreign investors which can help foster innovation that is related to new technologies, marketing, or management techniques (Javorcik & Spatareanu, 2008). The final and fifth variable that is included in the empirical analysis is the log of the expenditure on education. This variable is included as countries that are spending a larger part on education, might aim to create a higher educational attained population, which could return in higher levels of innovation. Gregoire

and Shepherd (2012) contributed to former literature, by stating that entrepreneurs with higher levels of education have a higher probability of succeeding. The reason that the expenditure on education is chosen instead of the share of the population who followed tertiary education is because the values that can be acquired through the World Bank Data give values above 100%, because it is calculated as the share of the population who are eligible to follow tertiary education. Therefore, this variable could potentially be driven by other factors that result in those percentage above 100%.

Table 6: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Perceived opportunities	494	36.515	15.122	7.58	87.28
GDP Growth	470	2.114	3.844	-14.839	25.176
Log(Population)	494	16.189	1.253	13.182	18.789
Log(FDI per capita)	432	6.574	1.563	-.534	11.994
Log(Education Expenditure per capita)	433	6.987	1.052	3.689	8.949

3.6 Correlation Matrix

To get a better understanding of the data prior to the empirical analysis a correlation matrix is constructed in table 7. Using this matrix, the possibility of biases due to multicollinearity could be checked and some brief insights in the correlation between the dependent and independent variables are discussed. First, it seems that the correlations between the independent variables, which are used in the same models, is not exceptionally high and therefore unlikely to create biases due to multicollinearity. Second, it can be seen in the correlation matrix is that the different control variables are all correlated with the dependent variable, which shows that they should be included in the model to reduce the chance of omitted variable bias. Finally, the correlation matrix shows negative estimates for the correlations between the log of the patent applications and the share of entrepreneurs. On the other hand, the correlation between the moderator variables and the proxy for innovation is positive.

Table 7: Matrix of correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Log(PatRes)	1.000													
(2) Total Entrepreneurship	-0.424	1.000												
(3) TEA	-0.414	0.717	1.000											
(4) Established entrepreneurs	-0.207	0.740	0.062	1.000										
(5) Motivational Index	0.340	0.061	-0.037	0.123	1.000									
(6) Governmental Support for Entrepreneurs	0.231	0.021	0.054	-0.022	0.378	1.000								
(7) Log (R&D expenditure pc)	0.597	-0.252	-0.326	-0.046	0.563	0.150	1.000							
(8) Log(Business registered)	-0.254	0.236	0.389	-0.038	0.120	0.132	0.022	1.000						
(9) Ease of Doing Business	0.282	-0.052	0.054	-0.126	0.226	0.429	0.203	0.194	1.000					
(10) Perceived opportunities	0.305	0.015	0.081	-0.057	0.556	0.288	0.435	0.120	0.413	1.000				
(11) Log(Population)	0.588	-0.430	-0.476	-0.157	-0.058	0.045	0.201	-0.414	0.131	-0.055	1.000			
(12) Log(FDI per capita)	-0.068	0.129	0.163	0.027	0.352	0.247	0.455	0.412	0.031	0.353	-0.328	1.000		
(13) GDP Growth	-0.510	0.192	0.286	-0.001	-0.031	0.059	-0.200	0.295	0.164	0.211	-0.434	0.254	1.000	
(14) Log(Education Expenditure per capita)	0.514	-0.178	-0.175	-0.087	0.603	0.234	0.960	0.192	0.265	0.517	0.083	0.547	-0.114	1.000

4. Methodology

To answer the research question of this paper an empirical analysis will be conducted with the data that is discussed in the previous section. The aim of this paper is to investigate whether entrepreneurial activity affects the level of innovation. To do so panel data will be used to determine whether this effect exists. As dependent variable the log of the number of patent applications by residents will be used. The main independent variable is the total percentage of entrepreneurship, which is calculated by taking the sum of the percentages of early-stage entrepreneurship and established business ownership, which are acquired through the GEM database. To visualize the effect that is expected the means of the patent applications per 100.000 capita's (figure 1) and the mean of the total percentage of entrepreneurial activity (figure 2) over the interval 2000-2020 are shown beneath. What can be derived from these figures is that the mean of the number of patent applications over the interval is decreasing, while the mean of entrepreneurial activity is increasing. However, these figures are measured as the average of all European countries combined and therefore it does not have to imply that the results correspond with these figures. Nevertheless, this visualization can still give some insights in the relationship that will be discussed.

Figure 1:

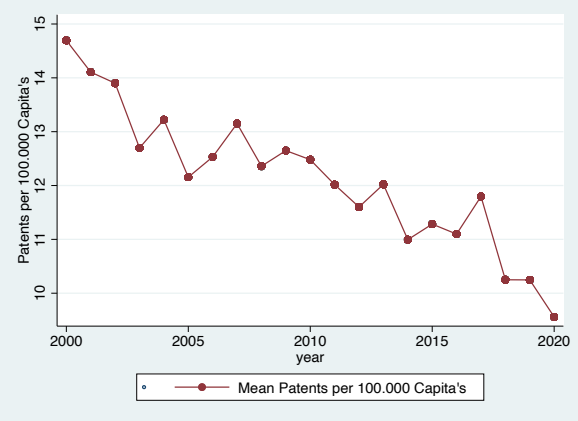
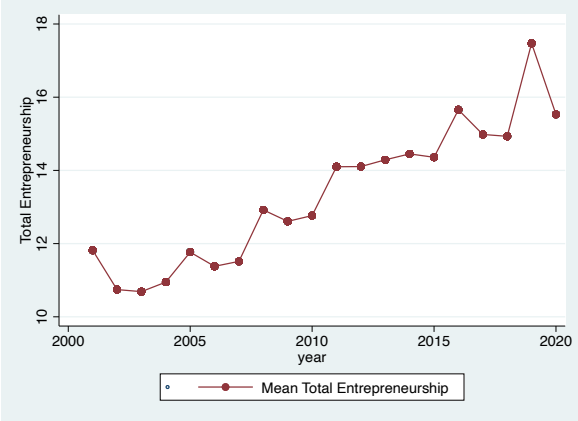


Figure 2:



4.1 Hypothesis 1

For the first hypothesis the relationship between entrepreneurship and the number of patent applications will be tested. As the data that is used has a panel format, the choice is between three estimation models: pooled OLS, fixed effects, and random effects. However, the pooled OLS method can be eliminated as the likelihood of the absence of an idiosyncratic show does not seem to be very high in this case. Therefore, the choice must be made between a fixed- or random effects model, this also holds for the other upcoming hypothesis. To test which

model is most appropriate for this research a Hausman test will be performed. To do so both models are performed, first the fixed effects and thereafter the random effects. The outcome of the Hausman test is shown in Table 8. The p-value that is found for this test is 0.279, and therefore the null-hypothesis that there are no systematic differences in coefficients for the two models is not rejected. This indicates that the appropriate model for the first hypothesis is a random effects model. A potential reason for this outcome is the relative time-invariant nature of some of the variables that are included in this research. This characteristic is encountered in earlier studies as well for the variables that are acquired through the GEM consortium database regarding entrepreneurship (Koellinger, 2008; Mrożewski & Kratzer, 2017).

Table 8: Hausman (1978) specification test

	Coef.
Chi-square test value	29.76
P-value	.278

Note: Hausman test for regression (1)

The regression that will be tested to determine whether entrepreneurship impacts the level of innovation is formulated below (1). The independent variable that is used is the number of resident patent applications by residents and is measured per country and year. The independent variable is the total percentage of entrepreneurs, which is also measured per country and year. Furthermore, a vector variable including the control variables is included in the model which is denoted as $X_{i,t}$, the variables that are included in this vector variable are also measured per country and year. The control variables that are used in the first regression are the GDP growth, the log of population size, the log of FDI inflow per capita, perceived opportunities and the log of the expenditure on education per capita. Besides, time fixed effects are included in the model, which is denoted as TFE_t and enables to control for time related heterogeneity.

$$(1)\ ResPat_{t,i} = \beta_0 + \beta_2 * TotEntre_{i,t} + \alpha_3 * X_{i,t} + \vartheta * TFE_t + \varepsilon_{i,t}$$

For the second part of the first hypothesis another Hausman test is performed and gives a p-value of 0.840 (Table 9), which means that the appropriate test for this regression again is a random effects model. The aim of this hypothesis is to test if early-stage entrepreneurship and established business ownership might affect the level of innovation in different ways and are therefore included separately in the model, instead of the previous regression that included the sum of these two variables, and it can therefore be used to determine a moderating effect. So, the test is used to determine when the tested association will be present.

Table 9: Hausman (1978) specification test

	Coef.
Chi-square test value	19.77
P-value	.840

Note: Hausman test for regression (2)

The equation that is used for this hypothesis is almost identical to the first equation, however as mentioned before $TEA_{i,t}$ and $Established_{i,t}$, are included separately in the model. Nevertheless, the control vector and time fixed effects are also included in this model.

$$(2)\ ResPat_{t,i} = \beta_0 + \beta_2 * TEA_{i,t} + \beta_3 * Established_{i,t} + \alpha_4 * X_{i,t} + \vartheta * TFE_t + \varepsilon_{i,t}$$

4.2 Hypothesis 2

The second hypothesis tests whether there are appearing differences on the relationship of early-stage entrepreneurship and innovation due to differences in the distribution of opportunity and necessity entrepreneurs. To do so the variable motivational index is used, which is acquired from the GEM consortium database. This variable is defined as the number of opportunity entrepreneurs divided by the number of necessity entrepreneurs. For this test another Hausman test is performed which shows a p-value of 0.980 (Table 10), which is higher than 0.050 and therefore a random effects model is the appropriate test for this hypothesis.

Table 10: Hausman (1978) specification test

	Coef.
Chi-square test value	14.84
P-value	.980

Note: Hausman test for regression (3)

The equation that is used to test the third hypothesis is formulated below (3). For this hypothesis the motivational index is included in the model as independent variable. The independent variable, motivational index, is measured per country and year. Besides, the vector variable including the controls and time fixed effects are also included in the model.

$$(3)\ ResPat_{t,i} = \beta_0 + \beta_1 * Motive_{i,t} + \beta_2 * TEA_{i,t} + \beta_3 * Established_{i,t} + \alpha_4 * X_{i,t} + \vartheta * TFE_t + \varepsilon_{i,t}$$

4.3 Hypothesis 3

The third hypothesis uses $GovSupport_{i,t}$ as independent variable to determine whether policies to support the entrepreneurial activity in a country are successful, this is tested as some papers found that governmental interference is not always the best solution to create a flourishing entrepreneurial environment (Shane, 2009). For the third hypothesis another Hausman test is performed. Same as the first two hypotheses the outcomes of the Hausman test also gives a p-value above 0.050, namely 0.955. This indicates that there are systematic differences in the coefficient between the fixed- and random effects model and therefore a random effects model is preferred.

Table 11: Hausman (1978) specification test

	Coef.
Chi-square test value	16.65
P-value	.955

Note: Hausman test for regression (4)

The equation that is constructed for the third hypothesis is stated below (4). The independent variable that is included in this hypothesis is $GovSupport_{i,t}$, and is defined by the GEM consortium database as the extent to which public policies are supportive to entrepreneurship. The independent variable that is included in this model is also measured per country and year. Besides, the inclusion of the governmental support, two distinct models are tested which includes an interaction term between early-stage entrepreneurship and governmental support and an interaction term between established business ownership and governmental support.

$$(4)\ ResPat_{t,i} = \beta_0 + \beta_1 * GovSupport_{i,t} + \beta_2 * TEA_{i,t} + \beta_3 * Established_{i,t} + \beta_4 * GovSupport_{i,t} * TEA_{i,t} + \alpha_5 * X_{i,t} + \vartheta * TFE_t + \varepsilon_{i,t}$$

4.4 Robustness Check

Finally, multiple robustness checks are performed to control the estimates that are found by the initial analysis. First, the dependent variable will be changed to the expenditure on R&D, to control the first hypothesis. Sharma et al. (2016), found evidence that R&D expenditure has a positive effect on product innovation. For this reason, R&D expenditure is chosen to check the robustness of the results that are found in this research. Whenever, the results hold that entrepreneurship has a positive effect on the expenditure on R&D, this might imply the robustness of the first hypothesis. For this robustness check a fixed effects model is tested (Appendix 2), with the equation that is formulated below (5).

$$(5)\ R\&DEXP_{t,i} = \beta_0 + \beta_2 * TotEntre_{i,t} + \alpha_3 * X_{i,t} + \vartheta * TFE_t + \varepsilon_{i,t}$$

As a second robustness check the number of new businesses registered is used as a proxy for entrepreneurship and replaces the variables TEA and Established business ownership. This variable is similar to the share of individuals that are involved in TEA. However, using the number of new businesses registered will also account for those individuals that started more than one business. Besides, the set of control variables is the same as in the first hypothesis. The aim of this robustness check is the same as the first one as it will be performed to determine whether it is likely if the relationship that is found in the first hypothesis holds when the variable number of new businesses registered is used. The tested equation is standing below, and a random effects model is performed again (Appendix 3).

$$(6)\ ResPat_{t,i} = \beta_0 + \beta_1 * NewBusiness_{i,t} + \alpha_2 * X_{i,t} + \vartheta * TFE_t + \varepsilon_{i,t}$$

The final robustness check that is performed, will test the reliability of the third hypothesis. The aim of the third hypothesis is to test whether there is a moderating effect of governmental support for entrepreneurs on the relationship with innovation. However, for this robustness check the variable governmental support is replaced with the ease of doing business rank. This is a rank on a scale of 0 to 100. Where 0 means the lowest performance and 100 the best performance. This variable is retrieved from the world bank data. The reason that this variable is not used as the main variable to determine whether such a moderating effect exist is because the variable is only available between 2015 and 2019, which is only a small sample of the interval that is investigated in this research. The model that is determined to be most appropriate is a fixed effects model (Appendix 4).

$$(7)\ ResPat_{t,i} = \beta_0 + \beta_1 * EaseofDoingBusiness_{i,t} + \beta_2 * TEA_{i,t} + \beta_3 * Established_{i,t} + \beta_4 * EaseofDoingBusiness_{i,t} * TEA_{i,t} + \alpha_5 * X_{i,t} + \vartheta * TFE_t + \varepsilon_{i,t}$$

5. Results

5.1 Hypothesis 1

For the first hypothesis the relationship between entrepreneurial activity and innovation is tested, with the log of the number of patent applications per 100.000 capita's as proxy for innovation. To do so, model 1 is tested with the total share of the population between 18 and 64 years old that is active in any stage of entrepreneurial activity as an independent variable. In model 1 it is found that there is no significant relationship between the log of the number of

patent applications by residents and the share of individuals that are active in any stage of entrepreneurship. However, to elaborate on this effect the total share of entrepreneurship is split in both early-stage entrepreneurship and established business ownership, these tests are shown in model 2. When including TEA and established business ownership in the model it is found that established business ownership seems to have a positive effect that is significant on a 99% confidence interval. While the estimate for TEA shows that there is a negative relationship between the log of the number of patent applications and the share of individuals that are either a nascent entrepreneur or owner-manager of a new business, on a 90% confidence interval. These estimates are in line with the findings of Wong et al. (2005), because they suggest that the overall TEA, has no significant effect on the economic growth. However, when they separated TEA in subcategories, they found that a small part of the TEA actually engages in true technological innovation, which are those that are defined as high potential TEA, which is tested for the second hypothesis. Concluding, the results of the models show that the total share of entrepreneurship has no significant relationship with innovation. However, when entrepreneurship is split in TEA and established business ownership it is found that there is a positive effect by those who are established business owners, while TEA seems to negatively impact innovation, which might induce the insignificant result if they are tested as the entire share of entrepreneurship.

Table 12: Model 1-2, Hypothesis 1

	(1) Log(ResPat)	(2) Log(ResPat)
Total Entrepreneurship	.004	
	(.004)	
TEA		-.011*
		(.006)
Established entrepreneurs		.022***
		(.007)
Perceived opportunities	.007***	.006***
	(.001)	(.001)
Log(Population)	.268***	.25***
	(.088)	(.094)
Log(FDI per capita)	.012	.012
	(.013)	(.013)
Log(Education Expenditure per capita)	.114**	.135**
	(.056)	(.057)
GDP Growth	-.014**	-.013**
	(.006)	(.006)
Constant	-2.956**	-2.704*
	(1.477)	(1.584)
Observations	361	361
Overall R^2	.435	.445
Year Dummies	YES	YES

Note: Regression model with the log of patent applications by residents as dependent variable, entrepreneurship as independent variable, and control variables, performing a random effects model.
Standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

5.2 Hypothesis 2

For the second hypothesis a random effects model is used to determine whether a higher proportion of 'opportunity' entrepreneurs as state by the GEM consortium database positively impacts the number of patent applications. This is tested in model 3. The Motivational index as defined by the GEM consortium database is the percentage of those that are active in TEA and are opportunity motivated, divided by those that are necessity motivated. These two types of TEA are self-assessed by the entrepreneurs that are included in the survey. In model 3, it is found that there is no significant relationship with the motivational index and the log of the number of patent applications. This however is not in line with former literature whereas Mrożewski and Kratzer (2016), have found a positive effect on the level of innovation of opportunity entrepreneurs and the contrary for necessity entrepreneurs. However, the variable specifications that they used are different than the ones in this research. To overcome the issue of the gaps in the GEM data they choose to take the mean of every variable instead of using inter- and extrapolation. They stated that some of the variables could be seen as time invariant, and although this could be the case for some variables it is not likely to hold for all

variables. Besides, it also limits their number of observations drastically, which might induce unreliable estimates. For example, in their robustness check they also use the number of patent applications per capita as an independent variable and get an outcome of -0.177 patent applications for the interaction term between opportunity entrepreneurship and perceived opportunities, which means that whenever the perceived opportunities stay the same and there will be a 1% rise in the share of opportunity driven entrepreneurs the number of patent applications decrease by 0.002 per capita. Consulting the World Data Bank, it could be found that the worldwide number of patent applications is around 2.3 million and a population around 7.8 billion. Taking these figures into account it could be derived that these numbers don't add up and there might be some defects in either their model, or variable specification. All in all, the results of this research suggest that there is no closing evidence for a relationship between the distribution of opportunity and necessity entrepreneurs and the level of innovation.

Table 13: Model 3, Hypothesis 2

	(3) Log(ResPat)
Motivational Index	.002
	(.005)
TEA	-.01*
	(.006)
Established Entrepreneurs	.022***
	(.007)
Perceived Opportunities	.006***
	(.001)
GDP Growth	-.013**
	(.006)
Log(Population)	.232**
	(.105)
Log(FDI per capita)	.012
	(.013)
Log(Education Expenditure per capita)	.115**
	(.058)
Constant	-2.323
	(1.75)
Observations	361
Overall R^2	.438
Year Dummies	YES

Note: Regression model with the log of patent applications by residents as dependent variable, motivational index and entrepreneurship as independent variable, and control variables, performing a random effects model.
Standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

5.3 Hypothesis 3

Finally, the moderating effect of governmental support on the relationship between entrepreneurship and innovation is tested. Former literature has found as already stated in the literature review that a higher level of governmental interference negatively impacts the entrepreneurial activity (Aidis et al., 2012). However, this does not have to imply anything about the type of entrepreneurs that still succeed to enter the market. It even could be possible that such support systems require some standards that must be matched before qualifying for the support and induces a fair system, which helps to select the right entrepreneurs. On the other hand, as already mentioned for hypothesis 1, Wennekers et al. (2005) suggests that for developed countries public policies are regarded as a promising approach to foster the economic growth, which in turn might lead to higher levels of innovation. In model 4-6, the effect of governmental support is tested, to determine whether such policies will impact the relationship of entrepreneurship and innovation. In model 4, the results suggest that there is no explicit impact of governmental support on the level of innovation, as the estimate for governmental support is insignificant. However, the results show some promising estimates as the estimate for the interaction term with TEA is positive and statistically significant, on a 90% confidence interval, while the estimate for the established business owners is negative and significant on a 90% confidence interval. Based, on these findings there seems to be a moderating effect of governmental support for entrepreneurship for both TEA and established business owners. However, the estimates show opposite results, where TEA can benefit from such policies, it negatively impacts the innovativeness of established business owners.

Table 14: Model 4-6, Hypothesis 3

	(4) Log(ResPat)	(5) Log(ResPat)	(6) Log(ResPat)
TEA	-.011*	-.055**	-.011**
	(.006)	(.025)	(.006)
Established Entrepreneurs	.027***	.027***	.074***
	(.007)	(.007)	(.025)
Governmental Support for Entrepreneurs	-.029	-.11**	.037
	(.025)	(.05)	(.042)
Perceived Opportunities	.004***	.004***	.004***
	(.001)	(.001)	(.001)
Log(Population)	.219**	.166	.203*
	(.103)	(.116)	(.115)
GDP Growth	-.012**	-.013**	-.01*
	(.006)	(.005)	(.006)
Log(FDI per capita)	.015	.017	.015
	(.013)	(.012)	(.012)
Log(Education Expenditure per capita)	.145**	.113*	.133**
	(.06)	(.061)	(.06)
Governmental Support for Entrepreneurs*TEA		.011*	
		(.006)	
Governmental Support for Entrepreneurs* Established Entrepreneurs			-.011*
			(.006)
Constant	-2.081	-.724	-2.056
	(1.719)	(1.962)	(1.927)
Observations	355	355	355
Overall R^2	.417	.406	.406
Year Dummies	YES	YES	YES

Note: Regression model with the log of patent applications by residents as dependent variable, governmental support on entrepreneurship and entrepreneurship as independent variable, and control variables, performing a random effects model. Standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

5.4 Robustness Check

Finally, a check is executed to test whether the results of the impact of entrepreneurship on innovation are robust. To test this the dependent variable that is used to conduct the initial research is replaced with the expenditure on R&D per capita. This independent variable is chosen as former literature found a relationship between the expenditure on research and development and the level of innovation. The literature that investigated this topic found that innovation is significantly and positively affected by R&D (Sharma et al., 2016). As is found in the first hypothesis there is a positive and significant effect between entrepreneurship and

innovation, however this effect seems to be mainly driven through established entrepreneurship. Therefore, the robustness check is executed using TEA and established business ownership separately. The results of the robustness check are shown in table 15, model 7. The findings of the robustness check are corresponding with the results of the first hypothesis. Model 7 shows that there is a negative and significant effect of TEA on the log of R&D expenditure per capita, on a 99% confidence interval. However, for the share of established business owners there is found a positive and significant estimate on a 95% confidence interval.

Table 15: model 7, robustness check hypothesis 1

	(7) log_rd_exp
TEA	-.016***
	(.004)
Established entrepreneurs	.011**
	(.005)
Perceived Opportunities	-.001
	(.001)
Log(Population)	-2.203***
	(.247)
GDP Growth	.002
	(.004)
Log(FDI per capita)	.015
	(.009)
Log(Education Expenditure per capita)	.662***
	(.046)
Constant	36.373***
	(4.054)
Observations	374
R-squared	.857
Year Dummies	YES

Note: Regression model with the log of R&D expenditure as independent variable, entrepreneurship as independent variable, and control variables, performing a random effects model.
Standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

The second robustness check that is performed is to test whether the relationship that is found in the first hypothesis still holds when the variable TEA is replaced with the log of the number of businesses registered per year. What the robustness check shows is that also for the log of new businesses registered the estimate is negative and significant on a 90% confidence interval. Based on the findings of the first hypothesis this robustness check confirms the result that TEA has a negative relationship with innovation. However, this robustness check is only

able to substantiate the findings of TEA and not the findings of the established business owners.

Table 16: model 8, robustness check hypothesis 1

	(8) Log(ResPat)
Log(Business Registered)	-.077*
	(.039)
Perceived Opportunities	.009***
	(.001)
Log(Population)	.288***
	(.08)
GDP Growth	-.015**
	(.006)
Log(FDI per capita)	-.01
	(.014)
Log(Education Expenditure per capita)	.258***
	(.074)
Constant	-4.723***
	(1.409)
Observations	284
Overall R^2	.593
Year Dummies	YES

Note: Regression model with the log of patent applications by residents as independent variable, the log number of businesses registerd as independent variable, and control variables, performing a random effects model. Standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

The final robustness check that is performed will test the reliability of the findings of hypothesis 3. What is found after performing the third robustness check is that the results are not quite similar to the results that were found in the initial tests that are performed. What the results of the robustness check show is that only model 10, with the interaction term between TEA and ease of doing business are found to be robust, while the test without the interaction effect and the test with the interaction effect between ease of doing business and the share of individuals that are established business owners are insignificant. A potential reason for these insignificant results is the number of measures for the variable ease of doing business that were available, as these were only measured between 2015 and 2019.

Table 17: model 9-11, robustness check hypothesis 3

	(9) Log(ResPat)	(10) Log(ResPat)	(11) Log(ResPat)
TEA	-.013	-.541**	-.012
	(.012)	(.261)	(.013)
Established entrepreneurs	-.006	-.006	-.155
	(.013)	(.012)	(.216)
Ease of Doing Business	-.018	-.09**	-.028
	(.023)	(.042)	(.028)
Perceived Opportunities	.001	.002	.002
	(.003)	(.003)	(.003)
Log(Population)	.855	.204	.932
	(1.766)	(1.757)	(1.777)
GDP Growth	-.002	.002	-.002
	(.017)	(.017)	(.017)
Log(FDI per capita)	.018	.027	.02
	(.023)	(.023)	(.024)
Log(Education Expenditure per capita)	.124	.259	.115
	(.28)	(.282)	(.282)
Ease of Doing Business*TEA		.007**	
		(.003)	
Ease of Doing Business* Established entrepreneurs			.002
			(.003)
Constant	-11.302	3.739	-11.704
	(29.635)	(29.915)	(29.756)
Observations	110	110	110
R-squared	.314	.354	.319
Year Dummies	YES	YES	YES

Note: Regression model with the log of patent applications by residents as dependent variable, ease of doing business and entrepreneurship as independent variable, and control variables, performing a random effects model.
Standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

6. Conclusion and Discussion

Finally, the results that were found will be summarized in this section and potential mechanisms that play a role at explaining the results will be discussed. Thereafter some limitations that came across this research will be discussed and possibilities for future research are given.

First, after conducting the empirical analysis for the first hypothesis gives some insight in the relationship between entrepreneurship and the number of patent applications by residents. However, the results suggest that there is no overall relationship between entrepreneurship and innovation. But when early-stage entrepreneurship and established business owners are split, it is found that the established business owners have a positive and significant effect on the number of patent applications by residents. This result is similar to the findings of Hagedoorn (1996), who shows that larger firms are the main drivers of innovation, as it is reasonable to suggest that established businesses are often larger in comparison to TEA. These results are also confirmed by two robustness checks that are performed using the R&D expenditure as a proxy for innovation and the log of businesses registered as a replacement for TEA. One potential mechanism that could drive the results that established business owners contribute to innovation are knowledge spillovers, as such spillovers are a strategy that is executed by existing organizations, which results in growth in industries, regions, and economies (Agarwal, Audretsch and Sarkar, 2007), which in turn might enhance innovation. Competitiveness also could be a mechanism that enhances innovation as Tang (2006) has found that more competitive markets show a positive relationship with the expenditure on R&D.

Besides, it is also tested whether there is a moderating effect of opportunity and necessity entrepreneurs on the relationship with innovation. However, contrary to the findings of Mrożewski and Kratzer (2016), there is no significant estimate found for a moderating effect of the motivational index, which is the share of opportunity entrepreneurs divided by the share of necessity entrepreneurs. However, the method and variables that are used by them are different than in this paper. Besides, they have also tested a different set of countries as they focused on all countries that are available at GEM, and this research focused on European countries.

Finally, the effect of governmental support is tested and shows a significant interaction with both TEA and established entrepreneurship. The interaction term that is found between TEA and governmental support for entrepreneurs is positive, while the interaction term with

established business owners is negative. Although the variable for governmental support only shows the extent to which a government supports entrepreneurs it does not define in which manner. However, former literature has given some examples of governmental interference that could have a moderating effect on innovation such as education (Sobel & King, 2008) and a fair legal justice system (Aidis et al.,2012). Concluding these findings an answer on the research question could be formulated:

Does a larger share of individuals active in any stage of entrepreneurship have a positive effect on innovation in Europe for the time interval, 2000-2020?

Although a higher share of individuals active in any stage of entrepreneurship does not seem to have a significant and positive relationship with innovation, it could be concluded that a higher share of established business owners has a positive relationship with innovation. On the other hand, it is also found that TEA has a negative relationship with the number of patent applications by residents. Besides, the findings of this research show that governmental support has a moderating effect on the relationship between entrepreneurship and innovation, while there is no significant estimate found for the motivational index.

The findings of these paper are likely to be subject to several limitations. First, the characteristics of the dataset are a disadvantage, as the data on entrepreneurial variables that are acquired through the GEM consortium database is unbalanced and had to be fixed using inter- and extrapolation. Using this method, could give some false estimates in comparison to the true events. This problem occurs more often in research using data from the GEM consortium database, for example Mrożewski and Kratzer (2016), also had to deal with this issue and therefore took the averages for each country and conducting an OLS regression, which potentially could be a solution. But the downside of that method is the impact on the size of the dataset, because it reduces the number of observations significantly. Another limitation is the fact that innovation is hard to measure and is commonly seen as an issue for research that want to test variables that impact innovation (Koellinger, 2008). He stated that for example R&D expenditure is not a proper measure of innovation, while countries might differ in their effectiveness in doing R&D. However, a measure that seems more appropriate is the Global Competitiveness Index (GCI), which is used by Mrożewski and Kratzer (2016), and potentially will be better able to capture innovation as this index includes multiple characteristics of innovation, such as R&D spending, patents, and university-industry collaboration.

Finally, also some suggestions for future research are discussed. As is already discussed as a limitation of this research is the innovation proxy. Therefore, it might be a potential idea to conduct the same research with the GCI as a proxy for innovation as this variable seems to be more all-encompassing. Besides, it is also very likely that this research is subject to omitted variable bias, thus for future research it might be reasonable to take this into account and look for other variables that might influence the level of innovation in a country. Another suggestion for future research is to include a broader set of control variables, as this research is most likely not able to include all the control variables that are needed to overcome the issue of omitted variable bias.

7. References

Aidis, R., Estrin, S., & Mickiewicz, T. M. (2012). Size matters: entrepreneurial entry and government. *Small Business Economics*, *39*(1), 119–139.

Agarwal, R., Audretsch, D., & Sarkar, M. B. (2007). The process of creative construction: knowledge spillovers, entrepreneurship, and economic growth. *Strategic Entrepreneurship Journal*, *1*(3-4), 263–286. https://doi.org/10.1002/sej.36

Anokhin, S., & Schulze, W. S. (2009). Entrepreneurship, innovation, and corruption. *Journal of Business Venturing*, *24*(5), 465–476. https://doi.org/10.1016/j.jbusvent.2008.06.001

Anokhin, S., & Wincent, J. (2012). Start-up rates and innovation: a cross-country examination. *Journal of International Business Studies*, *43*(1), 41–60.

Audretsch, D. B. (1995). Innovation, growth and survival. *International Journal of Industrial Organization*, *13*(4), 441–457. https://doi.org/10.1016/0167-7187(95)00499-8

Audretsch, D.B., Carree, M.A., & Thurik, A.R. (2001). Does Entrepreneurship Reduce Unemployment? *Working Paper*. Tinbergen Institute, Amsterdam and Rotterdam

Audretsch, D. B., & Keilbach, M. (2007). The theory of knowledge spillover entrepreneurship*. *Journal of Management Studies*, *44*(7), 1242–1254. https://doi.org/10.1111/j.1467-6486.2007.00722.x

Cefis, E., & Marsili, O. (2006). Survivor: the role of innovation in firms' survival. *Research Policy*, *35*(5), 626–641. https://doi.org/10.1016/j.respol.2006.02.006

Entrepreneurship, (n.d). Retrieved from https://eit.europa.eu/our-activities/entrepreneurship

Furman, J. L., Porter, M. E., & Stern, S. (2002). The determinants of national innovative capacity. *Research Policy*, *31*(6), 899–933. https://doi.org/10.1016/S0048-7333(01)00152-4

(n.d.). GEM Consortium. https://www.gemconsortium.org/

Gregoire, D. A., & Shepherd, D. A. (2012). Technology-market combinations and the identification of entrepreneurial opportunities: an investigation of the opportunity-individual nexus.(report). *Academy of Management Journal*, *55*(4), 753.

Hagedoorn, J. (1996). Innovation and entrepreneurship: schumpeter revisited. *Industrial and Corporate Change*, *5*(3), 883–896. https://doi.org/10.1093/icc/5.3.883

Hansen, J. A. (1992). Innovation, firm size, and firm age. *Small Business Economics*, *4*(1), 37–44.

Huergo, E., & Jaumandreu, J. (2004). How does probability of innovation change with firm age? *Small Business Economics*, *22*(3-4), 193–207.

Javorcik, B., & Spatareanu, M. (2008). To share or not to share: Does local participation matter for spillovers from foreign direct investment? Journal of Development Economics, 85(1–2), 194–217.

Kahn, K. B., & Candi, M. (2021). Investigating the relationship between innovation strategy and performance. *Journal of Business Research*, *132*, 56–66. https://doi.org/10.1016/j.jbusres.2021.04.009

Klippel, A. F., Petter, C. O., & Antunes, J. A. V. (2008). Management innovation, a way for mining companies to survive in a globalized world. *Utilities Policy*, *16*(4), 332–333. https://doi.org/10.1016/j.jup.2007.09.001

Koellinger, P. (2008). Why are some entrepreneurs more innovative than others? *Small Business Economics*, *31*(1), 21–37.

Mrożewski Matthias, & Kratzer, J. (2017). Entrepreneurship and country-level innovation: investigating the role of entrepreneurial opportunities. *The Journal of Technology Transfer*, *42*(5), 1125–1142. https://doi.org/10.1007/s10961-016-9479-2

Parrotta, P., & Pozzoli, D. (2012). The effect of learning by hiring on productivity. *The RAND Journal of Economics*, *43*(1), 167–185.

Parker, S. C. (2007). Law and the economics of entrepreneurship. *Comparative Labor Law and Policy Journal*, *28*(4), 695–716.

Reynolds, P., Camp, M., Bygrave, W., Autio, E., & Hay, M. (2002). Global entrepreneurship monitor 2001 executive report. Boston, MA: Global Entrepreneurship Research Association, Babson College.

Shane, S. (2009). Why encouraging more people to become entrepreneurs is bad public policy. *Small Business Economics*, *33*(2), 141–149.

Shane, S., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *The Academy of Management Review*, *25*(1), 217–226.

Sharma, P., Davcik, N. S., & Pillai, K. G. (2016). Product innovation as a mediator in the impact of r&d expenditure and brand equity on marketing performance. *Journal of Business Research*, *69*(12), 5662–5669. https://doi.org/10.1016/j.jbusres.2016.03.074

Sobel, R. S., & King, K. A. (2008). Does school choice increase the rate of youth entrepreneurship? *Economics of Education Review*, *27*(4), 429–438. https://doi.org/10.1016/j.econedurev.2007.01.005

StataCorp. 2019. *Stata Statistical Software: Release 16.* College Station, TX: StataCorp LLC.

Tang, J. (2006). Competition and innovation behaviour. *Research Policy*, *35*(1), 68.

Turut, O., & Ofek, E. (2012). Innovation strategy and entry deterrence. *Journal of Economics & Management Strategy*, *21*(3), 583–631. https://doi.org/10.1111/j.1530-9134.2012.00342.x

Wennekers, S., van Stel André, Thurik, R., & Reynolds, P. (2005). Nascent entrepreneurship and the level of economic development. *Small Business Economics : An International Journal*, *24*(3), 293–309. https://doi.org/10.1007/s11187-005-1994-8

Wong, P. K., Ho, Y. P., & Autio, E. (2005). Entrepreneurship, innovation and economic growth: evidence from gem data. *Small Business Economics : An International Journal*, *24*(3), 335–350. https://doi.org/10.1007/s11187-005-2000-1

World Bank, *World Development Indicators*. (2021). https://data.worldbank.org/

8. Appendix

Appendix 1: Description of variables

Variable	Label
Log(ResPat)	Log of resident patent applications per 100.000 capita's
Total Entrepreneurship	Sum of TEA and Established business ownership
Established business ownership	Percentage of 18-64 population who are currently an owner-manager of an established business, i.e., owning and managing a running business that has paid salaries, wages, or any other payments to the owners for more than 42 months.
TEA	Percentage of 18-64 population who are either a nascent entrepreneur or owner-manager of a new business.
Motivational Index	Percentage of those involved in TEA that are improvement-driven opportunity motivated, divided by the percentage of TEA that is necessity-motivated
Governmental Support for Entrepreneurs	The extent to which public policies support entrepreneurship - entrepreneurship as a relevant economic issue
Log(R&D Expenditure pc)	Log of the R&D expenditure per capita
Log(Business Registered)	Log of the annual number of businesses registered.
Ease of Doing Business	Rank between 1-100, a higher value indicates better regulations for businesses and stronger protection of property rights.
Perceived opportunities	Percentage of 18-64 population (individuals involved in any stage of entrepreneurial activity excluded) who see good opportunities to start a firm in the area where they live.
Log(GDP_PC)	Log of GDP per capita
Log(Population)	Log of Population
Log(FDI per capita)	Log of FDI inflows per capita
Log(Education Expenditure per capita)	Log of Expenditure on education per capita

Appendix 2: Hausman (1978) specification test

	Coef.
Chi-square test value	120.403
P-value	0.000

Appendix 3: Hausman (1978) specification test

	Coef.
Chi-square test value	11.273
P-value	.939

Appendix 4: Hausman (1978) specification test

	Coef.
Chi-square test value	37.775
P-value	0.000