



## **Master Thesis**

# **A Study of Relationship between Patents and Nascent Entrepreneurial Activity**

Program: Entrepreneurship and Strategy Economics

Thesis author: Alexandr Basso

Thesis supervisor: Dr Philipp Koellinger

## Table of content

Thesis abstract	3
Thesis introduction	4
Research question	5
Theory and Literature	5
Data	9
Research methodology	11
Analysis and Results	14
Test 1	14
Test 2	16
Test 3	19
Test 4	24
Test 5	26
Discussion and Conclusion	27
Used Literature	29
Appendix	31

## **THESIS ABSTRACT**

Entrepreneurship has gained an important role in eyes of both policy makes, as well as researchers around the world. The effects of that can be felt through the increasing support provided by the governments toward entrepreneurial activity. To date, several links have already been made between entrepreneurial activity and economic growth, but what about patents? Can patents be used to represent the knowledge factor in the economic growth? This paper performs regressions on nascent entrepreneurial activity and patent application for eleven OECD countries in order find what kind of relationship exists between the two. The product of this research provides indication to a possible relationship and further study of this topic could help explain entrepreneurship as a knowledge filter.

Keywords: nascent entrepreneurial activity, patents, knowledge, innovation

## THESIS INTRODUCTION

In recent years, *entrepreneurship* has been recognized to have a significant role in innovation and economic growth, as well as its contribution to the prosperity and economic welfare (Schramm, 2006, Acs & Armington 2006, Hoenig 2005, Audretsch 2006, Acs et al 2012). Following the economic crises (dot-com bubble 2000-2001, US housing bubble 2006-2007, global financial crisis 2007–2012) that happened in that last decade, a lot of problems have risen: high unemployment rates, drop in market revenues and demand. One of the measures taken by the policy makers to deal with those problems is the promotion of entrepreneurship. There are two types of entrepreneurship (Louis Jacques Filion, 1998) and out of the two it is the innovative, not the replicative entrepreneurship that brings new wealth to the economy. What policy makers aim to promote is entrepreneurship, however *nascent entrepreneurship* is what leads to future start-ups, as such many of policies have been aimed at assisting and encouraging those that are in the process of making a business.

Furthermore, in current day and age, to a certain degree, with enough funding it is rather easy to copy someone's product or idea, as such individuals and companies aim to protect themselves through *patenting* of those ideas. It can be seen that patents represent new ideas and those ideas in consequence create innovation and fuel entrepreneurship. Even though it is in the stage of emergence, it can be said that patents are a market of their own (Mark A. Lemley, 2007).

Some research has already been done on the topics related of patents and entrepreneurship (Scott Shane, 2002), however most of the papers only touch on the subject, but do not try to find the actual relationship between the two. To make the study narrower, this paper observes the *relationship between patents applied and nascent entrepreneurial activity*. In earlier studies, the relationship between the two has been overlooked to a point of being taken for granted, be that for lack of data, or taking a positive relationship as something that is a given. This study aims to clarify and bring some light to the matter of the whether there is strong correlation between the two, or if there is any relationship at all.

## RESEARCH QUESTION

The aim of the research is to bring increasing attention to the importance of the relationship between patents and entrepreneurship. As already mentioned, this relationship is not provided with enough attention and further consideration on patents policies can change entrepreneurial activity for the better. Thus, in order to bring more attention to the subject, this thesis is aimed answering the following specific research questions:

### **Research Question:**

"What kind of relationship do patents and nascent entrepreneurship activity have within OECD countries? Can high correlation be confirmed, or is there any relationship at all? Is it possible to relate one to the other with causation, not simply correlation?"

## THEORY AND LITERATURE

In order to achieve a coherent theory as to why would patents have an effect on entrepreneurship, let us take a look back on the first time that knowledge was recognized as not the same, as an economically relevant knowledge. In 1956 Solow's article triggered a large literature linking the traditional factors of production, capital and labor to economic growth. He observed that contribution of additional labor and capital could not explain increases in growth over time. Arrow (1962) was the first one to suggest that spillovers may not occur automatically, recognizing that knowledge is not the same thing as economical relevant knowledge. The endogenous growth proponents at a later time picked up that idea in order to incorporate knowledge in to the growth model - this is where history truly starts for the aggregate knowledge.

Endogenous growth theory core point is that economic growth is primarily the result of endogenous and not exogenous factors. Under this theory, investments into innovation, human capital and knowledge are the big contributors to economic growth. There are many factors that are considered under endogenous growth theory; some of them are positive externalities and spillover effects, which are also part of the main focus of this theory. Other

effects are also considered in some models, for example policy measures can have an impact on the long run growth rate of an economy. The simplest representation of endogenous model would be the AK model:

$$Y = A K^{\alpha} L^{1-\alpha}$$

This is a Cobb-Douglas function where  $Y$  represents the total product in an economy.  $A$  represents the total factor productivity,  $K$  is capital,  $L$  is labor and  $\alpha$  measures the output elasticity of capital. As already mentioned this is the simplest representation, but this should be enough to give an idea, of how the models look.

The main intellectual achievement by the endogenous growth theory is the recognition that an investment in the knowledge and human capital endogenously generates economic growth by means of spillover of knowledge. In history there have been several steps in the evolution of the endogenous growth theory. The first one is a natural response to the presence of market opportunities, which is the formation of knowledge and human capital. The first to provide endogenous growth models, with knowledge incorporated in to the model, where Romer (1986), Luca (1988), Rebelo (1991) and many others. One thing in common for all proposed models was that they focused mainly on the influence of knowledge spillovers on growth without providing any explanation as to how the knowledge actually spills over. Nevertheless, even without explaining how the knowledge spills over, the important issue of spillover of knowledge has been addressed, with aggregate knowledge capital being defined as a composition of R&D and human capital. To a degree the lack of explanation as to how did the knowledge spillover remained in the second step of evolution of the endogenous growth theory proposed by Schmitz (1989), Aghion and Howitt (1992), Chend and Dinopoulos 1993, Segerstrom (1995) and others. In this step, the neo-Schumpeterian models proposed where defining entry as an R&D race where only a small part of that R&D will bring about successful innovations. Even though compared to the first step it was an improvement, the problem with proposed models was that the Schumpeterian entrepreneur is not present in any of them. Schumpeter (1947) said: ‘the inventor produces ideas, the entrepreneur “gets things done” ..... an idea or scientific principle is not, by itself, of any importance for economic practice.’, therefor simple interpretation of R&D races, that mostly involve incumbents and focus on quality improvements of existing goods, is not enough to describe innovation process.

Later on, in paper of Acs et al. (2004) an attempt is made in extending the endogenous growth model and resulted in model that suggest how knowledge is converted into economic knowledge, as well as how economic knowledge influences growth. Building on Romer (1990) a model is elaborated in which there are two methods of developing new products. Same as in the original, incumbents undertake R&D by employing research labor, which generates new knowledge. Even though knowledge is generated, it does not constitute economic growth. By making use of entrepreneurship, economic growth can be distributed in countries that do not have high R&D spending through firm startups. In that paper entrepreneurship is taken as one type of mechanisms that reduces the knowledge filter. The process of starting a new firm commercializes knowledge that might not be commercialized under different circumstances. Those results are consistent with at that time existing empirical evidence, linking measure of entrepreneurship to economic growth. The end conclusion for that paper is that growth is enhanced through individual entrepreneurs exploiting knowledge, without producing any knowledge themselves. This conclusion reflects well on how entrepreneurship promotion has been targeted in the policies in most OECD countries.

Further work is done on the same subject in “Growth and Entrepreneurship” by Acs et al (2012), where authors suggest that spillover knowledge does not occur automatically as normally assumed in endogenous growth model. Entrepreneurship is considered as one of those mechanisms facilitating the spillover of knowledge, thereby contributing to economic growth. By using panel data of 18 countries, authors provide empirical evidence that besides R&D and human capital, entrepreneurial activity, also serves to promote economic growth, more specifically that it has a positive and systematic impact on economic growth. The model that is used by authors makes use of several testable predictions. Starting with growth being influenced by R&D spending and entrepreneurship at the country level, followed by assumption that countries with relatively low R&D spending may still be able to have high growth because of larger share of entrepreneurship. With that trend of thought R&D and entrepreneurship to a degree could be considered either substitutes, or complements to each other. To further point at the importance of knowledge spillover conduits, it is important to notice that the empirical results suggest that regardless of R&D spending, those countries that generate a greater degree of entrepreneurial activity exhibit systematically higher rates of

economic growth. These findings are consistent with the idea that nascent entrepreneurship is related to patents, which represent non-economical relevant knowledge.

Another paper that is of interest for this thesis is a survey and statically analysis done by Bo Carlsson and Ann-Charlotte Fridh (2002), where authors attempt to better understand the role of office of technology transfer. Those offices exist in order to better facilitate commercialization of research results in universities and similar institutions. The results of their study showed that technology transfer from universities to commercial sector needs to be understood in a broader context, hinting at the need of finding the proper balance between the basic function of teaching and researching within the universities on the one side and providing service to the larger community on the other. Several stages in the technology transfer process can be observed, each associated with its own outcome: invention disclosure, patent application, patent issued, license sold, license income and/or business start-ups. Interesting results come out from their statistics, only about a half of the invention disclosures result in patent applications; out of those a half of the applications result in patents; only a third of patents are licensed, and only a handful (ten to twenty percent) of licenses yield substantial income. That lead to median gross license income in 1996 of 2.2 million dollars and the median income per license was 80,000 dollars. This shows how little knowledge is actually made in to economically relevant knowledge.

Following the evolution of endogenous growth model to the latest development where knowledge spillovers are accounted for, as well as empirical results for the relationship between patents and commercialized knowledge, it is safe to assume that patent applications can serve well as an indicator for economical knowledge and nascent entrepreneurship as the knowledge filter. This assumption provides sufficient grounds to search for the relationship between nascent entrepreneurial activity and patents. From this the main hypothesis can be derived.

**Main Hypothesis:** Nascent entrepreneurial activity has positive relationship to the number of patent applied.



## DATA

This study makes use of GEM data in order to formulate the nascent entrepreneurial activity and OECD data to account for patents growth through the years. Control variables will be taken from those datasets: such as real GDP-growth, TEA, employment rates. The GEM is a worldwide research project, thus it allows for comparison across large number of countries. This database is designed to analyze and describe entrepreneurial activity, as well as institutional conditions. The GEM dataset has individual data, as such in order to make it useful, aggregate data has been computed by combining specific variables, using the right filters and computing the means in Stata. The OECD data is already aggregated; as such there should be no need for aggregation of values for those variables.

All the countries considered in this study participated each year from 2002 till 2008 in GEM and OECD surveys. The data has filtered people by age and limit them to the age of 18 to 64 years old, because this is considered to be the working population in most countries.

Countries used in this dataset are: USA, Netherlands, Belgium, France, Spain, UK, Denmark, Norway, Japan, Ireland, Finland.

The main use of GEM is to formulate the nascent entrepreneurial activity; however we take another important variable from them: TEA - an index to measure early-stage and new entrepreneurial activity. This index comprises nascent entrepreneurs and young business owners (business that existed for less than 42 months). Total Entrepreneurial Activity is used in order to derive the nascent entrepreneurial activity, as such it can be used in the model as a control variable.

The OECD data is mainly used to provide the information about the amount of patents generated for the selected countries and as a reference point the amount generated in the world for the relevant years. Four variables are made to represent patents: 1) Patents applied under European Patent Office worldwide, 2) Patents applied under Patent Cooperation Treaty worldwide, 3) Patents applied under European Patent Office in individual countries and 4) Patents applied under Patent Cooperation Treaty in individual countries. The first two represent worldwide level of applied patents under two means (EPO and PCT) and the second two represent patents applied in individual countries (11 OECD countries) under the same

two means. The main differences between the EPO and PCT, is that under the first one patents are applied for the euro zone – meaning that patents are granted for all members of European Patent Office, while the second one is called the international application for country members around the world – this application form reaches a lot more countries.

To present date, there is no data available or known way to formulate how relevant each patent is, as such we will consider all patents equality important. Two of the variables to be used from this data are the unemployment rate and real GDP-growth. Those control variables are necessary to see the relationship between patents and nascent entrepreneurship.

**Table 1. Description of all variables - partial.**

*Full Description of all used variables can be found in the Appendix Table 1.1*

**Dependent variables,  
GEM**

Variable notation	Definition
<b>Nascent entrepreneurship Entrep (suboanw)</b>	= the rate of individuals who claim to be engaged in starting a business they will at least partially own and that has paid wages or profits for no more than 6 months.

**Independent variables,  
OECD data**

<b>European Patent Office PA_EPO_world (PA_EPO)</b>	Patents applied under European Patent Office worldwide
<b>Patent Cooperation Treaty PA_PCT_world (PA_PCT)</b>	Patents applied under Patent Cooperation Treaty worldwide
<b>European Patent Office in individual country PA_EPO_country (PA_EPO_ind)</b>	Patents applied under European Patent Office in individual country
<b>Patent Cooperation Treaty in individual country PA_PCT_country (PA_PCT_ind)</b>	Patents applied under Patent Cooperation Treaty in individual country

**Other variables,  
GEM, OECD**

<b>TEA</b>	= Total early-stage entrepreneurial activity
<b>GDP</b>	= Billion US dollars, current prices and PPPs, real GDP
<b>Harmonized unemployment rate (unempl)</b>	= comparable unemployment rates
<b>Unempl1-Unempl6</b>	= time lags of unemployment rate

Notes:

For better understanding of the reader, variables mentioned in the paper are called differently from what they are called in the dataset.

Names in the brackets is how variables are named in the dataset - ( )

## RESEARCH METHODOLOGY

In this section the methodology to find the relationship between patents and nascent entrepreneurial activity will be discussed. Five tests will be performed in order to find out the relationship between nascent entrepreneurship activity and patents.

In the *First Test*, linear correlation is to be checked, to see if there is any correlation between the variables at all. This is done with the help of "pwcorr" function in Stata, it displays all the pairwise correlation coefficients between the variables specified. For this test all control variables are included, independent variables and dependent variables (Entrep, unempl, PA\_EPO\_country, PA\_PCT\_country, PA\_EPO\_world, PA\_PCT\_world, PA\_EPO\_country2, PA\_PCT\_country2, PA\_EPO\_world2, PA\_PCT\_world2, GDP), no lags or dummies are included. After running the test found correlations will be evaluated based on strength (None-Strong) and sign (Positive-Negative). This should give a general idea of the relationship between the variables in the dataset.

In the *Second Test*, tests will be performed in order to check for the possible multicollinearity issues. It is necessary to test for multicollinearity in all eleven OECD countries individually and on worldwide scale, for both patents applied under European Patent office and Patent Cooperation Treaty . This is done by the normal "collin" function in Stata. It is possible to establish, if there are multicollinearity problems by looking at the VIF (variance inflation factor) and the R-squared.

Furthermore, since the purpose is to test the main hypothesis, it is best to establish, if normality is present. Normality assumption provides greater certainty that p-values for the t-tests and F-test will be valid. This is done by examining the plot of a standardized normal probability ("pnorm" in Stata) and the plot of the quantiles of a variable against the quantiles of a normal distribution ("qnorm" in Stata). "pnorm" is sensitive to non-normality in the middle range of data and "qnorm" is sensitive to non-normality near the tails. In addition Shapiro-Wilk W test is performed to further check for normality, this test assumes that the p-value is based on the assumption that the distribution is normal, by examine the p-value we can see, if the residuals are normally distributed.

*Third Test* runs fixed and random effects linear models on the dataset, while applying Hausman test on both of them. In simplest terms, random effects model stands from the point where the independent variables have a random nature, while the fixed effect model considers the independent variables to have no random nature. It is fair to assume that more important test here is the fixed effects model, because within fixed effects model there is a consideration that exogenous effects can play a role in the model, however random effects model can provide better results, if the model is described well (more complete). However, since we perform Hausman test, we can see which of the two is more suited to be interpreted. Following Nickell (1981) and Diggle (2002), it can be concluded that the fixed effects model eliminates unobserved time-invariant of individual effects with use of the within transformation. As for the random effects model, it is assumed that all errors are already counted in the model - no exogenous effects exist, thus no outside effects are considered. (Christensen, 2002 and Baltagi& Chang, 1994). It is rather impossible for all possible error reasons to be included in to the model; as such random effect model seems to be of a lesser use.

This does not meant that random effects model will be ignored, RE model will be considered as much as fixed effects model and in the end, the two will be compared. Normally control variables should not be used in random and fixed effects models; however time dummies will be introduced, as well as unemployment rate and GDP as a control variable in order to at least to a degree account for necessity entrepreneurial activity. Afterwards, the Hausman specification test will be used in Stata to test the difference between the random and fixed effect model. This test is mostly to see how great of the effect do the exogenous factors make in our model and help decide which of the two tests is best suited to be considered for interpretation.

The tests will be performed in two steps: for country level data and for worldwide level data. This is necessary to see if there is a difference between the local and global effects of patents on entrepreneurship. The tests will also be performed first without control variables and then with control variables in order see the difference in effects.

The effect of patents on nascent entrepreneurship regardless, if it is positive or negative is very small. The reason for that is because only about half of all inventions disclosures in U.S. universities result in patent applications and half of those applications result in patents. About one third of patents are licensed and only ten to twenty percent of

licenses yield significant income (Carlsson and Fridh 2002, page 231). Meaning that only about two percent of inventions are successful in reaching the market and providing income and not all of those are used by nascent entrepreneurship, some are integrated in already established businesses. As already mentioned before, patent applications do not relate in a one to one ratio to successful new startups, existence of entrepreneurship is what slightly increases conversion of general knowledge in to economically relevant knowledge. This statistic forces us to expect very minuscule betas, as such for evaluation elasticity (  $d(\ln y)/d(\ln x)$  ) will be provided instead. Elasticity's are used to show the ration of the percentage change in nascent entrepreneurship to the percentage change in the patents applied under two policies (EPO and PCT).

In the *Forth Test*, tests are performed in order to check, if there is a non-linear relationship between the nascent entrepreneurship and patents. Compared to the third test, the squared value of the independent variable are added to the model, this should help ascertain, if there are hints to the relationship being non-linear. Same as in the previous test, this will be done without control variables and with control variables, however only for the individual countries.

The *Fifth Test* is more of a robustness check then an actual test. Same as in the fourth test an additional variable is added, however this time one year lag are used instead of squares of the main independent variable. The same random and fixed effects models will be used for country level and worldwide level for both types of patent application. This test is performed to see, if the coefficients change much compared to the original regressions. In case of not a significant change, it can be considered that coefficients are robust, also if signs and magnitudes of the estimated regression are also plausible, this is usually taken as evidence that regression coefficients can be reliable interpreted as the true causal effects Halbert White and Xun Lu (2010).

Note: World level tests they are not expected to yield any significant results and they are to be compared to the country level regressions for comparison. The tests that are performed on the world level are performed only for the referential purposes.

## Analysis and Results

### First Test - Testing for Linear correlation

First, testing for the correlation between the variables is performed, in order to see the measure of the strength and direction of the linear relationship between two variables. This is done with "pwcorr" command in Stata, which tests the pair wise correlation.

Pearson product-moment correlation coefficient, also known as  $r$ ,  $R$ , or Pearson's  $r$ , a measure of the strength and direction of the linear relationship between two variables that is defined in terms of the covariance of the variables divided by their standard deviations.

The strength of the relationship is in accordance to Cohen, J. (1988) and is defined as such:

Correlation	Negative	Positive
None	-0.09 to 0.0	0.0 to 0.09
Small	-0.3 to -0.1	0.1 to 0.3
Medium	-0.5 to -0.3	0.3 to 0.5
Strong	-1.0 to -0.5	0.5 to 1.0

**Table 2. Pair Wise Correlation table for all variables used in the regressions**

	suboanw	unempl	PA_EPO _country	PA_PCT _country	PA_EPO _world	PA_PCT _world	PA_EPO _country2	PA_PCT _world2	PA_EPO _country2	PA_PCT _country2	GDP
suboanw	1										
unempl	-0.1641	1									
PA_EPO _country	0.4362*	-0.1593	1								
PA_PCT _country	0.5220*	-0.1908	0.9778*	1							
PA_EPO _world	0.0396	-0.0190	0.0208	0.0535	1						
PA_PCT _world	0.7326	0.8695	0.8575	0.6437	0.1339	1					
PA_EPO _world2	0.0397	-0.0184	0.0210	0.0533	0.9999*	0.1401	1				
PA_PCT _country2	0.5852*	-0.1663	0.9675*	0.9752*	0.0294	0.0166	0.0297	1			
PA_PCT _world2	0.0344	0.0540	0.0077	0.0080	0.7999	0.8857	0.7976	0.0128	1		
PA_PCT _country2	0.7668	0.6409	0.9467	0.9452	0.0191	0.0000	0.0168	0.9123	0.0094	1	
GDP	0.6028*	-0.1035	0.9318*	0.9701*	0.0215	-0.0095	0.0211	0.9533*	0.0028	0.9728*	1
	0.0000	0.3703	0.0000	0.0000	0.8527	0.9348	0.8556	0.0000	0.9809	0.0000	

## **Results for first test:**

From table above we can see that Patents applied under European Patent Office and Patent Cooperation Treaty in individual countries have a positive medium level of relationship with Nascent Entrepreneurship of 0.4362 and 5220, with both of those results being statistically significant on 95% confidence interval.

It is interesting to note that the relationship between the nascent entrepreneurship and Patents applied around the world is not statistically significant. This should indicate that there are some effects on the world scale that are not present when observing the relationship in the individual countries.

It should be also interesting to point out that on average an increase in patent application both in individual countries and around the world is correlated to a decrease in unemployment.

Control variables GDP have a strong positive linear relationship with nascent entrepreneurship, while the unemployment rate has a small negative not statistically significant relationship. This strong correlation of GDP is expected, as there is strong relationship between growth and entrepreneurship Acs et al (2012).

GDP has a strong positive linear relationship with the Patents applied under both patent EOP and PCT, which indicates strong multicollinearity problems. Even though this problem is present, excluding GDP from regressions is not advised; as such GDP will still be used as a control variable. On a further note, GDP having a strong relationship patents applied seems intuitively correct, as wealthier countries generally do have a higher invest in Research and Development

## **Second Test - Testing for Multicollinearity**

Since in our models we are going to use several variables, it is best to test for multicollinearity, in addition to the pair wise correlation test. The main problem is that the more variables are used in models the more the degree of multicollinearity increases, as such the regression model estimates of the coefficients become unstable and the standard errors for the coefficients can get significantly inflated. This test will be done with "VIF" command



in Stata. VIF stands for *Variance Inflation Factor*, generally VIF greater than 10 is a problem, however VIF below 10 can be tolerated, even though it does not indicate a good degree of inflation. R-squared will also be a good indication of how well the variables fit together. Results for the VIF test can be seen below in the Table 3.

### **Table 3 Multicollinearity Test - VIF (variance inflation factor)**

**Table 3.1 Multicollinearity Test - for Patents applied under European Patent Office in individual country's**

Variable	VIF	SQRT VIF	Tolerance	R-squared
PA_EPO_country	7.82	2.80	0.1278	0.8722
GDP	7.71	2.78	0.1298	0.8702
unempl	1.04	1.02	0.9593	0.0407
Mean VIF	5.52			

**Table 3.2 Multicollinearity Test - for Patents applied under Patent Cooperation Treaty in individual country's**

Variable	VIF	SQRT VIF	Tolerance	R-squared
PA_PCT_country	19.71	4.44	0.0507	0.9493
GDP	19.20	4.38	0.0521	0.9479
unempl	1.18	1.08	0.8507	0.1493
Mean VIF	5.52			

**Table 3.3 Multicollinearity Test - for Patents applied under European Patent Office worldwide**

Variable	VIF	SQRT VIF	Tolerance	R-squared
PA_EPO_world	1.00	1.00	0.9993	0.0007
GDP	1.01	1.01	0.9889	0.0111
unempl	1.01	1.01	0.9890	0.0110
Mean VIF	1.01			

**Table 3.4 Multicollinearity Test - for Patents applied under Patent Cooperation Treaty worldwide**

Variable	VIF	SQRT VIF	Tolerance	R-squared
PA_PCT_world	1.01	1.00	0.9929	0.0071
GDP	1.01	1.01	0.9893	0.0107
unempl	1.02	1.01	0.9823	0.0177
Mean VIF	1.01			

After the VIF test, with Stata commands "pnorm" and "qnorm" tests are performed, to draw the graphs that will help check for non-normality in the middle range of data and near the tails range of data. With this we will see whether, the residuals have normal distribution or not, check for Table 4 in the Appendix.

To further test for normality, Shapiro-Wilk W-test is performed in order to test the null hypothesis that sample came from a normally distributed population. If the p-value is less than the alpha level of 0.05, then data is not from a normally distributed population and the other way around in case it is greater than five percent.

### **Results for the second test:**

With the results of the VIF test we can see a clear multicollinearity problem, for the patents applied under Patent Cooperation Treaty on the country level, VIF of 19.71 and R-squared of 0.9493. It is clear that the results for the regressions with this combination of variables will have problems.

The other VIF test for patents applied under European Patent Office individual countries is borderline clear. VIF of 7.82 and R-squared of 0.8722, are not good results, however still acceptable, which means this combination of variables should be used with care.

The problems arising in these two tests stem from the high correlation between GDP and Patents applied in the individual countries, nothing can be done about that, this can be observed in Table 2. On the other hand since GDP has less of a correlation on the world level it means that VIF and R-squared are rather low, hinting at little possibility of multicollinearity problems. This low effect of GDP on the world level is due to GDP referring to each individual country for specific years, not GDP around the world; as such the good results are not anything unexpected.

From Table 4 we can see that the residuals have small deviation from normal distribution, which indicates that non-normal distribution is possible both in the middle and near the tail end of the data. In addition the p-value of Shapiro-Wilk W test is 0.039, which tells us to reject the null hypothesis, meaning that the used data is not from a normally distributed population. This is rather troublesome, as this indicates that the data sample is not normally distributed, in addition to the multicollinearity problem. Over all such results tell us that hypothesis testing might have additional problems.

### **Third Test - FE & RE, Hausman Test**

In the third step, Housman specification test is being used in order to find out, if  $H_0$  is supported or not. The test will be performed in two stages, once without any control variables and then once again with control variables:

- Country level - NE & PA\_EPO\_country NE & PA\_PCT\_country
- World level - NE & PA\_EPO\_world NE & PA\_PCT\_world

From the test we observe R-Squared, Rho and Hausman values, which should tell us if the model is fully complete and determine, if there is a positive or a negative association between the dependent and independent variables. Also we perform this test to see whether fixed effect or random effect is efficient for our scenario or not. Unemployment rate is included in this mode in order to see how much of nascent entrepreneurship is being driven by unemployment. This is a significant factor as we assume that those that are forced in entrepreneurship by unemployment are not greatly affected by patents, since they are mostly not innovative. Even though it has high correlation with the independent variables, GDP is included as control variable as well.

## Results for the third test:

**Table 5.1 FE & RE, Hausman Test regression results of nascent entrepreneurial activity across 11 countries on patents applied on a Country level dataset, 2002-2008**

<i>Step 1: Regressions on Country level dataset (N=77) (11 countries) (with OUT control variables)</i>				
	Fixed effects	Random effects	Fixed effects	Random effects
PA_EPO_country	.380 (.15) [.000]	.166 (0.04**) [.000]		
PA_PCT_country			.155 (.28) [.000]	.151 (.01**) [.000]
Constant	2.168( 0.01**) [.826]	2.815(.00**) [.527]	2.799 (.0**) [.544]	2.813 (.0**) [.486]
<b>Model diagnostics</b>				
Hausman Test (Prob> Chi2)	.993		1.000	
R <sup>2</sup>	.053	.193	.037	.2755
Rho	.821	.725	.680	.6900
<i>Step 2: Regressions on Country level dataset (N=77) (11 countries) (WITH control variables)</i>				
PA_EPO_country	.165 (0.6) [.000]	-.044 (.68) [.000]		
PA_PCT_country			.289 (.08*) [.000]	.075 (.60) [.000]
GDP	-.177 ( 0.37)[.000]	.181 (.05*) [.000]	-.398 ( .03) [.000]	.070 (.62)[.000]
Unemploymentrate	-.404 (.13) [.145]	-.252 (0.16) [.097]	-.314 (.23) [.143]	-.229 (.24 ) [.106]
Constant	4.756 ( .01**) [1.845]	3.566 (.0**) [.702]	4.840 (.0**) [ 1.105]	3.522 (.0**) [.746]
<b>Model diagnostics</b>				
Hausman Test (Prob> Chi2)	.0008		0.000	
R <sup>2</sup>	.098	.408	.140	.321
Rho	.832	.430	.880	.504

Explanatory notes:

Instead of beta's, elasticity is displayed -  $d(\ln y)/d(\ln x)$

Significance level in first parentheses. = ( )

Standard errors in second parentheses. = [ ]

All models include time dummies.

\*\* significant at > 95% confidence

\* significant at > 90% confidence

The R<sup>2</sup> of the random effects model is the overall R<sup>2</sup> and the R<sup>2</sup> of the fixed effects models is the within R<sup>2</sup>

The first stage of this test demonstrates the relationship between the number of patents applied under the two patent options (EPO and PCT) on country level against nascent entrepreneurship, with and without control variables (Table 5.1). From the results of Hausman test we can see that for the models without control variables Random effects specification fits better. The Hausman test value fluctuates slightly, but basically stays at approximately value "1.0" in all three steps. Value of 1.0 is obviously greater than 0.05, which is an insignificant result for the Hausman test, meaning that  $H_0$  is supported and random effect model is safe to be used, since it is more efficient one out of the two. This means that Hausman test high value indicates that the unobserved heterogeneity is highly

significant, thus supporting random effects specification. On the other hand, the results of Hausman test for the model with control variables has fluctuates around "0", making the fixed effects specification test more appropriate. As such analysis will be done based on Random effects for the model without control variables (step 1) and fixed effects for the model with control variables (step 2).

For the model without control variables (step 1, table 5.1), R-squared in the regressions varies from 0.193 to 0.2755, which is not all that high. Meaning that our model does not predict the future outcomes too well and is not complete in terms of explaining the relationship between the variables, however R-squared is at best a descriptive measure and should not impact the interpretation of results too much. It should be noted that the greater values of 0.193 to 0.2755, are present for the random effects model. This goes in support of results of the Hausman test. As already mentioned the Rho value, reflects the positive or negative association between the dependent and independent variables. The regression results show Rho fluctuation between 0.69 and 0.725, which suggests a strong positive association, meaning that nascent entrepreneurial activity tends to increase when independent variables increase, our Rho shows a positive coefficient. The elasticity's show positive ratios for dependent and independent variables and for both patents under EPO, and PCT policies are statistically significant on 95% confidence interval, the same relationship can be observed for the constants. This is the first indication to a clear relationship between nascent entrepreneurial activity and patents applied, which supports the main hypothesis.

For the model with control variables (step 2, table 5.1), R-squared in the regressions varies from 0.098 to 0.14, which is too low. Meaning that our model does not predict the future outcomes well and is not complete in terms of explaining the relationship between the variables, however same as before R-squared is at best a descriptive measure and should not impact the interpretation of results too much. The regression results show Rho be around .832 and 0.88, which suggests a strong positive association, which means that nascent entrepreneurial activity, tend to increase when independent variables increase, this Rho shows a positive coefficient. The elasticity's for the second step for the independent variable and the constants are similar to the step one, however only the PCT variable is significant on 90% confidence interval, while EPO is not significant at all. This can be explained by the fact that control variables that are highly correlated with the independent variables cause multicollinearity issues, in addition to the already observed issues with R-squared. The

control variables seem to have a negative relationship with nascent entrepreneurship, which is counter intuitive (Acs et al 2007). It would be interesting to notice that for the random effects of step 2, the elasticity has a negative relationship; this is rather strange and hints at further problems in the model. This however could be dismissed, as the random model should not be used according to the Hausman test for that case.

**Table 5.2 FE & RE, Hausman Test regression results of nascent entrepreneurial activity across 11 countries on patents applied on a World level dataset , 2002-2008**

<i>Step 1: Regressions on World level dataset (N=77) (11 countries) (with OUT control variables)</i>				
	Fixed effects	Random effects	Fixed effects	Random effects
PA_EPO_world	.556 (.63) [.000]	-.468 (.83) [.000]		
PA_PCT_world			.019 (.83) [.000]	.019 (.83) [.000]
Constant	1.547 (.69)[3.891]	4.815 (.50) [7.293]	3.238 (.0**) [ .277]	3.238 (.0**) ( .513)
<b>Model diagnostics</b>				
Hausman Test (Prob> Chi2)	1.0000		Does not fit Hausman test	
R <sup>2</sup>	.018	.004	.018	.004
Rho	.763	.754	.763	.754
<i>Step 2: Regressions on World level dataset (N=77) (11 countries) (WITH control variables)</i>				
PA_EPO_world	.560 (0.6) [.000]	-.795 (.72) [.000]		
PA_PCT_world			.046 (.61) [.000]	.033 (.72) [.000]
GDP	-.235 (.37)[.000]	.115 (.06*)[.000]	-.235 (.15)[.000]	.115 (.06*)[.000]
Unemploymentrate	-.428 (.13) [ .145]	-.311 (0.12) [.111]	-.428 (.10) [.141]	-.311 (.12) [.111]
Constant	3.764 (.01**) [ 4.064 ]	6.399 (.39) [ 7.482 ]	5.356 (.0**) [ 1.002]	3.722 (.0**) [.792]
<b>Model diagnostics</b>				
Hausman Test (Prob> Chi2)	.365		.525	
R <sup>2</sup>	.094	.339	.094	.339
Rho	.911	.667	.911	.667
Explanatory notes:				
Instead of beta's, elasticity is displayed - d(lny)/d(lnx)				
Significance level in first parentheses. = ( )				
Standard errors in second parentheses. = [ ]				
All models include time dummies.				
** significant at > 95% confidence				
* significant at > 90% confidence				
The R <sup>2</sup> of the random effects model is the overall R <sup>2</sup> and the R <sup>2</sup> of the fixed effects models is the within R <sup>2</sup>				

The second stage of this test demonstrates the relationship between the number of patents applied under the two patent options on worldwide level against nascent entrepreneurship, with and without control variables (Table 5.2). From the results of Hausman test we can see that for the models without control variables Random effects specification fits better, however there arises a problem in model for the Patents applied

under Patent Cooperation Treaty, as the regression of Fixed and Random effects no longer fits the Hausman test, which means that independent variable have no significant effect on nascent entrepreneurship. As for EPO, Hausman test has a high value that indicates that the unobserved heterogeneity is highly significant, thus supporting random effects specification. The results of Hausman test for the model with control variables has fluctuates between 0.365 and 0.525, making the it hard to choose from fixed effects and random effects specification tests, as such in this study both of them will be used in the interpretation, with emphasis on Fixed Effect. As such analysis will be done based on Random Effects for the model without control variables (step 1) and both fixed effect & random effects specification test for the model with control variables (step 2).

For the model without control variables (step 1, table 5.2), R-squared in the regressions is 0.018, which is very poor, same as the results based on country level data. This meaning that this model does not predict the future outcomes well and is not complete in terms of explaining the relationship between the variables. As already mentioned the Rho value, reflects the positive or negative association between the dependent and independent variables. The regression results show Rho is 0.754, this suggests a strong positive association, which means that nascent entrepreneurial activity tend to increase when patents applied under European Patent Office for the world increases, our Rho shows a positive coefficient. The elasticity's show a negative ration for dependent and independent variables are not statistically significant. This is rather strange since these results are opposite to what is observed on the country level, which suggests presence other unobserved effects.

For the model with control variables (step 2, table 5.2), as both Random and Fixed effects are analyzed here and the values can be observed from the table above, it is best to concentrate on the similarities and differences between the two. Same as in the Step 1, no results are statistically significant and R-squared are rather low, even though random effects seem to provide a little bit better results, 0339 opposed to 0.094, suggesting that Random effects specification test could provide better statistical results, however in the end they are still questionable. For all regressions the Rho shows a positive coefficient. The elasticity for the Random effect of EPO on world level seem to provide a negative relationship, same as in step 1, but different from the results under the same conditions, but on a country level. Overall no conclusions can be drawn from the results of table 5.2, except for hints of greater unobserved factors that are not being accounted for in the model.

## **Fourth Test - FE & RE, Hausman Test, Testing for non-linear relationship**

In the fourth step, the possible non-linear relationship between nascent entrepreneurship and patents applied under two policies on a country level are being checked. The test will be performed with and without the control variables:

- Country level -            NE & PA\_EPO\_country                            NE & PA\_PCT\_country

Same as before, in the test we observe R-Squared, Rho and Hausman values. Unemployment and GDP will be included as the control variables, however this time the squared value of the independent variables, will also be included in order to observe the changes that occur in a non-linear relationship model in comparison to a linear relationship model that was tested in Test 3. Adding the squared value of the independent variable to the model should help ascertain, if there are hints to the relationship being non-linear. Same as in the previous test, this will be done without control variables and with control variables, however only for the individual countries.



## Results for the fourth test:

**Table 5.3 Non-Linear(SQR of the independent variable is included in the model) FE & RE, Hausman Test regression results of nascent entrepreneurial activity across 11 countries on patents applied on a Country level dataset, 2002-2008**

<i>Step 1: Regressions on Country level dataset (N=77) (11 countries)</i>		<i>(with OUT control variables)</i>		
	Fixed effects	Random effects	Fixed effects	Random effects
PA_EPO_country	-2.134 (.06*) [.000]	-582 (.04**) [.179]		
PA_PCT_country			.541(.15) [.377]	-.217 (.11)[.135]
SQR_independ_var	.914 (.02**) [.000]	.489 (.0**) [.116]	-.172 (.26)[.155]	.247 (.0**) [.088]
Constant	7.424 (.0**) [2.474]	3.820(.0**) [.422]	2.046(.02**) [.870]	3.277(.0**) [.392]
<b>Model diagnostics</b>				
Hausman Test (Prob> Chi2)	.6545		Does not fit Hausman test	
R <sup>2</sup>	.128	.614	.057	.554
Rho	.974	.468	.814	.344
<i>Step 2: Regressions on Country level dataset (N=77) (11 countries)</i>		<i>(WITH control variables)</i>		
PA_EPO_country	-2.203 (.05*) [ 1.146]	-.716 (.0**) [.189]		
PA_PCT_country			.388 (.29) [.373]	-.270(.1) [.167]
SQR_independ_var	.874 (.03**) [.407]	.473 (.0**) [.112]	-.049 (.76)[.166]	.221 (.01**) [.093]
GDP	-.140 (.46)[.192]	.148 (.05*) [.085]	-.372(.07*) [.205]	.087 (.53)[.140]
Unemploymentrate	-.386 (.13) [.258]	-.193 (.23) [.161]	-.320 (.23) [.267]	-.205(.24) [.174]
Constant	9.478 (.0**) [2.835]	4.315(.0**) [.648]	4.584 (.0**) [1.40]	3.851(.0**) [.676]
<b>Model diagnostics</b>				
Hausman Test (Prob> Chi2)	.966		0.001	
R <sup>2</sup>	.167	.658	.141	.576
Rho	.983	.439	.865	.362

Explanatory notes:

Instead of beta's, elasticity is displayed -  $d(\ln y)/d(\ln x)$

Significance level in first parentheses. = ()

Standard errors in second parentheses. = []

All models include time dummies.

\*\* significant at > 95% confidence

\* significant at > 90% confidence

The R<sup>2</sup> of the random effects model is the overall R<sup>2</sup> and the R<sup>2</sup> of the fixed effects models is the within R<sup>2</sup>

In the results above table 5.3, patent applied under European Patent Office in individual countries show signs of a non-linear relationship, compared to patents applied under Patent Cooperation Treaty. This can be due to the fact that most of 11 countries selected are from Europe. The data shows statistical significance for independent variable and surprisingly the elasticity's demonstrate a negative relationship for both models with and without control variables based on EPO. In both EPO models random effects test seems to be more appropriate in accordance with Hausman test, compared to fixed effects results R-

squared indicates that as well. This however is not observed for the PCT model, according to Hausman fixed effects is better suited for the PCT model with control variables and without control variables Hausman cannot be performed. Over all there are strong indications of a non-linear relationship between nascent entrepreneurship and patent applications.

**Fifth Test - FE & RE, Hausman Test; Robustness check;  
Testing with one year lags.**

In the fifth step, test is perform to check for robustness and it is done through the inclusion of one year lags in the models for both country and world level, with and without control variables. This test is performed in order to check how the “core” regression coefficient estimates behave when the regression specification is modified by adding one year lag regressor. In case of coefficients being plausible and robust, it is permutable to take that as evidence of structural validity.

The results can be found in Tables 6.1 and 6.2 in the Appendix. From these tables it is hard to ascertain with certainty that the regressions are robust, due to slight difference when compared to the main regressions in Tables 5.1 and 5.2. According to Halbert White and Xun Lu (2010), robustness checks can give neither necessary nor sufficient evidence for structural validity in case of bad use of additional regressor, as such the results should be handled with care.

## Discussion and Conclusion

At the start let us begin with the observed issues to this paper: there are several problems with the data used, starting with dataset being limited to only eleven OECD countries with time period of seven years, which makes it a total of 77 observation for the dataset. This is a rather small sample size to be performing fixed and random effects specification tests on. Other problems appear due to the strong Multicollinearity issues, caused by the strong correlation between main control variables and the independent variables, which cannot be avoided. Continuing with the number of patents applied not being able to represent the patents that got commercialized very well (patents turned in to economically relevant knowledge). This is a rather significant issue, as it causes problems for significance of regressions, decreasing the ability to detect a relationship between patents and nascent entrepreneurship which we assume to exist based on Acs et al (2012).

A careful examination of found results provides a blurry picture as to what kind of relationship exists between the nascent entrepreneurial activity and patent applications. However, even if a clear relationship was not confirmed there are indications that the relationship does exist and that the main hypothesis should not be rejected. That evidence can be observed in the regressions without the control variables for the country level under the both patent application policies and for regressions with control variables under the PCT. The relationship seems to indicate a positive relationship between the dependent and independent variables, supporting the main hypothesis. Regressions on the world level were performed only as reference for the country level and by comparing the two some similarities in elasticity's and signs do arise, however as initially expected the world level tests do not provide any significant results, as such no conclusions can be drawn. Thanks to the fourth test, signs of a non-linear relationship between applied patents and nascent entrepreneurial activity can be seen, however the sign of elasticity's and significances cause some doubt to be cast on the results. The one year lag test that was performed as robustness check, further confirms the uncertainty of the obtained results. There are some deviations from the main regressions in both coefficients and I magnitudes, indicating that the found coefficients may not be true casual effects. As was already mentioned, there are several problems observed in the course of the thesis, but some payoff can be drawn.

To put in a nutshell, there are indication that there is a relationship between nascent entrepreneurial activity and patents applied (under both European Patent Office and Patent Cooperation Treaty in individual countries), however they are not strong enough to draw a clear conclusion, be that to confirm or reject the main hypothesis. What can be said is that there is a positive correlation between the two, yet no causation can be drawn from it.

## Used literature

- Arrow, K., 1962, “The Economic Implication of Learning by Doing”, *Review of Economics and Statistics*, 80, 155-173.
- Acs, Z. J. & Szerb, L. (2008). A complex global entrepreneurship context index (CEC). *University of Pecs, Faculty of Business and Economics*
- Acs, Z. J. (2007), How is Entrepreneurship Good for Economic Growth? George Mason University, Max-Planck-Institute of Economics, Kauffman Foundation, Progress Foundation *The Beauty of Entrepreneurship*, Zurich, Switzerland, October
- Acs, Z. J. & Catherine Armington & Ting Zhang, (2006) "The Determinants of New-firm Survival across Regional Economies," Papers on Entrepreneurship, Growth and Public Policy 2007-04, Max Planck Institute of Economics, *Entrepreneurship, Growth and Public Policy Group*.
- Acs, Zoltan J., Audretsch, David B., Braunerhjelm, Pontus and Carlsson, Bo, (December 2004). “The Missing Link: The Knowledge Filter and Entrepreneurship in Endogenous Growth“, CEPR Discussion Paper No. 4783.
- Acs, Z. J., David B. Audretsch, Pontus Braunerhjelm and Bo Carlsson, (2012), “Growth and entrepreneurship”, *Small Business Economics*, Volume 39, p. 289-300
- Aghion, P., and Howitt, P., 1992, “A Model of Growth Through Creative Destruction”, *Econometrica*, v60, 323-51.
- Baltagi, B. H., & Y.-J. Chang. (1994). “Incomplete panels: A comparative study of alternative estimators for the unbalanced one-way error component regression model”, *Journal of Econometrics*, 62 (2), 67–89.
- Carlsson, B. and Fridh, A.-C., 2002, “Technology Transfer in United States Universities: A Survey and Statistical Analysis”, *Journal of Evolutionary Economics*, 12(1-2), 199- 232.
- Cohen, J. (1988), “Statistical power analysis for the behavioral sciences” (2nd ed.)
- Christensen, R. (2002). *Plane Answers to Complex Questions: The Theory of Linear Models*. New York: Springer.
- Cheng, L. and Dinopoulos, E., 1992, “Schumpeterian Growth and International Business Cycles”, *American Economic Review*, 82, 409-14.

- D. B. Audretsch & Max Keilbach, 2006. "Entrepreneurship, Growth and Restructuring," *Max Planck Institute of Economics, Entrepreneurship, Growth and Public Policy Group, Growth and Public Policy* 2006-13
- D.C. Mowery and S. Shane (2002), Special Issue on University Entrepreneurship and Technology Transfer: Introduction to the Special Issue on University Entrepreneurship and Technology Transfer, *Management Science*, 48:v-9
- Halbert White and Xun Lu (June 18, 2010), Robustness Checks and Robustness Tests in Applied Economics, *Department of Economics University of California, San Diego*
- Lucas, R., 1988, "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 3-39.
- M.A. Lemley, D Lichtman, (2007-2008 Rethinking Patent Law's Presumption of Validity, ) *Mark A.60 Stan. L. Rev.* 45
- Nickell, S. J., (1981). 'Economic policy and private investment since the oil crisis' by Artus et al, *European Economic Review, Elsevier*, vol. 16(1), pages 57-59.
- Oscar Torres-Reyna (2011), Panel Data Analysis, Fixed & Random Effects (using Stata 10.x) (ver. 4.1)
- P. J. Diggle, J. A. Tawn, R. A. Moyeed (2002), Model-based geostatistics, *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, Volume 47, Issue 3, pages 299–350, 1998
- Romer, P., 1986, "Increasing Returns and Economic Growth", *American Economic Review*, 94, 1002-1037.
- Rebelo, S., 1991, "Long-Run Policy Analysis and Long-Run Growth", *Journal of Political Economy*, 99, 500-521.
- Schramm, C. J. (2006). *The entrepreneurial imperative*. (New York: Collins)
- Schmitz, J., 1989, "Imitation, Entrepreneurship, and Long-Run Growth", *Journal of Political Economy*, 97, 721-739.
- Solow, R. (1956). "A contribution to theory of economic growth", *Quarterly Journal of Economics*, 70, 65–94.
- Segerstrom, P., 1995, "A Quality Ladders Growth Model With Decreasing Returns to R&D", *mimeo*, Michigan State University.

## Appendix

**Table 1.1 Description of all variables**

*Dependent variables,  
GEM*

Variable notation	Definition
<b>Nascent entrepreneurship Entrep (suboanw)</b>	= the rate of individuals who claim to be engaged in starting a business they will at least partially own and that has paid wages or profits for no more than 6 months.

*Independent variables,  
OECD data*

<b>European Patent Office PA_EPO_world (PA_EPO)</b>	Patents applied under European Patent Office worldwide
<b>Patent Cooperation Treaty PA_PCT_world (PA_PCT)</b>	Patents applied under Patent Cooperation Treaty worldwide
<b>European Patent Office in individual country PA_EPO_country (PA_EPO_ind)</b>	Patents applied under European Patent Office in individual country
<b>Patent Cooperation Treaty in individual country PA_PCT_country (PA_PCT_ind)</b>	Patents applied under Patent Cooperation Treaty in individual country
<b>PA_EPO1</b>	Time lag of Patents applied under European Patent Office worldwide
<b>PA_PCT1</b>	Time lag of Patents applied under Patent Cooperation Treaty worldwide
<b>PA_EPO_country1 (PA_EPO_ind1)</b>	Time lag of Patents applied under European Patent Office in individual country
<b>PA_PCT_country1 (PA_PCT_ind1)</b>	Time lag of Patents applied under Patent Cooperation Treaty in individual country
<b>PA_EPO_world2 (PA_EPO2)</b>	Patents applied under European Patent Office SQR
<b>PA_PCT_world2 (PA_PCT2)</b>	Patents applied under Patent Cooperation Treaty SQR
<b>PA_EPO_country2 (PA_EPO_ind2)</b>	Patents applied under European Patent Office on Country level SQR
<b>PA_PCT_country2 (PA_PCT_ind2)</b>	Patents applied under Patent Cooperation Treaty on Country level SQR

*Other variables,  
GEM, OECD*

<b>TEA</b>	= Total early-stage entrepreneurial activity
<b>GDP</b>	= Billion US dollars, current prices and PPPs, real GDP
<b>Harmonized unemployment rate (unempl)</b>	= comparable unemployment rates
<b>Unempl1-Unempl6</b>	= time lags of unemployment rate

Notes:

For the better understanding of the reader, variables mentioned in the paper are called differently from what they are called in the dataset.

Names in the brackets is how variables are named in the dataset- ( )

Table 2. Pair Wise Correlation table for all variables used in the regressions

	suboanw	unempl	PA_EPO _country	PA_PCT _country	PA_EPO _world	PA_PCT _world	PA_EPO _country2	PA_PCT _world2	PA_EPO _country2	PA_PCT _country2	GDP
suboanw	1										
unempl	-0.1641	1									
PA_EPO _country	0.4362*	-0.1593	1								
PA_PCT _country	0.5220*	-0.1908	0.9778*	1							
PA_EPO _world	0.0396	-0.0190	0.0208	0.0535	1						
PA_PCT _world	0.7326	0.8695	0.8575	0.6437	0.1339	1					
PA_EPO _country2	0.0285	0.0843	0.0091	-0.0052	0.2457	0.1401	1				
PA_PCT _world2	0.8057	0.4662	0.9374	0.9642	0.9999*	0.2243	0.1401	1			
PA_EPO _country2	0.0397	-0.0184	0.0210	0.0533	0.0294	0.0166	0.0297	0.0297	1		
PA_PCT _world2	0.7318	0.8740	0.8564	0.6454	0.7999	0.8857	0.7976	0.7976	0.0128	1	
PA_EPO _country2	0.5852*	-0.1663	0.9675*	0.9752*	0.0294	0.0166	0.0297	0.0297	0.0128	0.0128	1
PA_PCT _world2	0.0000	0.1484	0.0000	0.0000	0.7999	0.8857	0.7976	0.7976	0.9123	0.9123	0.0094
PA_EPO _country2	0.0344	0.0540	0.0077	0.0080	0.2666*	0.9771*	0.2718*	0.2718*	0.9611*	0.9611*	0.0094
PA_PCT _country2	0.7668	0.6409	0.9467	0.9452	0.0191	0.0000	0.0168	0.0168	0.9352	0.9352	0.9352
PA_EPO _country2	0.6452*	-0.1715	0.9081*	0.9678*	0.0709	-0.0063	0.0707	0.0707	0.0000	0.0000	0.0000
PA_PCT _country2	0.0000	0.1358	0.0000	0.0000	0.5402	0.9565	0.5413	0.5413	0.0000	0.0000	0.0000
GDP	0.6028*	-0.1035	0.9318*	0.9701*	0.0215	-0.0095	0.0211	0.0211	0.9533*	0.9533*	0.9728*
	0.0000	0.3703	0.0000	0.0000	0.8527	0.9348	0.8556	0.8556	0.0000	0.9809	0.0000



### Table 3 Multicollinearity Test - VIF (variance inflation factor)

Table 3.1 Multicollinearity Test - for Patents applied under European Patent Office in individual country's

Variable	VIF	SQRT VIF	Tolerance	R-squared
PA_EPO_country	7.82	2.80	0.1278	0.8722
GDP	7.71	2.78	0.1298	0.8702
unempl	1.04	1.02	0.9593	0.0407
Mean VIF	5.52			

Table 3.2 Multicollinearity Test - for Patents applied under Patent Cooperation Treaty in individual country's

Variable	VIF	SQRT VIF	Tolerance	R-squared
PA_PCT_country	19.71	4.44	0.0507	0.9493
GDP	19.20	4.38	0.0521	0.9479
unempl	1.18	1.08	0.8507	0.1493
Mean VIF	5.52			

Table 3.3 Multicollinearity Test - for Patents applied under European Patent Office worldwide

Variable	VIF	SQRT VIF	Tolerance	R-squared
PA_EPO_world	1.00	1.00	0.9993	0.0007
GDP	1.01	1.01	0.9889	0.0111
unempl	1.01	1.01	0.9890	0.0110
Mean VIF	1.01			

Table 3.4 Multicollinearity Test - for Patents applied under Patent Cooperation Treaty worldwide

Variable	VIF	SQRT VIF	Tolerance	R-squared
PA_PCT_world	1.01	1.00	0.9929	0.0071
GDP	1.01	1.01	0.9893	0.0107
unempl	1.02	1.01	0.9823	0.0177
Mean VIF	1.01			

## Table 4 - Checking Normality of Residuals

Table 4.1 Graph of a standardized normal probability (P-P)

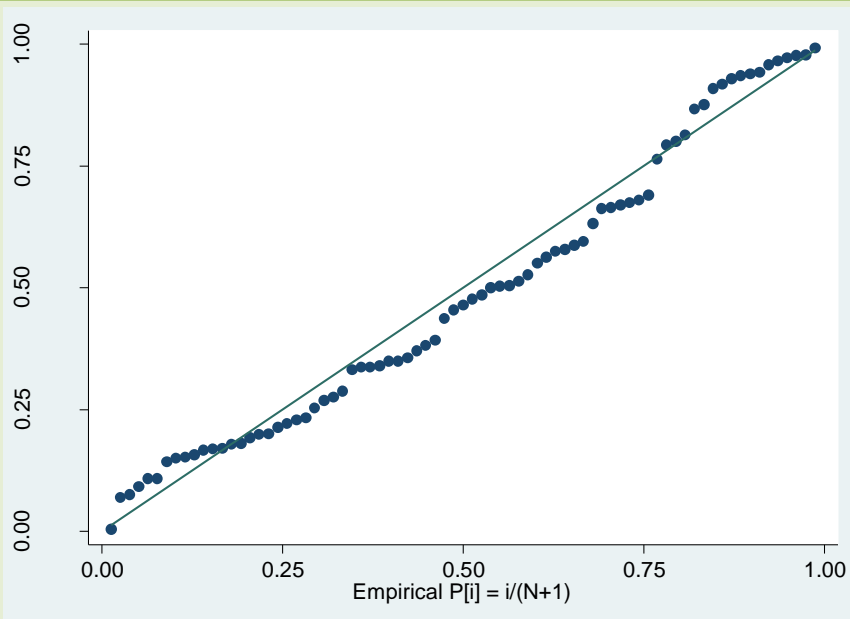


Table 4.2 Plot of the quantiles of a variable against the quantiles of a normal distribution

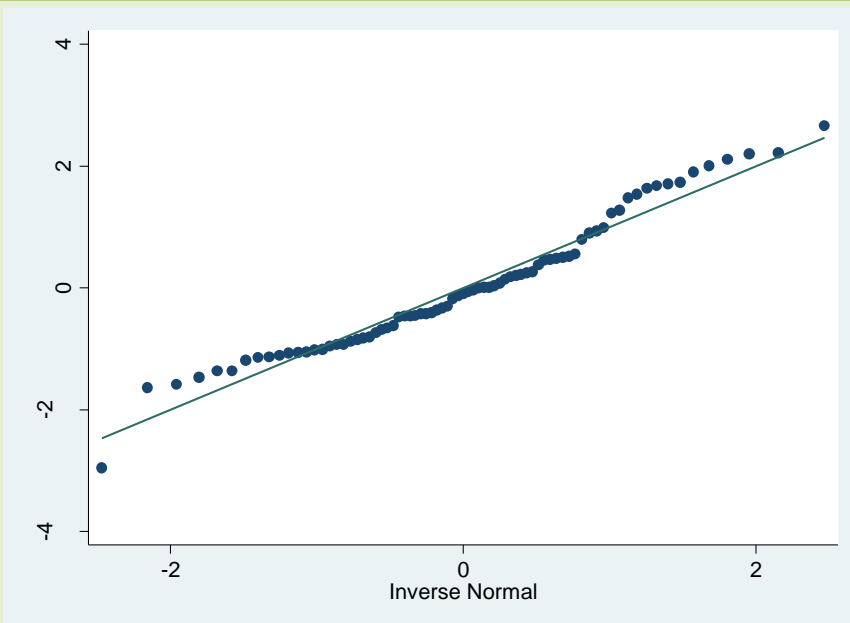


Table 4.3 Shapiro-Wilk W test - test for normality

Variable	Obs	W	V	z	Prob>z
Residuals	77	0.96641	2.234	1.758	<b>0.03939</b>

**Table 5.1 FE & RE, Hausman Test regression results of nascent entrepreneurial activity across 11 countries on Country level dataset, 2002-2008**

<i>Step 1: Regressions on Country level dataset (N=77) (11 countries) (with OUT control variables)</i>				
	Fixed effects	Random effects	Fixed effects	Random effects
PA_EPO_country	.380 (.15) [.000]	.166 (0.04**) [.000]		
PA_PCT_country			.155 (.28) [.000]	.151 (.01**) [.000]
Constant	2.168( 0.01**) [.826]	2.815(.00**) [.527]	2.799 (.0**) [.544]	2.813 (.0**) [.486]
<b>Model diagnostics</b>				
Hausman Test (Prob> Chi2)	.993		1.000	
R <sup>2</sup>	.053	.193	.037	.2755
Rho	.821	.725	.680	.6900
<i>Step 2: Regressions on Country level dataset (N=77) (11 countries) (WITH control variables)</i>				
PA_EPO_country	.165 (0.6) [.000]	-.044 (.68) [.000]		
PA_PCT_country			.289 (.08*) [.000]	.075 (.60) [.000]
GDP	-.177 ( 0.37)[.000]	.181 (.05*)[.000]	-.398 ( .03) [.000]	.070 (.62)[.000]
Unemploymentrate	-.404 (.13) [ .145]	-.252 (0.16) [.097]	-.314 (.23) [.143]	-.229 (.24 ) [ .106]
Constant	4.756 ( .01**) [1.845]	3.566 (.0**) [ .702]	4.840 (.0**) [ 1.105]	3.522 (.0**) [ .746 ]
<b>Model diagnostics</b>				
Hausman Test (Prob> Chi2)	.0008		0.000	
R <sup>2</sup>	.098	.408	.140	.321
Rho	.832	.430	.880	.504

Explanatory notes:

Instead of beta's, elasticity is displayed - d(lny)/d(lnx)

Significance level in first parentheses. = ()

Standard errors in second parentheses. = []

All models include time dummies.

\*\* significant at > 95% confidence

\* significant at > 90% confidence

The R<sup>2</sup> of the random effects model is the overall R<sup>2</sup> and the R<sup>2</sup> of the fixed effects models is the within R<sup>2</sup>

**Table 5.2 FE & RE, Hausman Test regression results of nascent entrepreneurial activity across 11 countries on World level dataset , 2002-2008**

<i>Step 1: Regressions on World level dataset (N=77) (11 countries) (with OUT control variables)</i>				
	Fixed effects	Random effects	Fixed effects	Random effects
<b>PA_EPO_world</b>	.556 (.63) [.000]	-.468 ( .83) [.000]		
<b>PA_PCT_world</b>			.019 ( .83) [.000]	.019 ( .83) [.000]
<b>Constant</b>	1.547 (.69)[3.891]	4.815 (.50) [7.293]	3.238 (.0**) [ .277]	3.238 (.0**) ( .513)
<b>Model diagnostics</b>				
<b>Hausman Test (Prob&gt; Chi2)</b>	1.0000		Does not fit Hausman test	
<b>R<sup>2</sup></b>	.018	.004	.018	.004
<b>Rho</b>	.763	.754	.763	.754
<i>Step 2: Regressions on World level dataset (N=77) (11 countries) (WITH control variables)</i>				
<b>PA_EPO_world</b>	.560 (0.6) [.000]	-.795 (.72) [.000]		
<b>PA_PCT_world</b>			.046 (.61) [.000]	.033 (.72) [.000]
<b>GDP</b>	-.235 ( 0.37)[.000]	.115 (.06*)[.000]	-.235 (.15)[.000]	.115 (.06*)[.000]
<b>Unemploymentrate</b>	-.428 (.13) [ .145]	-.311 (0.12) [.111]	-.428 (.10) [.141]	-.311 (.12) [.111]
<b>Constant</b>	3.764 ( .01**) [ 4.064 ]	6.399 (.39) [ 7.482 ]	5.356 (.0**) [ 1.002]	3.722 (.0**) [.792]
<b>Model diagnostics</b>				
<b>Hausman Test (Prob&gt; Chi2)</b>	.365		.525	
<b>R<sup>2</sup></b>	.094	.339	.094	.339
<b>Rho</b>	.911	.667	.911	.667

Explanatory notes:

Instead of beta's, elasticity is displayed -  $d(\ln y)/d(\ln x)$

Significance level in first parentheses. = ( )

Standard errors in second parentheses. = [ ]

All models include time dummies.

\*\* significant at > 95% confidence

\* significant at > 90% confidence

The R<sup>2</sup> of the random effects model is the overall R<sup>2</sup> and the R<sup>2</sup> of the fixed effects models is the within R<sup>2</sup>

**Table 5.3 Non-Linear(SQR of the independent variable is included in the model) FE & RE, Hausman Test regression results of nascent entrepreneurial activity across 11 countries on Country level dataset, 2002-2008**

<i>Step 1: Regressions on Country level dataset (N=77) (11 countries)</i>		<i>(with OUT control variables)</i>		
	Fixed effects	Random effects	Fixed effects	Random effects
<b>PA_EPO_country</b>	-2.134 (.06*) [.000]	-.582 (.04**) [.179]		
<b>PA_PCT_country</b>			.541(.15) [.377]	-.217 (.11)[.135]
<b>SQR_independ_var</b>	.914 ( .02**) [.000]	.489 (.0**) [.116]	-.172 (.26)[.155]	.247 (.0**) [.088]
<b>Constant</b>	7.424 ( .0**) [2.474]	3.820(.0**) [.422]	2.046(.02**) [.870]	3.277(.0**) [ .392]
<b>Model diagnostics</b>				
<b>Hausman Test (Prob&gt; Chi2)</b>	.6545		Does not fit Hausman test	
<b>R<sup>2</sup></b>	.128	.614	.057	.554
<b>Rho</b>	.974	.468	.814	.344
<i>Step 2: Regressions on Country level dataset (N=77) (11 countries)</i>		<i>(WITH control variables)</i>		
<b>PA_EPO_country</b>	-2.203 (.05*) [ 1.146]	-.716 (.0**) [.189]		
<b>PA_PCT_country</b>			.388 (.29) [.373]	-.270(.1) [.167]
<b>SQR_independ_var</b>	.874 ( .03**) [.407]	.473 (.0**) [.112]	-.049 (.76)[.166]	.221 (.01**) [.093]
<b>GDP</b>	-.140 ( .46)[.192]	.148 (.05*) [.085]	-.372( .07*) [.205]	.087 (.53)[.140]
<b>Unemploymentrate</b>	-.386 (.13) [ .258]	-.193 (.23) [ .161]	-.320 (.23) [.267]	-.205(.24 ) [ .174]
<b>Constant</b>	9.478 ( .0**) [2.835]	4.315(.0**) [ .648]	4.584 (.0**) [1.40]	3.851(.0**) [ .676]
<b>Model diagnostics</b>				
<b>Hausman Test (Prob&gt; Chi2)</b>	.966		0.001	
<b>R<sup>2</sup></b>	.167	.658	.141	.576
<b>Rho</b>	.983	.439	.865	.362

Explanatory notes:

Instead of beta's, elasticity is displayed -  $d(\ln y)/d(\ln x)$

Significance level in first parentheses. = ( )

Standard errors in second parentheses. = [ ]

All models include time dummies.

\*\* significant at > 95% confidence

\* significant at > 90% confidence

The R<sup>2</sup> of the random effects model is the overall R<sup>2</sup> and the R<sup>2</sup> of the fixed effects models is the within R<sup>2</sup>

**Table 6.1 Robustness check, one year lags included. FE & RE, Hausman Test regression results of nascent entrepreneurial activity across 11 countries on Country level dataset, 2002-2008**

<i>Step 1: Regressions on Country level dataset (N=77) (11 countries) (with OUT control variables)</i>				
	Fixed effects	Random effects	Fixed effects	Random effects
PA_EPO_country	.223 (.55) [ .376]	.216 (.56) [ .375 ]		
PA_PCT_country			.541(.15) [.377]	.240 (.31)[ .241 ]
One Year Lag	.274 ( .49) [.562]	-.054 (.88)[ .379 ]	-.172 (.26)[.155]	-.089 ( .7)[ .234 ]
Constant	1.769 ( .18) [ 1.328]	2.833 (.0**) [ .523]	2.046(.02**) [.870]	2.866 (.0**) [ .404 ]
<b>Model diagnostics</b>				
Hausman Test (Prob> Chi2)	.993		Does not fit Hausman test	
R <sup>2</sup>	.050	.209	.044	.272
Rho	.865	.687	.658	.382
<i>Step 2: Regressions on Country level dataset (N=77) (11 countries) (WITH control variables)</i>				
PA_EPO_country	-.526 (.33) [ .548 ]	.136 (.15) [ .440 ]		
PA_PCT_country			.100 (.74) [ .305]	.195 (.83) [ .236 ]
One Year Lag	.702 (.26) [ .623 ]	-.826 (.10)[ .510 ]	.318 (.28)[ .298]	-.309 (.24)[ .266]
GDP	-.598 ( .07*) [.330 ]	.323 (.01**) [.130]	-.771 ( .02**) [ .335 ]	.267 (.15)[ .189 ]
Unemploymentrate	-.316 (.29) [ .300 ]	-.251 (.19) [ .192 ]	-.278 (.34) [ .295]	-.223 (.28 ) [ .208 ]
Constant	5.918 ( .01**) [ 2.346 ]	3.623 (.0**) [ .748 ]	4.584 (.0**) [1.40]	3.536 (.0**) [ .789 ]
<b>Model diagnostics</b>				
Hausman Test (Prob> Chi2)	.048		.016	
R <sup>2</sup>	.136	.529	.165	.434
Rho	.954	.420	.946	.472

Explanatory notes:

Instead of beta's, elasticity is displayed -  $d(\ln y)/d(\ln x)$

Significance level in first parentheses. = ( )

Standard errors in second parentheses. = [ ]

All models include time dummies.

\*\* significant at > 95% confidence

\* significant at > 90% confidence

The R<sup>2</sup> of the random effects model is the overall R<sup>2</sup> and the R<sup>2</sup> of the fixed effects models is the within R<sup>2</sup>

**Table 6.2 Robustness check, one year lags included. FE & RE, Hausman Test regression results of nascent entrepreneurial activity across 11 countries on World level dataset, 2002-2008**

<i>Step 1: Regressions on World level dataset (N=77) (11 countries)</i>		<i>(with OUT control variables)</i>		
	<b>Fixed effects</b>	<b>Random effects</b>	<b>Fixed effects</b>	<b>Random effects</b>
<b>PA_EPO_world</b>	.233 (.89) [ 1.801 ]	3.173 (.73) [ 9.214 ]		
<b>PA_PCT_world</b>			.033 (.73) [.096 ]	.033 (.73)[ .096 ]
<b>One Year Lag</b>	-.060 ( .96) [1.244 ]	5.406 (.74)[ 16.279 ]	.012 (.13)[.092 ]	.012 ( .89)[ .092 ]
<b>Constant</b>	2.779 ( .69) [ 4.057 ]	11.357 (.44) [ 26.085 ]	3.175 (.0**) [ .524 ]	3.175 (.0**) [ .672 ]
<b>Model diagnostics</b>				
<b>Hausman Test (Prob&gt; Chi2)</b>	1.000		Does not fit Hausman test	
<b>R<sup>2</sup></b>	.013	.003	.013	.003
<b>Rho</b>	.742	.730	.742	.730
<i>Step 2: Regressions on World level dataset (N=77) (11 countries)</i>		<i>(WITH control variables)</i>		
<b>PA_EPO_world</b>	-.406 (.82) [ 1.857 ]	5.138 (.55) [ 9.420 ]		
<b>PA_PCT_world</b>			.034 (.36) [ .095 ]	.054 (.55) [ .098 ]
<b>One Year Lag</b>	.486 ( .73)[ 1.444 ]	-8.695 (.6)[ 16.6 ]	-.011 (.9)[ .090 ]	-.022 (.23)[ .093]
<b>GDP</b>	-.391 ( .06*)[ .214 ]	.129 (.03**)[ .063 ]	-.391 ( .06*) [ .214 ]	.129 (.03**)[ .063 ]
<b>Unemploymentrate</b>	-.358 (.22) [ .292 ]	-.252 (.24) [ .217 ]	-.358 (.22) [ .292 ]	-.252 (.24 ) [ .217 ]
<b>Constant</b>	5.571 ( .19) [ 4.239 ]	16.461 (.53) [26.748]	5.771 (.0**) [1.254 ]	3.402 (.0**) [ .932 ]
<b>Model diagnostics</b>				
<b>Hausman Test (Prob&gt; Chi2)</b>	.9973		.193	
<b>R<sup>2</sup></b>	.111	.362	.111	.626
<b>Rho</b>	.945	.626	.945	.626

Explanatory notes:

Instead of beta's, elasticity is displayed -  $d(\ln y)/d(\ln x)$

Significance level in first parentheses. = ( )

Standard errors in second parentheses. = [ ]

All models include time dummies.

\*\* significant at > 95% confidence

\* significant at > 90% confidence

The R<sup>2</sup> of the random effects model is the overall R<sup>2</sup> and the R<sup>2</sup> of the fixed effects models is the within R<sup>2</sup>