National differences in entrepreneurial activity
an institutional perspective

Zegel, S.J.

Award date:
2011

Link to publication
National differences in entrepreneurial activity: an institutional perspective

By: S.J. (Stijn) Zegel

BSc Innovation Sciences, TU/e 2010
Student identity number: 0614368

in partial fulfilment of the requirements for the degree of:

Master of Science
in Innovation Sciences

Supervisors:

Prof. Dr. K. Frenken
Dr. Z.O. Nomaler

Faculty of Innovation Sciences, TU/e
TU/e, Department of Industrial Engineering and Innovation Sciences

Subject headings: Entrepreneurship – Entrepreneurs – Entrepreneurship; developing countries – Institutions – Capitalism – Capitalism; history – Innovation: master theses
Summary

This study investigates whether there is a relation between the differences in institutional arrangements and the differences in entrepreneurial activity across countries. Entrepreneurship is a field that is very closely linked to the field of innovation studies since the entrepreneur is the actor driving the process of innovation. Schumpeter (1942) stated that the entrepreneur is the key person who converts inventions into successful innovations. Recently, policymakers have started to get increasingly more interested in the field of entrepreneurship as they see the entrepreneur as a driving factor for the economy. The field of innovation has had this attention earlier as the relation between innovation and economic development is supposedly very strong. Schumpeter (1942) describes this relation as “creative destruction”, which essentially explains that in order to stimulate development, it is necessary to break with the old used methods and technologies. Entrepreneurs are therefore the driver of this process. Institutional theory has proven to be of value in the explanation of innovation. Since innovation and entrepreneurship are so closely linked, it could well be the case that institutional theory will have the same explanatory power for entrepreneurship.

The theory of the Varieties of Capitalism (Hall & Soskice, 2001) uses institutions, important to firms, to identify three country groups. These three are called: Coordinated market economies (CMEs), Liberal market economies (LMEs) and Mediterranean market economies (MMEs). The contexts for firms in CMEs are dominated by collaboration. For LMEs they are dominated by competition, and for MMEs it is a combination of both.

The notion of institutions is well defined by North (1994):

“... the humanly devised constraints that structure human interaction. They are made up of formal constraints (e.g., rules, laws, constitutions), informal constraints (e.g., norms of behavior, conventions, self-imposed codes of conduct), and their enforcement characteristics. Together they define the incentive structure of societies and specifically economies.”

This means that in order to find the relation between institutions and entrepreneurship the relevant formal and informal constraints need to be identified. The relevant institutions for entrepreneurship are based on five pillars, each constructed from variables from their field. The pillars are: Education and skills pillar, Economy pillar, Culture pillar, Government pillar and Financial pillar.

In this study two analyses will be performed. In the first clustering analyses will be used to group countries for each of the pillars of entrepreneurship. These groups can then be matched with the groups from the theory of the Varieties of Capitalism to see whether these are also present for entrepreneurship. The second analysis will use univariate regressions and clustering analyses to find the relationships between the institutions and entrepreneurship. Because the entrepreneurship data is very hard to obtain, the sample size is limited. This is the main reason why this study cannot use multivariate regression analyses.

The data used in conducting these two analyses is acquired from the Global Entrepreneurship Monitor (GEM) and various other worldwide databases like the World Bank and the World Economic Forum. The GEM provides multiple types of entrepreneurship data, but the data used in this study is the Total Entrepreneurial activity (TEA). The TEA is roughly based on the percentage of people between the ages 15 and 64 who started a business in the past 42 months. However, this data only consists of entrepreneurs that start with an opportunity based intention. This excludes people who
start a firm because they do not see another option. The categories are explained in more detail in the thesis. The entrepreneurship data covers 46 countries.

From this study there are three conclusions:

1) The Varieties of Capitalism theory is not applicable on entrepreneurship.
2) The economic development has a heavy negative impact on entrepreneurship for weak economies. The reason for this is sought in economic growth, based on economies of scale and efficiencies.
3) For strong economies no effects of institutions on entrepreneurship were found. The reason for this is sought in other types of economic growth, based on knowledge, innovation and capital.

This study is subject to a few limitations regarding the used data, which can have a strong impact. The first is based on the sample size, which is not sufficient for multivariate regression analyses. This even becomes a bigger problem now that economic development has shown to be an indicator. Comparisons in future research can thus only be done between countries of the same economic level. The second data issue is based on inaccuracy, which plays a big role in research comparing countries, since indicators are often extrapolated and measured in a ‘soft’ way. As stated, future research should only focus on countries with the same economic development level. Furthermore, a focus on more local factors is recommended.
Table of Contents

1. Introduction ......................................................................................................................................... 9
2. Literature overview ........................................................................................................................... 10
   2.1 Varieties of Capitalism ................................................................................................................. 10
   2.2 Models of entrepreneurship ....................................................................................................... 13
   2.3 Comparative institutional analysis approach .............................................................................. 17
3. Data ................................................................................................................................................... 28
   3.1 Data collection ............................................................................................................................. 29
   3.2 Variable structuring ..................................................................................................................... 35
4. Analyses ............................................................................................................................................. 37
   4.1 Descriptive statistics .................................................................................................................... 37
   4.2 Clustering ..................................................................................................................................... 39
   4.3 Comparing clustering results with the VoC ................................................................................. 42
   4.4 Entrepreneurship analyses .......................................................................................................... 44
     4.4.1 Assumptions ......................................................................................................................... 45
     4.4.2 Univariate regression analyses ............................................................................................. 45
     4.4.3 Univariate regression analysis: ‘strong’ and ‘weak’ economies ........................................... 49
     4.4.4 Regression with Hierarchical Clustering (HC) dummies ....................................................... 55
     4.4.5 Final remarks on the analyses .............................................................................................. 59
5. Conclusion and discussion ................................................................................................................. 61
6. Bibliography ....................................................................................................................................... 64
7. Appendixes ........................................................................................................................................ 69
1. Introduction
Comparative research has gained a lot of interest as country leaders are interested in the performance of their countries vis-à-vis competitor countries. When speaking about the competition it is obvious that there are a lot of components within economies on which one can make a comparison. Currently, country leaders seem to have a great interest in stimulating entrepreneurship. Although the exact relation between entrepreneurship and economic growth is still heavily debated, the general assumption holds that entrepreneurship is closely linked to all kinds of processes that lead to economic growth.

An important process linked to entrepreneurship is innovation as this particular process is a key factor to the progress of the economy. It stands for all new ways of economic development. Schumpeter (Schumpeter, 1942) wrote that the entrepreneur is the person that can convert a new idea or invention into a successful innovation. The entrepreneur is the agent driving ‘creative destruction’, as Schumpeter calls it, which means the destruction of old technologies or processes by introducing challenging new ones. Central in this process is the transformation of technological knowledge to specific economic needs, which requires active involvement of entrepreneurs (Acs, Audretsch, Braunerhjelm, & Carlsson, 2004).

Although comparative research and entrepreneurship are both topics that are very central to all kinds of policies or researches there are very few studies on the differences between countries in terms of entrepreneurship. In order to find the appropriate way to investigate the possible connection between country differences and entrepreneurship, various disciplines need to be consulted. The most obvious field is the field of innovation studies as entrepreneurship is so closely linked to innovation. However, innovation studies are far more extensive than entrepreneurship studies, as the interest in explaining national differences in innovation goes back at least to the late 1980s with the introduction of the concept of national innovation system (for an overview, see Edquist 1997). The central tenet underlying these studies holds that differences in national institutions are responsible for differences in the rate and direction of innovation in a country. Similarly, in this study the connection of the institutional arrangements and entrepreneurship will be investigated. In particular, we will follow the more recent literature on Varieties of Capitalism (Hall and Soskice, 2001) to explain differences in entrepreneurship between countries. The following two research questions are central in this study:

Research question 1:

Do varieties of capitalism exist in the context of entrepreneurship?

Research question 2:

Do differences in institutions explain differences in the rate of opportunity-driven entrepreneurship across countries?

The organization of this study will be as follows. In the first chapter there will be an extensive overview of all kinds of literature, first explaining institutional analyses followed by an in depth view of what entrepreneurship is all about. The second chapter will present the exact research questions and explain the research design. In the next chapter the data collection and structuring will be discussed. The fourth chapter will cover the analysis and results followed by the final chapter consisting of conclusions and discussions.
2. Literature overview

2.1 Varieties of Capitalism

Institutions are used and defined in numerous theoretical frameworks. Probably the best known theory concerns the theory of Varieties of Capitalism (VoC). But the notion of institutions is also used in influential theories like ‘national innovation systems’ or ‘business systems’. In these works there are many definitions that often only differ in minor parts, but they all consist of elements that are intangible or sometimes even a bit fuzzy. Edquist and Johnson (1997) use terms like common habits, norms, routines, established practices, rules and laws, but one could question the exact meaning of these terms as it is unclear when and for whom (and between whom) these are valid. In the work of Fagerberg, Mowery and Nelson (2005) these institutions influence the relations between different actors like for instance firms and universities. So maybe these are meant to be a bit fuzzy as they need to explain the interaction rules for a wide range of different actors, in their article they explain that the situation can be different between countries for the same institution. The authors give an example in which the most important R&D performer in one country is company-based while in another it is university-based. Wennekers, Uhlaner and Thurik (2002) use a definition of institutions by North (1994), which makes a good distinction between formal and informal institutions. In his work institutions are:

“... the humanly devised constraints that structure human interaction. They are made up of formal constraints (e.g., rules, laws, constitutions), informal constraints (e.g., norms of behavior, conventions, self-imposed codes of conduct), and their enforcement characteristics. Together they define the incentive structure of societies and specifically economies.”

The two main components in institutional theories are therefore organizations and the institutions. As we have seen in the example by Fagerberg, Mowery and Nelson (2005) it is nearly impossible to create a definition of the specific organizations that deal with specific institutions as they may/will differ between countries. In the following paragraphs the theory of the varieties of capitalism is explained and discussed. This theory provides a good understanding of how to create a framework in which the connection between institutions and entrepreneurship can be tested.

The theory of the variety of capitalism introduced by Hall & Soskice (2001) makes the distinction between two types of economies on a continuum of varieties between them, the two ‘poles’ are liberal market economies and coordinated market economies (Akkermans, Castaldi, & Los, 2009). Liberal market economies (LMEs) dissolve problems purely through market mechanisms, which are based on competition, while coordinated market economies (CMEs) dissolve issues through collaboration. Their theory emphasizes the role of the differences between countries, which is another type of analysis than to try to define which of multiple economies is ‘the best’. This however does not mean that one certain type of economy can’t be better suited to reach a certain goal. Hall & Soskice (2001) describe that although different types of firm relationships vary systematically across nations, “firms gravitate towards the mode of coordination for which there is institutional support”. So this means that the firm relations in a country are based on the institutional arrangements within that country. Institutional theories emphasize that institutions may both constrain and enable the action choices of agents (Williamson, 1998) (DiMagio & Powell, 1983), therefore institutions can have an unintentional effect on entrepreneurship. When discussing which country has the strongest economic position the discussion is usually based on a specialization in the production of goods by using the available land, labor, capital etc. most effectively. The theory of the
varieties of capitalism, based on comparative institutional advantage theory, tries to explain why that country ‘chose’ to specialize in the production they have today. This reflects both market and non-market relations which may be just as important. Differences over time seem to be dominated by economic influences, whereas cross-country variations seem to be affected by cultural and institutional components (Wennekers, Uhlner, & Thurik, 2002).

Although the VoC approach shows the importance of formal institutions, formal institutions also play a role (Hall & Soskice, 2001). The way countries work nowadays result from a development through history, which results in formal institutions but also in informal rules and understandings. In fact this is closely linked to the development of a national culture.

In the theory of Hall & Soskice (2001) the firm is the unit of analysis, they use the definition of Teece and Pisano (1997) which states that “firms are actors seeking to develop and exploit core competencies or dynamic capabilities understood as capabilities for developing, producing and distributing goods and services profitably”. Keep in mind that this definition is quite different from the earlier given definition of the entrepreneur. This will have its implications as established firms are subject to very different dynamics than entrepreneurs. Hall & Soskice (2001) define five areas in which firms need to develop relationships, described in table 1.

**Table 1: Necessary network for firms described by Hall & Soskice (2001)**

<table>
<thead>
<tr>
<th>Five areas of interest</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industrial relations</td>
<td>Wage and productivity</td>
</tr>
<tr>
<td>Vocational training and education</td>
<td>Workforce with suitable skills</td>
</tr>
<tr>
<td>Corporate governance</td>
<td>‘Availability of’ &amp; ‘Ability to secure’ funds</td>
</tr>
<tr>
<td>Inter-firm relations</td>
<td>Stable demand, supplies &amp; access to technology</td>
</tr>
<tr>
<td>Own employees</td>
<td>Competences &amp; cooperation</td>
</tr>
</tbody>
</table>

In liberal market economies (LMEs) and coordinated market economies (CMEs) firms use different strategies to cope with these areas of interest. In LMEs relationships firms use to arrange these areas of interest are heavily subjected to market mechanisms. Therefore these relations are primarily based on the exchange of goods or services on a distance. Hall and Soskice (2001) speak about this exchanges being in the context of “competition and formal contracting”. What they mean to say is that the firms do not work on a coordinated plan like firms in CMEs where for instance the one firm clearly has the role of being the supplier and the other the buyer and as soon as contracting stops this relationship also stops. For firms choices will therefore be based on short-term supply and demand conditions of the market, while long-term considerations of stability are of lesser importance.

In CMEs on the other hand the relationships firms use to arrange the areas of interest are more subject to non-market related forces. For CMEs these forces are based on coordination and therefore these relations are heavily subjected to actor-networks. Hall and Soskice (2001) speak about the exchange of goods or services as being in the context of “extensive relational or incomplete contracting”. In CMEs the relation is easily just as important as the product or service the buyer tries to obtain, due to the exchange of private information inside networks and connections to important collaborations. For firms from CMEs will make decisions on strategic interactions amongst other firms.
For firms from both types of market economies market relations will be important. So although firms from CMEs use their market relations to create strategies with them this does not necessarily mean that firms from LMEs do not end up in the same deal. How the deal is made will however be completely different as for CMEs they will be based on collaboration and interactions while for LMEs they will be based on formal contracts and highly competitive markets. The same can be said about the long-term performance of firms. Hall and Soskice (2001) give an example about how firms would react on a sudden change in cost of the production of a certain product. The unexpected higher costs were passed along to the customer by the LME firms to maintain the same profitability while CME firms kept the same price to hold the market share.

Other examples of differences between the two market economies would be that CMEs use more specific and co-specific assets while LMEs use more switchable assets as CMEs optimize through coordination and LMEs optimize through adaptability. In addition, firms from CMEs have technology transfer through inter-company relations while LMEs have technology transfer through movement of scientific and engineering personnel across companies.

The third group the theory of VoC points at is the group of Mediterranean market economies (MMEs). These do not fit to the LMEs or CMEs because of their heavy state intervention over a longer period of time (Hall & Soskice, 2001). This long term influencing role that the state has had over the economy changed the business interactions and labor relations significantly. Although the influence of the state might have decreased a lot during the years the characteristics of these countries have not changed significantly (Schmidt, 2007). The aim of state intervention has always been to enhance business. However, the effect of state policies in practice may have been enhancing or hindering firms’ interactions and productive capabilities (Schmidt, 2007). How the firms in MME deal with the five issues as discussed for the firms from LMEs and CMEs is not entirely clear. However, Hall and Soskice (2001) argue that due to the heavy state influence the MMEs do tend to cluster into a single group. What they do say is that they counties are, besides extensive state intervention, also characterized by a large agrarian sector. They also tend to use non-market coordination for corporate finance and liberal arrangements for labor relations. So basically they tend to float in the middle of LMEs and CMEs.

The countries that are labeled LMEs, CMEs or MMEs by Hall and Soskice are presented in table 2.

<table>
<thead>
<tr>
<th>Liberal market economies</th>
<th>Coordinated market economies</th>
<th>Mediterranean market economies</th>
</tr>
</thead>
<tbody>
<tr>
<td>o Australia</td>
<td>o Austria</td>
<td>o France</td>
</tr>
<tr>
<td>o Canada</td>
<td>o Belgium</td>
<td>o Italy</td>
</tr>
<tr>
<td>o Ireland</td>
<td>o Iceland</td>
<td>o Spain</td>
</tr>
<tr>
<td></td>
<td>o Switzerland</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Germany</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Japan</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>o Portugal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>o Greece</td>
</tr>
<tr>
<td></td>
<td></td>
<td>o Turkey</td>
</tr>
<tr>
<td>o New Zealand</td>
<td>o Netherlands</td>
<td></td>
</tr>
<tr>
<td>o United Kingdom</td>
<td>o Norway</td>
<td></td>
</tr>
<tr>
<td>o United States of America</td>
<td>o Sweden</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Finland</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Denmark</td>
<td></td>
</tr>
</tbody>
</table>
For entrepreneurs these institutional factors will be very different as for instance funds have to be gathered from third parties, which can cause serious friction for the entrepreneur as most entrepreneurs do not have the previous discussed reputation to gain trust among investors. In the next chapters entrepreneurship will be examined in depth in order to get a better insight in the determinants for entrepreneurship and how the difference between countries affects these.

2.2 Models of entrepreneurship

One of the central questions in entrepreneurship research concerns the drivers for entrepreneurs to engage in entrepreneurial activity. Entrepreneurs are not servants of the economy in the way that they start a business in order to achieve innovation, job creation or economic growth (Hessels, Gelderen, & Thurik, 2008). These are motives for governments to support entrepreneurship. The entrepreneur, however, seeks personal benefits like profits, autonomy, and respect amongst others or is forced into entrepreneurship due to lack of better options. The motives of the entrepreneur can be split in two groups, necessity-driven and opportunity-driven entrepreneurs. The necessity-driven entrepreneurs are more or less pushed into entrepreneurship as there are no alternatives (often because of unemployment), while the opportunity-driven entrepreneur is pulled by motives like independence, wealth, challenge, recognition and status (Wilson, Marlino, & Kickul, 2004) (Gelderen & Jansen, 2006) (Cassar, 2007). For this research the motive must be opportunity-driven as these are the growth oriented entrepreneurs who contribute most to economic growth and productivity of a country.

Clearly the entrepreneur must first discover an opportunity before he or she can start any type of new firm, whether that opportunity is obtained though some scientific expertise, serendipity, good network abilities, imitation or some other process. Baker, Gedajlovic & Lubatkin (2005) state that once an opportunity arises it must be evaluated by the entrepreneur to see if the opportunity will deliver appropriate returns compared to its opportunity costs. This however does not always take place in a strictly rational way as the perception of certain barriers can be enough to withhold an entrepreneur from engaging in the opportunity. After this evaluation the entrepreneur must look at the exploitation possibilities and see if there are enough resources in reach to succeed at the risky undertaking. These resources can vary from personal abilities to availability of sufficient capital. In figure 1 the process is shown graphically.

![Schematic entrepreneurial process](image-url)
Although this model shows the stages of the entrepreneur very clearly it is far from the truth, the human mind is not able to think about evaluation without thinking about parts of the exploitation. So this indicates that the process after discovery is messier and more complicated than explained in the model above.

In the next model presented in a paper by Gnyawali & Fogel (1994), seen in figure 2 on the left, the ability to and the propensity to enterprise are split in two. Hereby we can make a difference between variables that have an effect purely on the question whether the entrepreneur ‘can’ exploit this opportunity and variables that influence whether the entrepreneur ‘wants’ to. This combined represents the likelihood that the entrepreneur will indeed exploit the opportunity.

Figure 2: Split between ability to and propensity to enterprise. (Gnyawali & Fogel, 1994)

The next step in their model, seen in figure 2 on the right, shows which types of variables influence the different parts of the model. These are the following four categories:

- **Entrepreneurial & business skills** influence the ability to enterprise by the types of education the entrepreneur has received and by the proportion of information he/she can access.
- **Socio-economic factors** which influence propensity to enterprise can consist of the public attitude towards entrepreneurship, successful role models, the extent of economic growth and the proportion of small firms in the country.
- **Government policies and procedures** directly influence the opportunity by variables, which are based on proprietary rights, number of involved institutions (bureaucracy) and required licenses.
- **Financial and Non-financial assistance** are based on all help an entrepreneur gets from outside his/her firm, all things that help with problems that cannot be solved within the firm. The whole capital availability is presented by these variables, like venture capital, loans or subsidies. Non-financial assistance can be represented by good business networks, information sources or government procurement programs for start-ups.

Although not mentioned in the work by Gnyawali & Fogel (1994), others press the importance of experience and talent which fit very well with ‘Ability to enterprise’ but these and other variables not mentioned in this model will be discussed in more detail at a later stage.

When talking about the determinants of entrepreneurship the general consensus amongst authors is that there are two types of determinants, namely personal variables and non-personal variables
These do not fully match with the four influencing categories we have seen in the previous section, as personal variables contain both variables that influence the ability to enterprise and the propensity to enterprise (for instance, an entrepreneur’s experience and confidence). Non-personal variables, however, do not influence the ability to enterprise, but they do influence the propensity to enterprise (for instance the business climate).

In this research the scope it to look at the connection between the difference in institutional arrangement and the difference in entrepreneurial activity. This means that personal variables will be hard to incorporate as these are very micro oriented variables like ‘confidence’ or ‘age’. Although some of these will return in terms of cultural differences between countries. These cultural variables represent the entire population of a country and not just the single entrepreneur. In the following section, the four aforementioned categories are discussed in more detail.

2.2.1 Entrepreneurial, business & technical skills
In an article by Chrisman, Bauerschmidt & Hofer (1999) an extensive overview is shown of personal variables, among which skills, experience and education. Skills, experience and education are the direct link to the ability to enterprise as these reflect ones capabilities and talents for entrepreneurship. When the possessed skills are of a more general nature they can be applied not only in the firm where the entrepreneur is currently working, but also in other situations, such as a new venture (Blundell R., Dearden, Meghir, & Sianesi, 1999). In starting a new firm it will often be the case that the skills of the entrepreneur do not completely match with the skills they need. They may enterprise in a field that is new but especially the skills to build up a firm in general will be new to the entrepreneur. However the entrepreneur can also gain experience without having the actual experience him/herself, for instance entrepreneurial parents or close connections may also spill over their experiences onto the entrepreneur (Verheul, Thurik, Hessels, & Zwan, 2010) (Chrisman, Bauerschmidt, & Hofer, 1999). Furthermore it seems to matter whether the entrepreneur is still connected to other knowledge by having a part-time job, as entrepreneurs that start a firm and still are part-time employed seem to perform better than those without a part-time job (Sanditov & Verspagen, 2011) (Verheul, Thurik, Hessels, & Zwan, 2010), the authors however do not mention whether the part-time job and the new firm are operating in the same field.

2.2.2 Socio-economic factors
As discussed before the motives of the entrepreneur can be split in two groups, necessity- and opportunity-driven motives, the necessity-driven entrepreneurs is pushed into entrepreneurship as there are no alternatives while the opportunity-driven entrepreneur is pulled by motives like independence, wealth, challenge, recognition and status (Wilson, Marlino, & Kickul, 2004) (Gelderen & Jansen, 2006) (Cassar, 2007). Necessity-driven entrepreneurs involve themselves in entrepreneurship because they feel that it is their only choice, while opportunity-driven entrepreneurs engage themselves because the opportunity is too good to let pass. The common sense about entrepreneurship is that it is a risky business to be involved in as it takes a lot of hard work and long our while payment is often far away. This makes the incentives even more important as you need a strong will in order to succeed, therefore any obstacles that decrease the incentives of potential entrepreneurs will decrease the number of entrepreneurs. Furthermore the possible inputs to come up with ideas and the possible areas in which these opportunities can be pursued needs to be as big as possible. So this means that the more the potential entrepreneur is engaged in
challenging areas and the more markets there are available the more possibilities the entrepreneur has to run into an opportunity.

Personality characteristics from the previously mentioned article by Chrisman, Bauerschmidt & Hofer (1999) are likely to influence the propensity to enterprise as entrepreneurship should be ‘in your veins’. The values and beliefs of the entrepreneur should also match the entrepreneurial lifestyle (Chrisman, Bauerschmidt, & Hofer, 1999). However one could question the notion of values and beliefs being purely a personal variable as a lot of values and beliefs are formed through the culture the entrepreneur lives in and the entrepreneur’s attitudes towards religions. A quick and relevant example could be ‘the American dream’. Country specific culture may be a big influencer for entrepreneurship as it can determine for instance the attitude towards risk or the orientation towards a day to day job. As cultural differences can have their impact on entrepreneurship they may predict part of the cross country differences (Freytag & Thurik, 2007) (Wennekers S., 2006). One specifically important aspect of a country’s culture is social capital as it is important for the entrepreneur because it facilitates support for opportunity perception and access to finance (Kwon & Arenius, 2010) (Shane & Cable, 2002). Investors gain information about entrepreneurs through their networks, by which they find the right finance seeking entrepreneurs.

The business climate can have a serious impact on entrepreneurship as well, since starting a firm is dependent on both consumer demand and the availability of finance. Therefore, the level of economic development will play a role, most commonly incorporated as the level of GDP per capita. There are also other economy driven factors like the R&D expenditure, labor related variables and available markets that play a role. R&D expenditure is a measurement for knowledge creation which can be a cornerstone for entrepreneurs in terms of new possibilities (Sanditov & Verspagen, 2011). In terms of labor related variables the dispersion of wages and unemployment benefits can have a great impact on the motivation of the potential entrepreneurs. Available markets can play a role when it comes to available opportunities as well.

2.2.3 Government policies and procedures
In an article by Hartog, Stel & Storey (2010) they discuss the implications of four different government policies and procedures for entrepreneurship. The four they discuss are: social security entitlements, taxes, regulation and the rule of law. Taxes and social security entitlements can have a big impact on the attractiveness of entrepreneurship, as they influence the benefit difference between entrepreneurship and a payroll job. Regulations and the rule of law both can hamper the startup of a new firm or even increase the competition the entrepreneur receives from established firms. When talking about governmental services which the entrepreneur needs it seems evident to also incorporate corruption and bureaucracy quality in the dynamics as these both can have a huge impact on the efficiency of these governmental services (Hartog, Stel, & Storey, 2010) (van de Walle, 2005).

2.2.4 Financial and Non-financial assistance
To build up a firm the entrepreneur needs a lot of capital which usually is far more than the average savings account. Therefore financial assistance is essential for the success of the start-up. However there are multiple stages in which different investors may be interested in investing in the firm. These financial investments for start-ups can be differentiated into three stages. The first is the development stage: this stage is financed by the entrepreneur himself together with strategic
partners, friends and family. During this period numerous bootstrapping techniques can be used to overcome the capital shortages (Smith & Smith, 2004). After this stage the start-up is confronted with the start-up/early growth phase. In this phase seed capital is the source of income mostly funded by venture capitalists or angel investors. The final stage is based on rapid growth and exit strategies. At this stage, there are multiple possible financial sources like venture capitalists, trade credit and eventually the IPO (Initial Public Offering). When discussing these types of finance it is important to keep in mind that entrepreneurs could not only be influenced by the actual availability of capital but merely by their own perception of this availability (Verheul, Thurik, Hessels, & Zwan, 2010).

Bootstrapping techniques are merely consisting of different types of loans (credit cards, friends and family, banks and loans against property) and personal savings. The biggest two, covering somewhere between 80% and 90% of the total amount, are the personal savings and credit card loans (Smith & Smith, 2004). This, indicates that especially in the early phases of start-ups the richer the entrepreneur is, the bigger the chance is that the entrepreneur can pull his business through the development phase.

The next round of capital inflow is called seed capital. This is the capital used to really start building a firm just after initial development is done. At this point in time the entrepreneur can really demonstrate the prospects of the firm whereas in the previous stage it could only show an idea. Although still hard it is possible to attract investors from outside his/her network to the start-up, like for instance venture capitalists. This is also the stage in which most subsidies can be claimed, as governments require some prospective plans before committing their funds.

Financing in the later stage is often a far lesser problem as success (or failure) is far easier to assess at that point in time, decreasing the risk on investment which is a huge benefit as both risk tolerant investors (like venture capitalists) and risk averse investors (like banks) are now potential investors. In the end there is the exit strategy, which for a lot of start-ups results in an IPO (initial public offering), but there are numerous other possibilities. At this point the founders of the company and the investors get paid for the risks they have taken and effort they put into the firm. This may have an impact on new entrepreneurs as they will have some point of reference of the prospects of success.

Non financial assistance comprises of all the kinds of help the entrepreneur receives for free. One of the main types is free information, like for instance the “innovation voucher” project in the Netherlands. Often these programs are temporary and highly country-specific. This makes it almost impossible but also less relevant to this study as these temporary programs are not the cornerstones of the institutional arrangement within a country.

2.3 Comparative institutional analysis approach
Along the same line as the varieties of capitalism (VoC) approach explains differences in innovative activity, the aim in this study is to find the institutions that facilitate the factors that influence entrepreneurship as we have seen in the previous chapter. These institutions can then be used as a measurement of similarity amongst countries to see whether there are specific varieties of countries, and whether, based on these institutions, countries have different rates of entrepreneurial activity. The same approach has been used by Hall and Soskice as they created three groups of countries and investigated whether these groups showed a difference with respect to innovation. The two main
research questions here will be focused on whether the groups found in the VoC approach are also present in this research and whether the groups found have any relation with the rate of entrepreneurship in countries. The two main research questions will be as following:

Research question 1:
*Do varieties of capitalism exit in the context of entrepreneurship?*

Research question 2:
*Do differences in institutions explain differences in the rate of opportunity driven entrepreneurship across countries?*

The hypothesis for the first question is rather trivial as the groups will rather correspond or they don’t. Since innovation and entrepreneurship are dependent on slightly different institutions, one can expect that the groups of countries are not entirely the same. For the second research question the hypothesis is that each of the key institutional areas will indeed show a significant difference. The underlying factors that facilitate the differences between the country groups will be explained through the variables that will be used to create the country groups.

In the next sections different institutional areas will be explained. In addition, the relation between each institutional area and entrepreneurship will be made concrete creating hypotheses for each of these effects. Please note that the earlier mentioned socio-economic factor will be split into economic factors and culture factors to create a clearer overview, so a distinguishing will be made between five institutional areas.

2.3.1 Entrepreneurial, business & technical skills

To increase the knowledge base and skills the entrepreneur needs to start up his/her firm different countries offer different possible types of education. In the previous sub-chapter the notions of firm-specific and sector-specific knowledge was already discussed but not is the utter detail. Lam (2000) distinguished between four types of education configurations, which are a combination between two different dimensions. By creating a matrix of these two dimensions you get the four different types of education. The first dimension is the education within a specific firm or within a specific sector. Firm specific skills are very much focused on the skills needed for the position within the firm, while sector specific skills are far more equalized between firms. The second dimension is based on the formal education system which can create a great divide between ‘the academic elite’ and the working class, making intermediary skills far less appreciated. This could push more people to work hard to gain this theoretical knowledge while the people in the counterpart economies will be more likely to gain knowledge through experience which will be based on more tacit and contextual knowledge.

*Skills*

The incentives for employees to expand their knowledge base may not always be as straightforward. Current literature states that the dependence of the employee of the market is reduced by employment and unemployment protection. However, Estevez-Abe, Iversen and Soskice (1999) argue that these conclusions about this process are misleading. They argue that the dependence is actually increased as welfare measures like employment protection, unemployment protection and wage protection create an incentive to expand their skills more specifically to the firm/industry, which makes the employees more vulnerable in case of being fired. By contrast, employees without a
safety net will aim to make themselves less vulnerable in terms of skills, resulting in investments in more generic skills, including skills that can be used to set up an own firms. They conclude that employment protection will result in higher investments in firm-specific skills and unemployment protection is industry-specific skills.

Hypothesis 1:

*The more employees within a country have firm-specific skills the fewer entrepreneurs in a country.*

**Education**

Education is a great asset for employees, therefore also for entrepreneurs who want to start a new firm. However the effectiveness of education decreases as the level of education of the individual rises. In other words when an entrepreneur starts a firm in a field in which he/she has no education at all he/she will benefit the most from education and vice-versa (Hanoch, 1967) (Vaillancourt, 1995). This also relates to the type of education as practical education is much easier to directly apply to the working environment than scientific education. When education is specified to the topic of entrepreneurship it positively relates to the entrepreneurial activity of opportunity-driven entrepreneurs. The effects of education are however different for men and women, Butcher and Case (1994) find that women are able to gain more from education than men.

Hypothesis 2:

*The more people in a country have a practical education the more entrepreneurs in a country.*

Hypothesis 3:

*The higher the education level of people in a country is the more entrepreneurs in a country.*

**Experience**

When specifically looking at entrepreneurial education it is important to look at the differences between the types of information learned at universities/business schools and entrepreneurial education. In an overview article by Garavan and O’Cinneide (1994), it is explained that most of the learning done in schools is based on judging about information, while the entrepreneurial learning is about judging people and their incentives (Gibb, 1993). The authors also stress the importance of different ways of learning to be effective, including concrete experience, reflective observation, abstract conceptualization and active experimentation (Davies & Gibb, 1991). Thus, there are quite a couple of experiences that can only be gained through the actual process of starting up a firm, or through networking with existing entrepreneurs who can share their experiences.

Hypothesis 4:

*The more entrepreneurs are present in the local networks of a country the more entrepreneurs in a country.*

**2.3.2 Economic factors**

**Dispersion of wages**

Earlier in this work the distinction was made between coordinated market economies and liberal market economies. In terms of labor market institutions one could make the same type of distinction. Bassanini and Ernst (2002) define a country coordinated as a number of industrial relations are controlled. The first is the wage bargaining process which needs to be centralized or
coordinated by a certain actor or a group of actors. This can be done among employers or headed by trade unions. Other authors come to the same result as it comes to the notion that labor union institutions decrease the dispersion of wages (Freeman, 1982) (Freeman, 1998). The second and third points concern decision-making and free-riding problems, which both do not really influence the number of starting entrepreneurs.

When talking about the dispersion of wages, one should always keep in mind that the decision to look for a new job or a better paying job is relative to the current financial situation (Freeman, 1998). A person will not risk starting a new firm if the possible return is not worth the risk, meaning the money he receives today (from salaries or unemployment benefits) affects the decision to start a new venture. The dispersion of wages is closely linked to the incentive to work a lot or just work lesser hours. One could argue that the more you earn the less you need to work for the same payment but evidence points in a different direction. Work by Juhn, Murphy, & Topel (1991) shows that actually the opposite is true as the higher payment also gives the worker the incentive to spend more time at their job. This means that the worker reasons that an extra hour of work is worth more with a high salary than when paid less. They also argue that the same goes for other occupations as the worker will purely weight the pro’s and con’s of each possibility including all kinds of behavior (changing jobs, studying instead of working, not working at all, etc.). In this process institutional arrangements can fulfill an important and influential role.

Hypothesis 5:

*The higher the dispersion of wages the more entrepreneurs in a country.*

**Unemployment benefits**

This relates to the availability of unemployment benefits as many researchers have found that higher unemployment benefits, the lower the incentives to look for a job (Katz & Meyer, 1990) (Atkinson & Michlewright, 1991). As this gap gets smaller the incentives to look for a new source of income also get smaller. They also found that as soon as eligibility for unemployment ends the incentive for this new source increases dramatically resulting in a drastic increase in the job-finding rate. This also means that generous social security benefits for employees increase the opportunity costs of entrepreneurship. However, social security may also have a positive effect on entrepreneurial activity by creating a safety net in case of business failure. However, empirical results show that higher social security mostly increases the opportunity costs (Hessels, Stel, Brouwer, & Wennekers, 2007) (Wennekers, Uhlauer, & Thurik, 2002) (Parker & Robson, 2004) (Hartog, Stel, & Storey, 2010). Secondly, the higher social security contribution the entrepreneur has to offer his employees increase the wage costs (Hessels, Stel, Brouwer, & Wennekers, 2007).

Hypothesis 6:

*The higher the unemployment benefits the fewer entrepreneurs in a country.*

**Sectors**

Other direct implications of the welfare state system that are based on the institutional arrangements are linked to market possibilities. The fact that a couple of big sectors are controlled by the government, like for instance health care, decreases the number of possible sectors in which entrepreneurs can exploit their ideas (Henrekson, 2005). Furthermore, there could also be limits on the amount of trade a firm can pursue across borders. Therefore the openness to export is important.
Hypothesis 7:

*Countries with more available sectors have more entrepreneurs:*

7a) The bigger the public sector the fewer entrepreneurs.

7b) The higher the export (per GDP) in a country the more entrepreneurs.

**GDP**

Many researchers have found that the differences between countries in the relation between economic development and entrepreneurship are U-shaped (Blau, 1987) (Acs, Audretsch, & Evans, 1994) (Carree, Stel, Thurik, & Wennekers, 2002) (Wennekers, Stel, Thurik, & Reynolds, 2005). The rate of new start-ups is thus declining with economic development only to pick up again after the country reaches higher levels of development. However here we need to refer again to the work on the split up between necessity- and opportunity-driven entrepreneurs as in weak economies there will generally be more necessity-driven entrepreneurs. Considering this, it turns out that higher levels of economic performance indeed relate to higher numbers of entrepreneurs.

Hypothesis 8:

*The higher the GDP level of a country the more entrepreneurs in a country.*

**Business cycles**

Sanditov and Verspagen (2011) note that it is important to take into account the short term dynamics of economic growth. Therefore they include GDP growth in their model. The amount of entrepreneurs will gain momentum on the same moments as the business cycle of a country. High economic growth often comes with a lot of possibilities in terms of finance. All types of capital are freed up in times of prosperity as they are not negatively affected by the economic cycle. For entrepreneurs this is really important as investing capital in a new firm is never without risk, meaning that this capital is only available in times when risks are acceptable. Opportunities may also be a lot scarcer in times of crisis for both high growth entrepreneurs as for small entrepreneurs. High growth entrepreneurs are often bought by other firms or work in projects together with other firms and governments. Small entrepreneurs will often deal with consumers directly making them also very dependent on the market. For both these types of entrepreneurs it will be very hard to make their business work as they are dependent on others that are not willing to corporate due to the economic climate. This can also be argued the other way around as a better economic climate corresponds to more opportunities.

Hypothesis 9:

*The higher the country’s GDP growth the more entrepreneurs in a country.*

**Knowledge creation**

The more under-utilized knowledge a country has the more opportunities there are for knowledge intensive entrepreneurs (Acs, Braunerhjelm, & Audretsch, 2009). This knowledge that can spill over to the entrepreneur is developed through R&D expenditures which can be done by very different actors in different countries, as it can be done by firms, universities or public research organizations (Acs, Braunerhjelm, & Audretsch, 2009). Either way these R&D expenditures are important to entrepreneurs.

Hypothesis 10:

*The higher the R&D-expenditures in a country the more entrepreneurs in a country.*
2.3.3 Cultural factors

When including cultural differences the most commonly used data are the cultural dimensions measured by Geert Hofstede in the end of the 20th century (Hartog, Stel, & Storey, 2010). He makes a distinction between four cultural dimensions relevant to business: inequality, individualism, roles between genders and uncertainty (Hofstede G., 2001). Later a fifth dimension was added by Michael Minkov based on the differences between long-term and short-term orientation (Hofstede, Hofstede, & Minkov, 2010).

Power distribution

The first dimension is called power distance, which indicates how the less powerful actors in the society accept the unequal distribution of power within the country (Hofstede G., 2001). For entrepreneurs this plays a key role as it shows that people within a country are more into following a leader or taking matters into their own hands. In the first case they accept the inequality, and will therefore not try to change it. In the second case they do not accept power distribution inequality hence will more likely start their own business, with as reason to be more independent of the present authority.

Hypothesis 11:

The more people are concerned about the distribution of power within a country the more entrepreneurs in a country.

Individualism

The second dimension is called individualism, which is very closely linked to social capital as this dimension explains how the ties between individuals within the country are build up. Some people will be very self-centered only looking after their own family which is very individualistic, while there are also countries that consist of more cohesive groups in which individuals are strongly interconnected with their society (Hofstede G., 2001). We have seen that social capital is extremely important for entrepreneurs, even if entrepreneurs do not necessarily have a bigger network than non-entrepreneurs. Yet, entrepreneurs do seem to be able to use their network to a far better extent than non-entrepreneurs (Liao & Welsch, 2005). Having better access to social capital than others can be beneficial to entrepreneurs, which leads us to the differences between countries as they are evident and correlated to differences in entrepreneurial activity (Kwon & Arenius, 2010). Being very individualistic or not relates to the amount and the types of people that will be able to help the entrepreneur. Very individualistic people often have better access to external help while collectivistic people often have more help from people close to them. External help consists of access to services that are often not (totally) free but even though hard to get, one could think of market information or capital investment possibilities. Help from people who are close to the entrepreneurs often consists of free services, like for instance family loans or free labor. As both types of resources can be crucial in the easy stages of the firm the effects of individualism on entrepreneurship will be U-shaped, as both cohesive as non-cohesive networks can create benefits for the entrepreneur.

Hypothesis 12:

There is an inverted U-shaped relation between individualism and the number of entrepreneurs in a country.
Gender
The third dimension is based on the differences between the sexes. For societies where the roles between sexes are very clear risk taking is a male task, which means that there are less people who have the potential of becoming an entrepreneur.

Hypothesis 13:
The clearer the gender roles are within a country the fewer entrepreneurs in a country.

Uncertainty avoidance
The fourth dimension is called uncertainty avoidance, which relates to how people feel in unstructured situations (Hofstede G., 2001). Unstructured situations are different from the usual situations in which people know what they can expect. For entrepreneurship, this dimension is expected to be important as the entrepreneur will usually face many new and unstructured situations. Not only is starting a firm, and all the processes involved, new, but maybe the product or service they intend to offer as well. In countries with lower uncertainty avoidance, “not only familiar but also unfamiliar risks are accepted, such as changing jobs and starting activities for which there are no rules.” Low uncertainty avoidance thus implies “willingness to enter into unknown ventures” (Wennekers, Thurik, Stel, & Noorderhaven, 2007).

Hypothesis 14:
The higher the uncertainty avoidance within a country the fewer entrepreneurs in a country.

Future orientation
The final dimension is based on the difference between long-term and short term orientation. The type of orientation within a country can be very different, as people in some countries are focused on future benefits and others are more focused on the here and now. This might have a strong implication on the number of new firms, as a short-term orientation does not match with entrepreneurship. When starting a new firm you generally get paid for your efforts in the end of the startup phase which is not at all a short-term orientation. However when starting a new firm the entrepreneur loses the certainty of career growth, a pension and material benefits like a home, car or just pure savings. Therefore the best orientation for the entrepreneur should be somewhere in the middle of the two extremes.

Hypothesis 15:
There is an inverted U-shaped correlation between future orientation and the number of entrepreneurs in a country.

Role models
The performance for previous entrepreneurs may differ a lot between countries as well the total number of new start-ups. The entrepreneurs that become very successful may perform an important function as role model, which is also emphasized in the work of Gnyawali & Fogel (1994). Keeping this in mind there might be a connection between the number of extremely successful firms from a country and the number of entrepreneurs from that country, as these successful firm act as an example for all entrepreneurs who are skeptical about the success chance. Although these firms might not all have been started from scratch they do tell the success story that some people may be just as successful. This will of course only play a role for entrepreneurs that want to grow out to become a big firm. However one should keep in mind that this will be correlated to opportunity-
driven entrepreneurs. Extremely successful firms will in this research be limited to the number of firms which made it to the world top, because these are the firms that people will come in contact with the most and these are thus the companies that show the (business) pride of the country. However the bigger the number of companies a country has the less impressive a single firm becomes, meaning that the absolute number will have to be related to the number of domestic firms a country has.

Hypothesis 16:

*Countries with more companies in the top 2000 have more entrepreneurs.*

2.3.4 Government policies and procedures

**Taxes**

Taxes have multiple benefits for entrepreneurs when compared to employees as the entrepreneur declares his own taxes directly to the authorities and not through an employer, the entrepreneur can pay his taxes at the end of the year and he can claim more expenses (Hartog, Stel, & Storey, 2010). Therefore, it is not surprising that Parker & Robson (2004) found a positive and significant relation to the personal income tax rates. However, higher tax rates also negatively influence the entrepreneur as the income they get is usually the “prime source of funding” (Michaelas, Chittenden, & Poutziouris, 1990). Secondly, by increasing the income tax the entrepreneurial chances in the service sector decrease a lot as the expenses in this sector are largely related to personnel time. Therefore higher service costs will cause the customer to consider dealing with the problem him/herself (Henrekson, 2005).

Hypothesis 17:

*The higher income taxes within a country the fewer entrepreneurs in a country.*

**Regulations**

Djankov, Porta, Lopez-de-Silanes, & Shleifer (2002) found in their research that due to different regulations both the time and costs of creating a business differed a lot between countries. As a result there are numerous studies done to check whether regulation of the governments created to protect the customer and creditors were suffocating the entrepreneur through bureaucracy. Many of them have found negative results concerning this issue (Ardagna & Lusardi, 2009) (Stel, Storey, & Thurik, 2007).

Because policymakers try to create support for entrepreneurs, a lot of these time consuming and costly regulations have been eliminated, significantly decreasing the earlier mentioned differences between countries (Hartog, Stel, & Storey, 2010). Adding to that Baumol (2007) and van Stel, Storey & Thurik (2007) found that the regulations mostly influence the distribution of entrepreneurs between different activities. In the end bureaucracy forces the entrepreneur to put a lot of unnecessary effort in creating his/her firm, whether it is in the official or unofficial circuit.

The quality of the bureaucracy in a country is something that can hamper or speed up the flexibility with which a new firm can be started. To determine what the performance or quality of the bureaucracy within a country is very hard. As work by van de Walle (2005) shows the notion of bureaucracy is a very vague term consisting of a broad range of concepts. The same can be seen in the work by Rainey & Steinbauer (1999) as they try to create a theory that shows how to cope with
all the different tasks that governments deal with. They use metaphors to express that a government is this large intangible thing that can be very bad organized while it still excels at a couple of sub-tasks. Van de Walle (2005) quotes the work of Putnam, Leonardi & Nanetti (1993) when concluding that the “bottom line” evaluation is practically impossible. Either way, the issue remains that the entrepreneur has to cope with the government to gain licenses and so on, as we have seen in the regulation part of this chapter that can have big impacts.

Hypothesis 18:

The more efficient the bureaucracy in a country is the more entrepreneurs in a country.

Rule of law

Concerning the rule of law, intellectual property rights (IPR) may affect the rate of entrepreneurship in a country. Some argue that these rights are very useful to entrepreneurs as they clearly need the protection from competitors (Nyström, 2008). The counter argument holds that when entrepreneurs are benefitting so are large firms. The latter should be able to more easily respond to, and perhaps even manipulate it in their favor. Especially if you investigate the increase of lawsuits in the past decade, IPR may not work in favor of the low-capital-entrepreneurs as they cannot afford a long term lawsuit.

Hypothesis 19:

Countries with strong intellectual property rights have fewer entrepreneurs.

Corruption

When discussing corruption it is important to make a clear distinction between the actual rules that are active in a country and the enforcement of these laws. In countries in which corruption is a problem the enforcement of the rules is sub-optimal. The enforcement of rules is done to keep both the citizens and the government of a country in line. Because of these two parties there can also be two types of problems for entrepreneurs as in the first scenario other people or firms within the country can hinder the startup and in the second scenario it can be the government that slows down the startup. When people or other companies hinder the startup they can bend the enforcement of the rules in their favor. While in case of the government, or its employees, the startup is dependent on their work making them easy target to all kinds of schemes. The work by Freeman (1998) relates this to the incentives of the different actors as they all weigh the benefits of different actions to the costs. The costs can even be paying a penalty for doing something illegal, which may be lower than the benefits from bending the rules. Startups are very vulnerable in these countries as they are often ‘weak’ even if they can generate great value.

Hypothesis 20:

The more corrupt the country is the fewer entrepreneurs in a country.

2.3.5 Financial assistance

Development capital

As we have seen in the previous chapter about the determinants of entrepreneurship the development stage is mostly financed through personal savings and credit card loans. These two may differ a lot between countries. The saving behavior of individuals is greatly influenced by the institutions as numerous arrangements like unemployment insurance, income-dependent pensions,
sick-leave benefits, higher education, highly subsidized healthcare remove the incentive for people to have spare capital in case of emergencies (Henrekson, 2005). This in combination with an even distribution of wealth, often seen in these welfare states, creates problems when an entrepreneur tries to raise enough capital to realize a business opportunity, as the capital is more spread amongst individuals. However the fact that savings are related to the institutional arrangement only decreases the differences between the different societies as people with savings will not invest it in new opportunities if they are extremely dependent on it in terms of emergency security.

Finance hypothesis 21:

Countries with more available development capital have more entrepreneurs:

21a) The higher the savings people within a country have the more entrepreneurs.
21b) The more property people within a country possess the more entrepreneurs.
21c) The easier an entrepreneur can get a loan the more entrepreneurs.

Seed capital

According to a WRR report by Nootbeoom and Stam (2008) the availability of seed capital is a huge problem for entrepreneurs in Europe as venture capitalists seem to enter the process at a later state. Also Pearce, de Yonge, Haemmig, Davila and Foster (2011) present big differences across countries on numerous issues concerning venture capitalists (VCs) in their part in the World Economic Forum Report of 2011. When looking at the stages in which the different countries invest there are big differences: European countries tend not to invest in the early phases while China invests the most of all countries in the latest phase. As second difference they point out is the amount of money spent by VCs. The US is leading in this respect. Not surprisingly, there is the same type of difference when looking at the number of VC-backed companies. When looking at the industries in which these investments are made there are also substantial differences. This could be linked to the level of development in the country as Europe and the US show similar investments while very different from China and India, who also show similarities. As alternative or complement there are a lot of different subsidies, created to help entrepreneurs cross these gaps.

Finance hypothesis 22:

Countries with more available seed capital have more entrepreneurs:

22a) The more access to early VC-capital within a country the more entrepreneurs.
22b) The more access to subsidies within a country the more entrepreneurs.
In table 3 an overview of all the hypotheses is presented. After this table the next chapter covering data collection starts.

**Table 3: Overview of the hypotheses**

<table>
<thead>
<tr>
<th>Hypotheses overview</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Entrepreneurial, business &amp; technical skills</strong></td>
</tr>
<tr>
<td>1  The more employees within a country have firm-specific skills the fewer entrepreneurs in a country.</td>
</tr>
<tr>
<td>2  The more people in a country have a practical education the more entrepreneurs in a country.</td>
</tr>
<tr>
<td>3  The higher the education level of people in a country is the more entrepreneurs in a country.</td>
</tr>
<tr>
<td>4  The more entrepreneurs are present in the local networks of a country the more entrepreneurs in a country.</td>
</tr>
<tr>
<td><strong>Economic factors</strong></td>
</tr>
<tr>
<td>5  The higher the dispersion of wages the more entrepreneurs in a country.</td>
</tr>
<tr>
<td>6  The higher the unemployment benefits the fewer entrepreneurs in a country.</td>
</tr>
<tr>
<td>7  Countries with more available sectors have more entrepreneurs:</td>
</tr>
<tr>
<td>• The bigger the public sector the fewer entrepreneurs.</td>
</tr>
<tr>
<td>• The higher the export (per GDP) in a country the more entrepreneurs.</td>
</tr>
<tr>
<td>8  The higher the GDP level of a country the more entrepreneurs in a country.</td>
</tr>
<tr>
<td>9  The higher the country’s GDP growth the more entrepreneurs in a country.</td>
</tr>
<tr>
<td>10 The higher the R&amp;D-expenditures in a country the more entrepreneurs in a country.</td>
</tr>
<tr>
<td><strong>Cultural factors</strong></td>
</tr>
<tr>
<td>11 The more people are concerned about the distribution of power within a country the more entrepreneurs in a country.</td>
</tr>
<tr>
<td>12 There is an inverted U-shaped relation between individualism and the number of entrepreneurs in a country.</td>
</tr>
<tr>
<td>13 The clearer the gender roles are within a country the fewer entrepreneurs in a country.</td>
</tr>
<tr>
<td>14 The higher the uncertainty avoidance within a country the fewer entrepreneurs in a country.</td>
</tr>
<tr>
<td>15 There is an inverted U-shaped correlation between future orientation and the number of entrepreneurs in a country.</td>
</tr>
<tr>
<td>16 Countries with more companies in the top 2000 have more entrepreneurs.</td>
</tr>
<tr>
<td><strong>Government policies and procedures</strong></td>
</tr>
<tr>
<td>17 The higher income taxes within a country the fewer entrepreneurs in a country.</td>
</tr>
<tr>
<td>18 The more efficient the bureaucracy in a country is the more entrepreneurs in a country.</td>
</tr>
<tr>
<td>19 Countries with strong intellectual property rights have fewer entrepreneurs.</td>
</tr>
<tr>
<td>20 The more corrupt the country is the fewer entrepreneurs in a country.</td>
</tr>
<tr>
<td><strong>Financial assistance</strong></td>
</tr>
<tr>
<td>21 Countries with more available development capital have more entrepreneurs:</td>
</tr>
<tr>
<td>• The higher the savings people within a country have the more entrepreneurs.</td>
</tr>
<tr>
<td>• The more property people within a country possess the more entrepreneurs.</td>
</tr>
<tr>
<td>• The easier an entrepreneur can get a loan the more entrepreneurs.</td>
</tr>
<tr>
<td>22 Countries with more available seed capital have more entrepreneurs:</td>
</tr>
<tr>
<td>• The more access to early VC-capital within a country the more entrepreneurs.</td>
</tr>
<tr>
<td>• The more access to subsidies within a country the more entrepreneurs.</td>
</tr>
</tbody>
</table>
3. Data

In the varieties of capitalism theory, groups of countries are formed because the institutions in their countries are structured in the same way, which means that each of these countries approach certain key topics in the same way. For instance, in coordinated market economies government involvement is very high (hence the name coordinated) and in liberal market economies this involvement is very low. The involvement of the government is only one of the dimensions that indicate the institutions of a country. There are many institutions that are structured in a specific way which can be the result of active changes but also the shaping of many years of politics and events.

The main reason why institutions of countries can be grouped is because a certain combination of institutions will complement each other. The ideas in the work of Hall and Soskice on the varieties of capitalism were focused on innovation as they argued that the institutional arrangement of the two main types of economies also influenced the capabilities to innovate and the type of innovativeness. In this case we are still unsure whether the two types of economies discussed by Hall and Soskice are also of relevance for entrepreneurship.

The various institutions are headed under institutional pillars. These pillars are therefore the representation of the related institutions. As we have seen in the literature review about entrepreneurship there are a couple of core institutional pillars, namely: Education and skills, Economic factors, Cultural factors, Governmental factors and the Financial factors. Now before these will be called pillars it is very important that these variables form a combination that really represents the nature of this pillar, because as soon as these do represent a pillar these pillars can be used to test whether the same varieties of countries exist as were found for innovation in the research of Hall and Soskice (2001). In the next table the variables are divided by the way the literature on entrepreneurship suggested this. Please note that these pillars, showed in table 4, are still preliminary as in the next chapter the data gathering will be discussed, through unavailable data or other data related issues the variables can become a bad representation to the rest of the cluster therefore the final clusters will be discussed after the data gathering chapter.

Table 4: Preliminary pillars

<table>
<thead>
<tr>
<th>Preliminary pillars</th>
<th>Education &amp; Skills</th>
<th>Economy</th>
<th>Culture</th>
<th>Government</th>
<th>Finance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firm-specific &amp; Sector-specific skills</td>
<td>Wage dispersion</td>
<td>Power distribution</td>
<td>Income tax</td>
<td>Development capital: • Savings</td>
<td></td>
</tr>
<tr>
<td>Practical education</td>
<td>Unemployment benefits</td>
<td>Social capital</td>
<td>Bureaucracy inefficiency</td>
<td>Development capital: • Property</td>
<td></td>
</tr>
<tr>
<td>Education level • Size tertiary education • Quality tertiary education</td>
<td>Available sectors: • Public sector size • Export</td>
<td>Gender roles</td>
<td>Intellectual property rights (IPR)</td>
<td>Development capital: • Loans</td>
<td></td>
</tr>
<tr>
<td>Experience through local networks</td>
<td>GDP per capita</td>
<td>Uncertainty avoidance</td>
<td>Corruption</td>
<td>Seed capital: • Venture capital</td>
<td></td>
</tr>
<tr>
<td>GDP growth</td>
<td>Future orientation</td>
<td></td>
<td></td>
<td>Seed capital: • Subsidies</td>
<td></td>
</tr>
<tr>
<td>R&amp;D expenditure</td>
<td>Role model firms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.1 Data collection

3.1.1 Dependent variable
This variable is key to this analysis as this is the subject of this study. This means that mistakes in this variable can have serious effects on the results and conclusions of our study. Currently the only database that supports measures of entrepreneurship across a large amount of countries is the Global Entrepreneurship Monitor (GEM) database.

The data
In the GEM database there are three classifications of people who run/own (part of) a firm, the first category is the group of firms which are just started, no older than three months, and the questioned person should not have received any wage or salaries, the second group are firms that have been existing a bit longer, three to 42 months, but are not fully developed yet therefore again the questioned person should not have received any wage or salaries yet. The last category consists of firms that exist longer than 42 months and the questioned person was paid for his/her service. In all these three groups it is important to note that the entrepreneur should both be (part) owner as manager.

The dependent variable extracted from the GEM database is the percentage of entrepreneurs on the total population between the ages 15-64 who belong to the first or second group. The GEM database calls this group the TEA group, which stands for Total Entrepreneurial Activity, the abbreviation TEA will come back in the analysis part of this report but will be explained again at that point. The GEM database makes a difference between opportunity-driven entrepreneurs and necessity-driven entrepreneurs. In our case, the focus is on opportunity-driven entrepreneurs. Note, however, that the notions of opportunity-driven and necessity-driven are to some extent intuitive. What counts as an opportunity may be to some extent country specific. For example, what counts as an opportunity in a developing country (say, to open a shop) may be seen as necessity-driven in a develop country.

The entrepreneurship data used in the research is mostly coming from the GEM survey of the year 2007, but as the teams that gather the entrepreneurship data are not always available every year the participating countries differ a bit through the years. Therefore this was used to increase the number of countries in the research by using the data of some countries from the year 2006, hereby increasing the total number of usable countries from 36 to 46 countries. In this case ‘usable’ means that for these countries there is also enough data available on the independent. The countries for which 2007 data was available are the following:

ARGENTINA, AUSTRIA, BELGIUM, BRAZIL, CHILE, CHINA, COLOMBIA, CROATIA, DENMARK, FINLAND, FRANCE, GREECE, HONG KONG, HUNGARY, INDIA, IRELAND, ISRAEL, ITALY, JAPAN, LATVIA, NETHERLANDS, NORWAY, PERU, PORTUGAL, ROMANIA, RUSSIA, SLOVENIA, SPAIN, SWEDEN, SWITZERLAND, THAILAND, TURKEY, UNITED KINGDOM, UNITED STATES OF AMERICA, URUGUAY, VENEZUELA

For 2006:

AUSTRALIA, CANADA, CZECH REPUBLIC, GERMANY, INDONESIA, MALAYSIA, MEXICO, PHILIPPINES, SINGAPORE, SOUTH AFRICA
Especially the additions of Australia and Canada are important for the theory of the varieties of capitalism as they a big part of the liberal market economies.

The survey
The GEM database is formed by a survey that is conducted every year on a country level basis. The countries should have at least 85% telephone network coverage for the GEM to rely on telephone interviews, while countries below this 85% coverage will be approached in a face-to-face style. Each country has its own national team that must conduct the survey with at least 2,000 adults using ‘the best practice social survey techniques’. These survey techniques are however not discussed in detail as each national team will employ a professional survey firms to conduct the survey on a random Digit Dialing basis. GEM does submit an annual report with a lot of facts of the data gathering of that particular year, plus the actual survey that was used that year. The survey includes all questions that are asked to people between the ages of 15-64, this gives great insight into how certain details are obtained. The annual report includes matters such as the members of the national teams and the exact number of observations. From this, it becomes apparent that some national teams consist of over 10 people while others are a lot smaller. Their background and financial sponsors also seem to differ quite a bit. Nevertheless, most of the teams are backed up by university institutions and most sponsors are government related.

Some countries do not achieve the goal of reaching 2,000 adults while in others many more than 2,000 have been interviewed. To give an example the US, Japan, France, Croatia and Hungary are all in between 1.500 and 1.600 observations while both the UK and Spain go beyond the 25.000.

The survey the GEM uses to categorize entrepreneurs in the opportunity-driven or necessity-driven group does try to capture this effect by asking about the availability of other employment possibilities. However it seems to me that they are not able to facilitate a 100% correct measurement because of the previous discussed process. In figure 3 a snapshot of the question from the used questionnaire is shown.

Figure 3: Entrepreneurial motivation GEM data

Other databases
There are other databases that cover the same type of subjects as the GEM database but then they are a lot less specific or are not freely available to any given researcher. One of the possibilities that is available through for instance the World Bank databases, which represent the number of new firms that are newly registered in each country. This however gives a wrong impression of countries which don’t have a strong registration system. Furthermore, and this is probably the strongest reason to choose for the GEM database, the fact that the GEM database gives a distinction between opportunity-driven and necessity-driven entrepreneurs makes it the only possible database to use.
for this research. So even if the measurement of the number of opportunity-driven entrepreneurs is not perfect it is the only measurement available at this point in time.

**Final remarks on dependent variable**

As a final remark about the dependent variable it is worth mentioning that the opportunity-driven TEA group in the GEM database makes a distinction between different types of motivations, they mainly distinct income, independence and mixed reasons. A problem here holds when adding these three groups the total number of opportunity-driven entrepreneurs is not always corresponding to the data given by GEM. To clarify this contact was sought with the people responsible for the GEM organization. A response was however never received. Therefore the total number of opportunity-driven entrepreneurs (TEA) was used in this analysis.

### 3.1.2 Independent variables

For the independent variables the data will be discussed per hypothesis. For some hypotheses there will not be enough data available, these hypotheses will be dropped. Therefore these will be marked with a strikethrough. The sources are discussed in more detail in a separate appendix, appendix 1, as multiple variables were extracted from the same source.

**Education and skills pillar**

1. **The more employees within a country have firm-specific skills the fewer entrepreneurs in a country.**

   The first hypothesis revolves around firm-specific skills and sector-specific skills. Exact data about the nature of the skills people in a country have acquired is not available, but this will be approximated by the amount of persons working the service sector and the industry sector. People in the industry sector will be more bound to their firm while people in the service sector will have more portable skills. The reasoning behind this is that skills in the industry sector are more heavily based on knowhow than the skills in the service sector. The data of the percentage of people working in these sectors was freely available from the World Bank database, as explained before the way they have acquired this will be discussed later on.

2. **The more people in a country have a practical education the more entrepreneurs in a country.**

   The second hypothesis needs specific information about the nature of the education and the amount of practical education that is followed in a country. In this case data was found in the Global Education Digest 2009 made by Unesco Institute for Statistics. They specify the percentage of people following an academic study and a practical study, while both these were still classified as higher education. However, this variable should be used with caution as countries apply different definitions. To give an example, in the Netherlands almost all higher education is also partly academic based as the education prepares the student to be able to follow a masters at a university. Therefore the percentage practical versus academic education in the Netherlands is categorized as 1% to 99%. So although there are semi-practical studies in the Netherlands they are still categorized as partly academic and therefore not purely practical.

3. **The higher the education level of people in a country is the more entrepreneurs in a country.**

   The third hypothesis is based on the education level in a country, which will be measured by the percentage of people following tertiary education and the quality of the education system. The
percentage of people attending a tertiary education was extracted from the World Bank database, while the education system quality was extracted from the World Economic Forum (WEF).

4 The more entrepreneurs are present in the local networks of a country the more entrepreneurs in a country.

The fourth hypothesis requires data about the local network of the entrepreneur, which was in fact a question in the questionnaire of the GEM. They asked the entrepreneurs whether they are in contact with an entrepreneur that started a business in the past two years. This business should at least cost some time or resources, but did not necessarily had to have succeeded. As the GEM has been discussed in such detail in the dependent variable section this will not be repeated in at the end where the other data sources for the independent variables are discussed.

Economic pillar

5 The higher the dispersion of wages the more entrepreneurs in a country.

The fifth hypothesis deals to the dispersion of wages. From the Human Development Report 2007/2008 for the United Nations Development Program the difference between the richest 10%/20% of the population and the poorest 10%/20% of the population could be obtained. This shows the share of the income the rich get divided by the share of income the poor get.

6 The higher the unemployment benefits the fewer entrepreneurs in a country.

The sixth hypothesis is focused on the amount of unemployment benefits that each person is entitled to. A problem with this variable is that it is often not a fixed amount but that it is very dependent on the situation under which someone is unemployed. Although the International Labour Office provides a good estimation for a number of countries in their World Social Report 2010/11, it is not covering enough countries to include the variable in the analysis. Therefore, this hypothesis is dropped.

7 Countries with more available sectors have more entrepreneurs:
   • The bigger the public sector the fewer entrepreneurs.
   • The higher the export (per GDP) in a country the more entrepreneurs.

The seventh hypothesis is split up into two variables, namely the size of the public sector, and the amount of export a country has. The size of the public sector and the amount of export are both present in the database of the World Bank.

8 The higher the GDP level of a country the more entrepreneurs in a country.

9 The higher the country’s GDP growth the more entrepreneurs in a country.

10 The higher the R&D-expenditures in a country the more entrepreneurs in a country.

The eighth, ninth and tenth hypotheses are based on the GDP level, GDP growth and R&D-expenditure, which are all accurately measured by many sources. The measurements of the World Bank were used. The GDP growth was recalculated over a period of 5 years preceding the year 2008, so the average was taken over the years 2003 till 2007. This was necessary to measure the long-term growth of a country and not just temporary growth of one year.
Cultural pillar

11 The more people are concerned about the distribution of power within a country the more entrepreneurs in a country.
12 There is a U-shaped relation between individualism and the number of entrepreneurs in a country.
13 The clearer the gender roles are within a country the fewer entrepreneurs in a country.
14 The higher the uncertainty avoidance within a country the fewer entrepreneurs in a country.
15 There is a U-shaped correlation between future orientation and the number of entrepreneurs in a country.

Hypotheses 11 to 15 are all measurements from Hofstede. They were included because the theory fits the realm of entrepreneurship, and the data of Hofstede of course fits these perfectly. The Hofstede data is freely available on his own website.

16 Countries with more companies in the top 2000 have more entrepreneurs.

The sixteenth hypothesis is based on the notion of role models. Not only is it very hard to have a country-wide measurement of the amount of role models individual entrepreneurs have, it also has to be clearly distinct from the local network variable discussed earlier. Therefore role models were defined as leading firms, as everybody knows the biggest firms from his/her own country. In order to obtain a measurement for these role model firms, the total amount of firms that are in the top 2000 for each country was obtained from the Forbes website. In order to capture the chance an entrepreneur has to become such a leading figure the number of firm in this top 2000 was divided by the total number of domestic firms that a country has. The total number of domestic firms was obtained from the World Bank. One should be aware that this is only an approximation of the effect of role models.

Governmental pillar

17 The higher income taxes within a country the fewer entrepreneurs in a country.

The seventeenth hypothesis is dealing with amount of income tax the entrepreneur has to pay. Here the same problem as unemployment benefits arises as the income tax is not only depending on the country but also on the height of the income itself. Therefore was the highest marginal income tax picked as the most appropriate measure, while knowing that this is not perfect.

18 The more efficient the bureaucracy in a country is the more entrepreneurs in a country.

The eighteenth hypothesis revolves around the effectiveness of the bureaucracy, which is separately measured by the World Bank in their Worldwide Governance Indicators (WGI) project.

19 Countries with strong intellectual property rights have fewer entrepreneurs.

The nineteenth hypothesis is concerning the strength of the intellectual property rights. This is more an opinion than a fact as many of these systems work the same on paper but work differently in practice. We used data from the WEF.

20 The more corrupt the country is the fewer entrepreneurs in a country.

The twentieth hypothesis needs a measurement of corruption, which is available from many sources, but for this research the corruption index from the Global Corruption Report 2008 was used, as they compile a lot of other research into one number.
Financial pillar

21 Countries with more available development capital have more entrepreneurs:
   - The higher the savings people within a country have the more entrepreneurs.
   - The more property people within a country possess the more entrepreneurs.
   - The easier an entrepreneur can get a loan the more entrepreneurs.

The twenty-first hypothesis is split up in three items: the amount of personal savings, the amount of personal property and ease in which an individual can get a loan. Unfortunately there was no global data available of the amount of savings or property, so the ease to get a loan will have to do. The data on loans was obtained from the WEF.

22 Countries with more available seed capital have more entrepreneurs:
   - The more access to early VC-capital within a country the more entrepreneurs.
   - The more access to subsidies within a country the more entrepreneurs.

The twenty-second hypothesis revolves around the availability of venture capital and subsidies for the entrepreneur. The venture capital availability was described in a secondary GEM database, where they asked national expert on their opinion about the availability, the reason for this expert group was mostly based on their wide knowledge base about the possibilities which often is not sufficient for starting entrepreneurs. The availability of subsidies was also in this secondary GEM database in which they rated the availability of subsidies for the firm.
3.2 Variable structuring

Now that the data collection has been discussed it is become clear that some data was not available. For clarification they will be summed up once more: unemployment benefits, amount of personal savings and property ownership. As these were unavailable they will not be part of the pillars, and although this sound very logical this does implicate that the remainder of the variables will play a bit more of an important role in their pillar as the number of variables the pillar is constructed from is now lower. Now there are a couple of other variables that need some discussion as the data found on these variables may not be entirely representative for the pillars they belong to.

The first to discuss is the variable ‘Local Networks’. This variable indicates, as discussed earlier, the percentage of people who know an entrepreneur who has started a business in the past two years. The local network variable is part of the education and skills pillar because it should account for experience spillovers from other entrepreneurs, and the data available is very suitable for this role. However it also is a measure for role models on a local level, so in this case it is not the experience spillover that makes entrepreneurship more attractive but the effects of the role model function of the local network that creates this attractiveness. Simply put, people can be attracted to become entrepreneurs because of their skills basis or because they want to match the achievements of people they admire. Because of these dynamics it seemed best to exclude this variable from the pillar and to use it in the analysis as an independent variable separately from the pillars.

A different but same type of logic is applicable on the variable that describes successful firms that figure as role models, this variable is called ‘Role model firms’ and is part of the culture pillar. However the variable is measured as the number of firms from a country that are present in the top 2000 of firms in the world divided by the total number of domestic firms a country has. This means that it is a measure of two things. This first being the probability of extreme success for firms of a certain country, the second being the availability of role models for entrepreneurs. However this does not match with the cultural aspect of the pillar, as it does not represent the way in which people are feeling connected with those firms. So although it is a measure for potential role models it is not measuring the meaningfulness of them, consequently this variable will as well be tested separately from the pillar.

Finally there are two variables in the economic pillar that need to be addressed: these are ‘Export’ and ‘GDP growth’. Export (as % of GDP) is a complicated variable that is often removed from country analysis for the same reason, the reason being that this variable clusters small countries together. This clustering effect is actually really logical as the internal trading is not labeled export, so smaller countries are more likely to have a higher percentage of export than larger countries. Export will be deleted from the analysis in total. Last but not least is the variable ‘GDP growth’ which will no longer be part of the economy pillar as it will not match with the upcoming cluster analysis due to the five years measurement. GDP growth will be included in the analysis as an independent variable separately from the pillars as it can have big influences on the amount of new entrepreneurs due to its opportunity predicting nature.

The final pillars are displayed in table 5A, showing the other independent variables in the bottom. To make the following analyses paragraphs easier to read the pillars will be presented again in table 5B with the actual variable names.
Table 5A: Final pillars

<table>
<thead>
<tr>
<th>Final pillars</th>
<th>Education &amp; Skills</th>
<th>Economy</th>
<th>Culture</th>
<th>Government</th>
<th>Finance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firm-specific &amp; Sector-specific skills</td>
<td>Wage dispersion</td>
<td>Power distribution</td>
<td>Income tax</td>
<td>Development capital: • Loans</td>
<td></td>
</tr>
<tr>
<td>Practical education</td>
<td>Available sectors: • Public sector size</td>
<td>Social capital</td>
<td>Bureaucracy inefficiency</td>
<td>Seed capital: • Venture capital</td>
<td></td>
</tr>
<tr>
<td>Education level</td>
<td>GDP per capita</td>
<td>Gender roles</td>
<td>Intellectual property rights (IPR)</td>
<td>Seed capital: • Subsidies</td>
<td></td>
</tr>
<tr>
<td>• Size tertiary education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Quality tertiary education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R&amp;D expenditure</td>
<td>Uncertainty avoidance</td>
<td>Corruption</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Future orientation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other independent variables

| GDP growth | Local Networks | Role model firms |

Table 5B: Final pillars using variables

<table>
<thead>
<tr>
<th>Final pillars (variable overview)</th>
<th>Education &amp; Skills</th>
<th>Economy</th>
<th>Culture</th>
<th>Government</th>
<th>Finance</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Services sector</td>
<td>Inequality (10% - 10%)</td>
<td>Power distance</td>
<td>Highest marginal income tax rate</td>
<td>Loans</td>
<td></td>
</tr>
<tr>
<td>% Purely practical tertiary education</td>
<td>Public sector (% of GDP)</td>
<td>Individualism</td>
<td>Government effectiveness</td>
<td>Venture capital (VC) availability</td>
<td></td>
</tr>
<tr>
<td>% Tertiary education</td>
<td>GDP (per capita)</td>
<td>Masculinity</td>
<td>Intellectual property rights quality</td>
<td>Subsidies</td>
<td></td>
</tr>
<tr>
<td>Education system quality</td>
<td>R&amp;D expenditure (% of GDP)</td>
<td>Uncertainty avoidance</td>
<td>Corruption index</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Long term orientation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other independent variables

| GDP growth (% of GDP) | Local Networks | Role model firms |

Page | 36
4. Analyses

In the previous chapters the structure of the variables into topic specific pillars has been explained. These pillars were created to get an in depth look into how the countries differ on these core topics to entrepreneurship. With this pillar analysis the two main research questions will be answered, but in order to do so groups have to be formed from each pillar and the data needs to be examined for irregularities. Recall that the two research questions are:

Research question 1:  
*Do varieties of capitalism exist in the context of entrepreneurship?*

Research question 2:  
*Do differences in institutions explain differences in the rate of opportunity-driven entrepreneurship across countries?*

4.1 Descriptive statistics

With every statistical analysis come model assumptions that need to be checked to make sure that the analysis is appropriate for the data that is used. In our study there are two types of analysis which will each answer one of the questions. In other words this means that there are also two sets of model assumptions checks necessary. The first analysis will be to test whether the groups found in the theory of VoC are also present in the pillars of this research. For this there is no real available statistical test which will become clear in the following section in which the process of group creation will be explained in depth. Now this does not mean that there are no assumptions at all as every analysis needs to discuss issues like missing data and outliers.

In the second analysis linear regressions will be used to confirm that the found variables show a significant and strong contribution to the prediction of the dependent variable, in this case entrepreneurship. The mayor issues that need to be addressed for a linear regression are a lot more due to the assumption of linearity and such. The issues that will be discussed for the first analysis will also be relevant for the second analysis. Therefore these assumption checks are split in two. The assumptions that will be addressed are listed below showing which will be discussed in this chapter and which will be discussed after the first analysis.

For both research question 1 and research question 2:

1. The number of cases
2. Accuracy of the data
3. Missing data
4. Outliers
5. Multicollinearity

Only for research question 2:

6. Normality of the data
7. Linearity
8. Homoscedasticity
4.1.1 The number of cases
When doing country comparative research a common problem is the number of cases that are available. In the case of for instance consumer research it is often just a matter of time and budget how many cases you can get, but for country comparative research there is of course a maximum of countries that you can possible get. Furthermore the number of countries that actually provide decent information sources or participate in yearly measurements is a lot lower. In this research the number of countries of which the dependent variable is available is limited to 53. However as noted before the total number of countries for which most of the independent variables were available too were 46 countries, so this means that 7 countries were discarded from the analysis due to a lack of data which will be addressed in detail in the subsection ‘Missing data’. The implications of this somewhat low sample size will mostly be the precision of the statistical analysis which will be discussed in the second analysis and the discussion. For the first analysis it will not have such a negative impact as most of the countries of the VoC analysis are still within the remaining countries.

4.1.2 Accuracy of the data
In the previous chapter the data and its’ sources were already discussed, concerning the accuracy by which the data was collected. The only type of accuracy that needs to be addressed is coding and other numerical problems. However no exceptional accuracy problems were found.

4.1.3 Missing data
There were quite a few missing values in the data. Some countries had so many missing data that were excluded from the analysis. The countries are the Dominican Republic, Iceland, Kazakhstan, Puerto Rico, Serbia, the United Arab Emirates and Jamaica. Note here that Puerto Rico is not officially a separate country but is often considered so in several databases. Besides these seven countries most other countries have complete coverage in the data, and if they do not it are usually only one or two data points.

4.1.4 Outliers
Outliers are quite rare in the data. Actually there are only two values that are so distant from the rest that they are considered an outlier. They are both in the export variable: the countries Hong Kong and Singapore both have export values above 200%, while the two closest followers are Malaysia and Belgium with respectively 110% and 83%. The reason that Hong Kong and Singapore are above 200% is not hard to find as these are the only two ‘countries’ that only exist out of one city. This makes it very clear that anything that goes out of the city is in fact export. Besides this they are both very big harbor cities which connect the sea trade with the main land, making them extremely big trading hubs. Please note that Hong Kong is in fact not a country of its own but a state of the People’s Republic of China. Nevertheless it is considered as a separate country in almost every database just like the earlier mentioned Puerto Rico.

4.1.5 Multicollinearity/correlations
For the first analysis it is important to understand the correlations between the variables within each pillar, because high correlation might reduce the explanatory value of the pillar. If for instance all variables correlate the formation of groups is purely one dimensional, meaning that this does not add any value in a regression with respect to using a normal variable. However, it is possible to see which countries are more close together in this single dimension, which can still be used to see if the same groups show up in this analysis as in VoC. In the case that there is high correlation this is
strangely enough only a ‘good thing’ as this means that the formed groups of countries are really different in all the aspects, thus creating a better distinction between groups.

Each pillar will now be discussed in terms of correlation. For each of these there will be a corresponding correlation matrix to be found in appendix 3 to 7. The threshold that will be applied will be 0.700. The correlation matrix for the education and skills pillar showed only low correlations as the highest correlation factor was only 0.504. Also the correlations for the economy pillar are below the 0.700 although GDP per capita and R&D expenditure have a correlation of 0.654. The same type of logic can be applied to the financial pillar as again there are correlations that are fairly high but not higher than the 0.700 threshold. Still this cluster should show clusters of country on a capital availability basis.

The culture and government pillars however have values that exceed the 0.700 threshold. In the culture pillar there is only one, namely the relation between the ‘power distance’ and ‘individualism’ variable. When putting these together in a regression this could have misleading results but for the group creation this is not a problem as this only means that countries differ in the same way on these two aspects of culture. The government pillar shows a very high number of high correlations, except for the variable ‘Highest marginal income tax rate’ all variables correlate with at least 0.922. Now as mentioned earlier the highest marginal income tax rate variable is not the best representative variable because it does not show the tax rate that will actually hit the entrepreneurs but it shows the tax orientation of governments. This together with the fact that the other three variables (government effectiveness, Intellectual property rights (IPR) quality, corruption index) are so highly correlated creates a bad representation of the cluster. Thus this cluster will now be represented by one of the three correlation variables, and from this three the IPR quality was chosen to be the most important. Please note that this choice is pretty insignificant as the variables correlate above 0.900 and thus are almost the same. The implication for the analysis will be that for the first analysis groups will be based on a single dimension, and for the second analysis only the IPR quality variable will be used to represent this pillar. To make things clear the government pillar will still be tested for the second analysis by the use of univariate regressions. However groups based on a single variable are obsolete when using regression analyses.

Now that the correlation within pillars has been discussed it is important to emphasize that there are also correlation between variables from different pillars. These will however be discussed during the analysis as then the meaning of these correlation will be a lot more clear.

Finally, there were some double measurements of variables for which the value has to be validated. For the inequality variable there was a comparison of the richest 10% of the population with the poorest 10% and a comparison of the 20% of each available, but after correlation analysis it turned out to show the same information. Therefore only inequality (10%-10%) will be included. The same goes for the % of people in the service or industry sector, these were highly negatively correlated therefore only the % service sector will be used. So in the end the double measurements were superfluous.

4.2 Clustering

In this research the squared Euclidean distance method was chosen. It was used to measure the distance between the different countries in an article by Hanusch and Pyka (2006), who stated that this method is frequently used as a distance measure for these kinds of research, and they
themselves use this method in their analysis of Europe and the Lisbon Agenda. This method calculates the distance between the different countries and creates links between them showing which countries are in the same position on a certain domain. Yet again the Euclidean distance method itself can be performed in at least two ways. The first way is using the Hierarchical Clustering (HC) method from the statistical program SPSS and the second option is to calculate the distance manually and to use the Multi Dimensional Scaling (MDS) method from the network analysis program UCINET. In both the HC method as in the MDS method the ‘squared Euclidean distance’ and ‘z-scores’ will be used which will both be explained in more detail. Although these two methods both use the same type of calculations to create groups out of the multiple variables it is not the case that they will both show the same end result. The main reasons why the end results are different are that SPSS and UCINET have a black box in which it is not explained how the calculation of the end result is done and that the representation of the end result is also different. The end representation usually does not matter as it is just a different graph using the same type of input, and although this is partly true it will also make a difference. This difference is due to the fact that it is hard or even impossible for people to understand representations that go beyond two dimensions, so the pictures and tables that are produced by the programs only obtain two dimensions while the input variables are higher than two. This means that the programs have to use some type of ‘force’ to fit the countries in the position that is most optimal in a two dimension picture. If this is still a bit abstract it will most likely become clearer in later portions of this chapter when the actual pictures are presented.

4.2.1 Calculation process

Now let’s have a look at the calculation process of both the HC method and the MDS method, but before we can go into the details of the process of each of the methods the notions ‘squared Euclidean distance’ and ‘z-scores’ will be explained in more detail. Please note that these are used in both HC as MDS but that for the MDS this is calculated manually as in the HC method SPSS calculates this in the black box. The choice of input variables can have a big effect on the formed groups, so to make sure that big variables like GDP per capita are not overshadowing other variables the variables were standardized using Z-scores. This allows for equal importance of the variables independent of the actual values the variable has. The calculation of Z-scores is rather simple and is expressed as following:

\[
Z\text{ - scores} = \frac{X - \mu}{\sigma}
\]

\[X = a\text{ data point of the variable, } \mu = \text{the mean of the variable, } \sigma = \text{the standard deviation of the variable}\]

The squared Euclidean distance can be calculated from these z-scores as the influence of the different variables is now equal. The distance calculation takes the ‘coordinates’ of a country and subtracts the ‘coordinates’ of another country to see the absolute distance between them, coordinates is place between quotes as coordinates are often linked to the two or three dimensional space however in this case countries can have more than three dimensions for each pillar. To give a very simple example the next two pictures show a two dimensional space for which the distance will be calculated.
Now if we use the following formula of the Euclidean distance the calculation of the distance will be simple. (Note that this is not the squared distance!)

\[
\sqrt{(1 - 4)^2 + (5 - 0)^2} = \sqrt{34}
\]

Under the square root the first subtraction is purely based in the horizontal dimension namely the coordinate of country A (1) minus the coordinate of country B (4), while the second subtraction is based on the vertical dimension. So when there are more than three dimensions we are not able to create pictures that are interpretable like this easy two dimension picture but we can still calculate the distance in the same way by adding more of those subtractions for the extra dimensions. Now that we have discussed the way to calculate the distance between two countries by using the Euclidean distance it is important to explain why the distance is squared in this research. The squared distance is a measure that is used by other researchers because the squared measure creates even more distance between countries which are further away while keeping close countries relatively closer. This way the groups will appear much clearer making it easier to determine which countries are arranged the same way for a certain pillar.

### 4.2.2 Accuracy of the HC and MDS methods

As explained both the HC method as the MDS method use both the Z-scores as the squared Euclidean distance. The difference between the two methods is the output: the HC method produces a ‘dendrogram’ and the MDS method creates a 2D scatter plot on which the countries are presented with the appropriate distance between them. A dendrogram is a picture in which all the countries are presented by one line each on the left and as a single line on the right, in between the left and the right side there are numerous point on which two lines merge into one line because the belong together. So the later a line joins the bigger the distance between the countries (groups) as on the right all the countries belong to the same group. In figure 4A is an example which shows that Brazil and Venezuela are most close followed by Argentina, South Africa and in the end Colombia joins this group. Still in the end they all belong to the same group although the distance between the all will be higher than the distance between the original Brazil and Venezuela. The MDS method uses a different approach as it creates a map on which the countries are scattered according to the distance between them, as can be seen in figure 4B.
In the HC method each country is forced to connect to the group were it is closest to, but often there are variables between groups which makes it nearly impossible to judge which of the two groups is the right one. However as this doubt is not displayed in the figure it is very tempting to conclude that certain groups are clearly distinct while they are actually quite close to each other. For the MDS this is the other way around: the method does not ‘create’ the groups for you but it just displays the 2D overview as best as possible, and that is exactly where mistakes can be made as the program has to reduce the initial >2 dimensions to a two dimensional figure it forces the points into two dimensions. Picture yourself a 3D view of a pyramid with a country on each corner. If you then try to create a 2D picture with all the countries on it you will find that this is hardly possible. Because of this reason the creators of UCINET included a stress-factor which shows the amount of ‘force’ the MDS calculation needed to get all countries a decent spot. To counter this problem there is another calculation method that can severely lower this stress-factor and enhance the credibility of the figure. This calculation method is the cosine transformation for the distance matrix, this calculation considers the distance between the two countries but also the distance to others to calculate the distance between the countries. Note that the transformation is thus applied to the distance matrix, the formula is as follows:

\[
\text{Cosine transformation} = \frac{\sum_{i=1}^{N} X_i Y_i}{\sqrt{\sum_{i=1}^{N} X_i X_i \times \sum_{j=1}^{N} Y_j Y_j}}
\]

\(X_i, Y_i = the\ distance\ of\ country\ X, Y\ to\ any\ of\ the\ other\ countries\)

4.2.3 Comparison of the HC and MDS methods

Now in order to make sure that the groups are not based on the sometimes fictional groups of the HC method but at the same time not using slightly distorted groups from the MDS method they will both be used and compared to validate each other. The difference between the two methods has been discussed in quite some detail accept for the difference in the way groups should be extracted from the end products. In the HC method SPSS produces dendrograms that due to their characteristics already have some sort of grouping, even though the actual groups are still to be chosen by the researcher by choosing the appropriate allowed degree of distance. In the MDS methods groups are not determined and are thus entirely subjected to the choice of the researcher.

4.3 Comparing clustering results with the VoC

When we look back on the theory of the varieties of capitalism we have seen that Hall and Soskice defined three big groups of countries who were either belong to Coordinated, Liberal or Mediterranean market economies. For the overview please consult table 2.

In appendix 8 to 12 the dendrograms and MDS maps for each pillar are presented. In each of the MDS maps there are colors that represent the groups of the VoC theory. Each of the pillars will be
discussed in terms of these groups, but also the differences between the two types of results will need to be covered. When comparing the hierarchy of the connection in the dendrogram from the HC method to the MDS map it becomes clear that they have a lot in common but are far from the same. The groups that are most obvious in the dendrogram are often not that obvious on the MDS map, meaning that when using the one separate from the other would result in different sets of countries. For each of the pillars some examples will be shown before discussing the comparison with the VoC groups.

4.3.1 Education and skills pillar
For the education and skills pillar the most important differences between the methods will be the position of essential countries to the VoC theory. When looking at the position of Switzerland and the USA it is very typical to see that Switzerland is part of the group of France, UK, Japan, Israel and Ireland while the USA is part of the group with Scandinavian countries. Although they differ on quite some points it is still very clear that they both do not correspond with the VoC groups. The first notable thing is that Asian, South American and Eastern Europe countries seem to be on the opposite side of the figures than the well known strong economies, which may cause the others to be grouped together due to a lack of development. However, it seems there is enough room in the ‘strong’ or ‘weak’ groups to differentiate from each other. Even in the ‘strong’ area the coordinated and liberal groups do not entirely hold as Australia is in the middle of Scandinavian countries and Belgium and Japan are nowhere near their so called peers. Still there is some resemblance to be found so if the other pillars show a clear distinction this could be sort of an outlier.

4.3.2 Economy pillar
For this pillar there are again a lot of placement differences between the methods. For this actually the placement of Germany is very peculiar, in the HC method Germany is combined with other coordinated market economies while in the MDS map it is clearly closer to the more liberal market economies. Again we see the split between ‘strong’ and ‘weak’ countries which in this case is totally logical as it is the economy pillar. However the coordinated and liberal group is again mixed up in both the figures.

4.3.3 Culture pillar
Here we see a very different picture than in the other two pillars. In both the figures a lot of country groups are very clearly distinct, the liberal market economies are sticking together really nice, as are the Scandinavian countries and the Anglo-Saxon countries. However in terms of coordinated market economies this shows a disturbed view as the group is now split up in two very distinct groups, and what makes it even worse is that the Scandinavian group is connecting sooner with the liberal group than with the Angle-Saxon countries, in both methods.

4.3.4 Government pillar (IPR)
As this pillar was initially disbanded because it only consists of IPR now it is not totally fair to call it a pillar. Nevertheless it is still possible to see if the countries are in the place that VoC found. It is clear that the biggest difference in this picture is again the economic development, and as these figures are only based on one variable the differences between the countries of the same economic development soon become too small. Therefore this is not a useful addition to the analysis. In the following analysis for entrepreneurship the effect of IPR will be used as a separate variable which means that this will not be an issue there.
4.3.5 Finance pillar
In this case Denmark, Switzerland and Japan are leaps away from the other coordinated countries. Furthermore Belgium is closer to the liberal market economies than the coordinated market economies. But even the countries that are connected in the MDS picture are very wide spread, meaning that these groups would not be formed without the knowledge of these groups. So in these figures it is definitely the case that the groups used in VoC will not hold. In this case it is also interesting to look at the difference between the HC and the MDS methods, the thing is that in the MDS map Chile and Germany are very close to each other, while in the HC they are in the two different main groups.

Implication for the next analyses
Before continuing with the entrepreneurial analysis it is good to state that there was little evidence found for VoC. One of the reasons for the lack of evidence might be that the VoC theory is Eurocentric. Although the implication of this will be discussed in the end conclusion this does tell us that there might be a need for continent specific dummies for the next regression analysis. Finally the level of economic development creates a strong separation in the data as in almost every picture a cluster of strong economies and a cluster of weak economies is present.

4.4 Entrepreneurship analyses
In this chapter the focus will be to see the relationships between the independent variables and entrepreneurship. The most common way to test the relation between dependent variable and a set of independent variables in a linear way is to apply a multivariate regression. However, as discussed at length before, the amount of data is limited. In total there are only 46 data points available, which is not enough to conduct a multivariate regression. The opinions about the exact number of necessary cases are rather divided but commonly the range is about 10 to 20 cases per independent variable. For this study this would mean that only 2 to 4 variables are allowed, while knowing that four variables is risky. Keeping this in mind it is practically impossible to measure all the hypotheses in one model. Therefore the design of this analysis will be as following: first all the hypotheses will be tested in univariate regression analyses and after that the groups created in the preceding chapter will be used to shed some more light on the institutional differences with relation to entrepreneurship.

The univariate regressions consist of three regressions: the first is a univariate regression using the entire dataset and the second and third will be focused on weak and strong economies. In the descriptive statistics there was found that the countries seemed to be split between lesser developed countries and countries with high GDP levels. With the second and third univariate regressions the univariate regression on the entire data set can be validated as there is a big chance these development differences have an impact on the relation between the independent variables and entrepreneurship.

After the univariate regression tests the earlier found groups will be used to test the influence of the entire pillars on the data. Each of these pillars exists out of multiple independent variables which now do not need to be in the same model. By using these groups the influences of the pillars can be tested without using a multivariable regression.

First the univariate regressions will be discussed but before going into these regressions the assumption check started before the first analysis needs to be finished.
4.4.1 Assumptions
Before the first analysis the following issues were already discussed: Number of cases, Accuracy of the data, Missing data, Outliers and Multicollinearity. Now before the data can be used to run linear regressions it is important to see whether some other assumptions also hold, these are: Normality of the data, Linearity and finally Homoscedasticity. These issues will be discussed in appendix 2. No particular problems were encountered except for the dependent variable that needed to be transformed. Therefore the new dependent will be the logarithmic function of the Total Entrepreneurial Activity (TEA).

4.4.2 Univariate regression analyses
The univariate analyses will be performed on the independent variables from the pillars, on the other independent variables, the dummies for the VoC country groups and the continent dummies. These regressions will show whether there is any ground to validate the hypotheses on. As discussed just before it will be impossible to check these hypotheses in a multivariate regression. Therefore the conclusions drawn from these univariate regressions need to be interpreted very carefully. To facilitate some type of validation the results of these analyses will be reexamined in the following subchapters.

Some variables were hypothesized to have a U-shaped relation with entrepreneurship so here is the chance to test those hypotheses as well, it involves two variables namely: ‘individualism’ and ‘long term orientation’. In case to get this U-shaped relation the variable is put into the regression twice of which one is squared. By adding the normal independent variable together with the squared version of the independent variable the data can create a U-shaped or inverted U-shaped relation. The only way these two relations can be achieved is if one of the two variables is negative and one in positive. In case the squared version is negative the relation is inverted U-shaped, and in case of a negative normal version the relation is U-shaped.

In table 6 all the univariate regressions are presented, in the first column on the left are the variables followed by the unstandardized coefficient B, standardized coefficient Beta, the significance and totally on the right is the R-square of the model.
Table 6: Results of the univariate regressions

<table>
<thead>
<tr>
<th>Variables</th>
<th>B - unstandardized coefficient</th>
<th>Beta - standardized coefficient</th>
<th>Significance</th>
<th>R-square</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education &amp; Skills pillar</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Services sector</td>
<td>-.020</td>
<td>-.406</td>
<td>.005</td>
<td>.165</td>
</tr>
<tr>
<td>% Purely practical tertiary education</td>
<td>-.002</td>
<td>-.008</td>
<td>.959</td>
<td>.000</td>
</tr>
<tr>
<td>% Tertiary education</td>
<td>-.010</td>
<td>-.360</td>
<td>.016</td>
<td>.129</td>
</tr>
<tr>
<td>Education system quality</td>
<td>-.149</td>
<td>-.274</td>
<td>.065</td>
<td>.075</td>
</tr>
<tr>
<td><strong>Economy pillar</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inequality (10% - 10%)</td>
<td>.631</td>
<td>.309</td>
<td>.037</td>
<td>.095</td>
</tr>
<tr>
<td>Public sector (% of GDP)</td>
<td>-.011</td>
<td>-.175</td>
<td>.257</td>
<td>.031</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>-.305</td>
<td>-.264</td>
<td>.077</td>
<td>.069</td>
</tr>
<tr>
<td>R&amp;D expenditure</td>
<td>-.306</td>
<td>-.511</td>
<td>.000</td>
<td>.261</td>
</tr>
<tr>
<td><strong>Culture pillar</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power distance</td>
<td>.004</td>
<td>.166</td>
<td>.269</td>
<td>.028</td>
</tr>
<tr>
<td>Individualism</td>
<td>-.010</td>
<td>-.396</td>
<td>.006</td>
<td>.157</td>
</tr>
<tr>
<td>Individualism + individualism^2</td>
<td>-.041</td>
<td>-1.668</td>
<td>.023</td>
<td>.218</td>
</tr>
<tr>
<td>Masculinity</td>
<td>.000</td>
<td>.015</td>
<td>.919</td>
<td>.000</td>
</tr>
<tr>
<td>Uncertainty avoidance</td>
<td>.001</td>
<td>.023</td>
<td>.879</td>
<td>.001</td>
</tr>
<tr>
<td>Long term orientation</td>
<td>-.010</td>
<td>-.362</td>
<td>.013</td>
<td>.131</td>
</tr>
<tr>
<td>Long term orientation + long term orientation^2</td>
<td>-.020</td>
<td>-.701</td>
<td>.340</td>
<td>.135</td>
</tr>
<tr>
<td><strong>Government pillar</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPR quality</td>
<td>-.154</td>
<td>-.328</td>
<td>.026</td>
<td>.107</td>
</tr>
<tr>
<td><strong>Finance pillar</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loans</td>
<td>-.160</td>
<td>-.226</td>
<td>.130</td>
<td>.051</td>
</tr>
<tr>
<td>VC availability</td>
<td>-.193</td>
<td>-.193</td>
<td>.211</td>
<td>.037</td>
</tr>
<tr>
<td>Subsidies</td>
<td>-.240</td>
<td>-.217</td>
<td>.157</td>
<td>.047</td>
</tr>
<tr>
<td><strong>Other independent variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local Networks</td>
<td>.017</td>
<td>.297</td>
<td>.045</td>
<td>.088</td>
</tr>
<tr>
<td>GDP growth (% of GDP) [over a period of 5 years]</td>
<td>-.038</td>
<td>.104</td>
<td>.493</td>
<td>.011</td>
</tr>
<tr>
<td>Role model firms</td>
<td>-.139</td>
<td>-.104</td>
<td>.491</td>
<td>.011</td>
</tr>
<tr>
<td>Coordinated dummy</td>
<td>-.187</td>
<td>-.133</td>
<td>.377</td>
<td>.018</td>
</tr>
<tr>
<td>Liberal dummy</td>
<td>-.147</td>
<td>-.079</td>
<td>.601</td>
<td>.006</td>
</tr>
<tr>
<td>Mediterranean dummy</td>
<td>-.070</td>
<td>-.041</td>
<td>.789</td>
<td>.002</td>
</tr>
<tr>
<td>Asia dummy</td>
<td>.336</td>
<td>.231</td>
<td>.123</td>
<td>.053</td>
</tr>
<tr>
<td>Europe dummy</td>
<td>-.386</td>
<td>-.333</td>
<td>.024</td>
<td>.111</td>
</tr>
<tr>
<td>South America dummy</td>
<td>.581</td>
<td>.361</td>
<td>.014</td>
<td>.131</td>
</tr>
</tbody>
</table>
From the table it becomes clear that there are quite a few variables which have a significant relation with entrepreneurship.

In the education and skills pillar most of the variables are significant. Only the information about practical education is not of any help in predicting the levels of entrepreneurship. As can be seen from the R-square values the impact from other variables is, although significant, rather low. The amount of people working in the service sector is the best predictor in this pillar. It is also noteworthy to say that the impact of these variables is negative, while the prediction was that a higher service sector would result in more entrepreneurs due to skill portability and higher education levels and quality should also have been positively influencing entrepreneurship. So in the education and skills pillar effects are present but they have a different impact than expected.

In the economy pillar again three out of four variables are significant, and except for the effect of R&D expenditure the effects are of low explanatory power. The effect of inequality is as predicted as higher levels of inequality indeed result in higher levels of entrepreneurship. However, the signs of the GDP and R&D effects are opposite to what was hypothesized: higher levels of GDP and R&D expenditure were expected to be pair with higher levels of entrepreneurship while evidently the opposite is found here. The effect for the GDP dummy seems very low but one needs to consider here that it is a dummy based on a threshold.

The cultural pillar shows three significant models. The first model uses individualism alone, the second uses individualism together with the square of individualism and the last model uses the variable long term orientation. For both individualism and long term orientation the hypothesis was based on an inverted U-shaped relation with entrepreneurship. For individualism we do see a U-shaped relation but it is not inverted as the original variable is negatively related while the square is positively related. Although also the single variable is significant the R-square value of the U-shaped relation shows a higher value than the single variable. The long term orientation variable was also hypothesized to have this U-shaped relation but in this case only a single relation is found, so the people are short term oriented the more entrepreneurs will be present.

The effects from the government pillar are restricted to the IPR variable. In the analyses IPR quality was found significant and with the correct sign.

In the financial pillar there are no significant variables to be found.

Amongst the other independent variables the local networks variable shows a positive relation like expected. GDP growth and the role model variable do not show any significance. The last two continent dummies are significant, meaning that European countries have fewer entrepreneurs and South American countries have more.

Table 7 presents whether the hypotheses are confirmed or not.
Table 7: Overview of the hypotheses confirmation based on univariate regressions

<table>
<thead>
<tr>
<th>Hypotheses overview</th>
<th>Confirmed?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education and skills pillar</strong></td>
<td></td>
</tr>
<tr>
<td>1. The more employees within a country have firm-specific skills the fewer entrepreneurs in a country.</td>
<td>No*</td>
</tr>
<tr>
<td>2. The more people in a country have a practical education the more entrepreneurs in a country.</td>
<td>No</td>
</tr>
<tr>
<td>3. The higher the education level of people in a country is the more entrepreneurs in a country.</td>
<td>No*</td>
</tr>
<tr>
<td><strong>Economy pillar</strong></td>
<td></td>
</tr>
<tr>
<td>5. The higher the dispersion of wages the more entrepreneurs in a country.</td>
<td>Yes</td>
</tr>
<tr>
<td>7. Countries with more available sectors have more entrepreneurs:</td>
<td></td>
</tr>
<tr>
<td>- The bigger the public sector the fewer entrepreneurs.</td>
<td>No</td>
</tr>
<tr>
<td>8. The higher the GDP level of a country the more entrepreneurs in a country.</td>
<td>No*</td>
</tr>
<tr>
<td>10. The higher the R&amp;D-expenditures in a country the more entrepreneurs in a country.</td>
<td>No*</td>
</tr>
<tr>
<td><strong>Cultural pillar</strong></td>
<td></td>
</tr>
<tr>
<td>11. The more people are concerned about the distribution of power within a country the more entrepreneurs in a country.</td>
<td>No</td>
</tr>
<tr>
<td>12. There is an inverted U-shaped relation between individualism and the number of entrepreneurs in a country.</td>
<td>No</td>
</tr>
<tr>
<td>13. The clearer the gender roles are within a country the fewer entrepreneurs in a country.</td>
<td>No</td>
</tr>
<tr>
<td>14. The higher the uncertainty avoidance within a country the fewer entrepreneurs in a country.</td>
<td>No</td>
</tr>
<tr>
<td>15. There is an inverted U-shaped correlation between future orientation and the number of entrepreneurs in a country.</td>
<td>No</td>
</tr>
<tr>
<td><strong>Government pillar</strong></td>
<td></td>
</tr>
<tr>
<td>18. The more efficient the bureaucracy in a country is the more entrepreneurs in a country.</td>
<td>Not tested</td>
</tr>
<tr>
<td>19. Countries with strong intellectual property rights have fewer entrepreneurs.</td>
<td>Yes</td>
</tr>
<tr>
<td>20. The more corrupt the country is the fewer entrepreneurs in a country.</td>
<td>Not tested</td>
</tr>
<tr>
<td><strong>Finance pillar</strong></td>
<td></td>
</tr>
<tr>
<td>21. Countries with more available development capital have more entrepreneurs:</td>
<td>No</td>
</tr>
<tr>
<td>- The easier an entrepreneur can get a loan the more entrepreneurs.</td>
<td></td>
</tr>
<tr>
<td>22. Countries with more available seed capital have more entrepreneurs:</td>
<td>No</td>
</tr>
<tr>
<td>- The more access to early VC-capital within a country the more entrepreneurs.</td>
<td></td>
</tr>
<tr>
<td>- The more access to subsidies within a country the more entrepreneurs.</td>
<td></td>
</tr>
<tr>
<td><strong>Other hypotheses</strong></td>
<td></td>
</tr>
<tr>
<td>4. The more entrepreneurs are present in the local networks of a country the more entrepreneurs in a country.</td>
<td>Yes</td>
</tr>
<tr>
<td>9. The higher the country’s GDP growth the more entrepreneurs in a country.</td>
<td>No</td>
</tr>
<tr>
<td>16. Countries with more companies in the top 2000 have more entrepreneurs.</td>
<td>No</td>
</tr>
</tbody>
</table>

* Significant effects were found but the sign of the effect was opposite to the hypothesis

When having an indept look at the results table it shows that there are many variables that show an effect that is opposite to the expected effect. Especially effects like corruption and government effectiveness seem very unlikely to have a negative effect on entrepreneurship. Also the GDP dummy variable shows a significant effect, with an unpredicted sign, which means that the earlier found split in the data, has predictive value for the amount of entrepreneurs. The next section will go deeper into this split into ‘strong’ and ‘weak’ economies on a basis of GDP per capita. Before going into these two groups it is important to analyze why the many of the effects found have an unexpected sign. Therefore the correlations between all the significant variables will be examined.
Inequality, individualism, long term orientation and local network did not show any correlation that exceeded or was close to the .700 threshold. Besides the long term orientation variable the other effects are the effects that were expected, notably these are almost all cultural variables. This means that all the variables that showed a significant effect but in the opposite direction than expected showed very high correlations amongst each other. The relations between these variables to the GDP dummy clearly showed the highest correlation values. The variables from the education and skills pillar all showed correlations with the GDP dummy with values just below the .700 threshold, so even though the values are lower they relation is still quite strong if one considers that the GDP variable only has values of 1 and 0. This means that in order to correlate with this variable the numbers of the correlating variable must show a clear distinction between the ‘1’ countries and ‘0’ countries.

The upcoming analyses on ‘strong’ and ‘weak’ economies allows for the possibility to check whether there are big differences between the two groups. Although significance levels are likely to generally be lower than in the total group regressions as the sample size will be twice as small. For the found effects on the entire data set will also be checked whether they are also applicable on ‘strong’ and ‘weak’ economies. Furthermore this analysis allows the effects of the VoC country groups to be stronger available as all the countries in these groups are in fact in the ‘strong’ economies group.

4.4.3 Univariate regression analysis: ‘strong’ and ‘weak’ economies
To create a ‘strong’ and a ‘weak’ group of countries the data had to be split on reasonable terms. The first variable that is a good indicator for economic development is ‘GDP per capita’. Secondly the level of R&D expenditure was considered to be a good indicator of this R&D expenses are only done in case of economic prosperity. In figure 5 these two variables were put into one graph to show which the economic developed countries were. However, note that the labeling of the ‘weak’ countries is so crowded that labels might not be in the best position. From this picture it is clear that less developed economies are all grouped in the lower left corner and are very cohesive, while the more developed economies are far more spread out. Therefore the split will take part just after the cohesive group. The appropriate level of GDP will then be 20,000. This is not only a good level when looking at the spreading but is also happens to be around the middle of the number of countries. The following summation shows the countries sorted from lowest GDP to highest GDP.

**Weak economies:**
INDIA, PHILIPPINES, INDONESIA, CHINA, THAILAND, PERU, COLOMBIA, SOUTH AFRICA, ARGENTINA, MALAYSIA, BRAZIL, URUGUAY, ROMANIA, VENEZUELA, TURKEY, RUSSIA, MEXICO, CHILE, LATVIA, CROATIA, HUNGARY, CZECH REPUBLIC

**Strong economies:**
PORTUGAL, ISRAEL, SLOVENIA, GREECE, HONG KONG, SPAIN, JAPAN, ITALY, SINGAPORE, GERMANY, FRANCE, AUSTRALIA, BELGIUM, CANADA, AUSTRIA, UNITED KINGDOM, FINLAND, UNITED STATES OF AMERICA, NETHERLANDS, SWEDEN, DENMARK, SWITZERLAND, IRELAND, NORWAY
Now that the two groups are created, one can run again the univariate regression on both the groups of countries as reported in tables 8 and 9. It is important to keep in mind that these variables are tested on only half the data set, meaning that they are tested on only a bit more than twenty countries. When the sample size gets so low it is not warranted to draw strong conclusions based on small differences. Therefore, only the differences that stand out are considered important. First the regressions on the ‘weak’ economies will be discussed, and afterwards the ‘strong’ economies will be addressed before concluding on the findings of these analyses.

‘Weak’ economies analyses
The table showing the univariate regression results for the weak economies is shown in table 8. Overall the significance of all variables has dropped as expected, due to the lower sample size. The tertiary education quality, % tertiary education, inequality, intellectual property rights (IPR) and local network variable all showed not significant effects anymore. However for the education variables and the IPR variable it could also be the case that they were found significant for the entire data set because they predicted the split in the data, so now that the data is already split they lose their significance. There are six variables that do show significant predictive value for the number of entrepreneurs for the ‘weak’ economies, but within these there are again high correlations. Due to these high correlations there are essentially two groups of effects present in the data:

- The first effect is based on GDP per capita, Public sector (as % of GDP), Individualism and % Service sector which correlate with a value of which all correlate highly with each other. The correlation matrix is presented in appendix 13. Although in this table not all values are above .700 the values are
still very high, especially considering the small sample size. An increase of the public sector in development countries is often linked to the total development of the country as these public sectors are regularly weak. This is also confirmed by the fact that the public sector variable was not found significant for the total data set and also not for the strong economies set. The growth of the service sector is also characteristic to economic development as a lot more services will become available through new efficiencies. An increase in individualism could be a logical result as well due financial independence of the people. Lastly the increasing GDP per capita is essentially the definition of economic development. So as these processes are so linked to economic development it shows that economic development has a negative impact on entrepreneurship.

The second effect is based on R&D expenditure and Long term orientation which also have a correlation. The correlation can be observed in appendix 14, the table however does not show the correlation with the other significant variables but these values all stay below 0.5. The correlation does show a very logical effect because R&D expenditure is always focused on future gains, which links to the extent in which a long term orientation is important in the country. In the regression results for the entire dataset it became clear that cultural effects were not related to the economic split investigated in this analysis. However here it seems as if this effect shows the growth potential of a country, which is linked to the economic development. Whether or not this effect will have its impact on the growth does not really matter in terms of entrepreneurship as clearly this process already has a negatively influence on entrepreneurship. It could well be the case that this future orientation and higher R&D expenditure favors the bigger projects and firms, which is bad for entrepreneurs.
Table 8: Results for the univariate regressions for weak economies

<table>
<thead>
<tr>
<th>Variables</th>
<th>B - unstandardized coefficient</th>
<th>Beta - standardized coefficient</th>
<th>Significance</th>
<th>R-square</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education &amp; Skills pillar</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Service sector</td>
<td>-.038</td>
<td>-.593</td>
<td>.004</td>
<td>.352</td>
</tr>
<tr>
<td>% Purely practical tertiary education &gt; (Square root)</td>
<td>.058</td>
<td>.173</td>
<td>.441</td>
<td>.030</td>
</tr>
<tr>
<td>% Tertiary education</td>
<td>-.010</td>
<td>-.296</td>
<td>.192</td>
<td>.088</td>
</tr>
<tr>
<td>Education system quality</td>
<td>-.082</td>
<td>-.090</td>
<td>.690</td>
<td>.008</td>
</tr>
<tr>
<td><strong>Economy pillar</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inequality (10% - 10%) &gt; (Logarithmic)</td>
<td>.556</td>
<td>.278</td>
<td>.210</td>
<td>.077</td>
</tr>
<tr>
<td>Public sector (% of GDP)</td>
<td>-.033</td>
<td>-.389</td>
<td>.073</td>
<td>.151</td>
</tr>
<tr>
<td>R&amp;D expenditure</td>
<td>-.835</td>
<td>-.527</td>
<td>.012</td>
<td>.278</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>-.000</td>
<td>-.615</td>
<td>.002</td>
<td>.379</td>
</tr>
<tr>
<td><strong>Culture pillar</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power distance</td>
<td>.007</td>
<td>.183</td>
<td>.416</td>
<td>.033</td>
</tr>
<tr>
<td>Individualism</td>
<td>-.022</td>
<td>-.630</td>
<td>.002</td>
<td>.397</td>
</tr>
<tr>
<td>Individualism + individualism²</td>
<td>-.055</td>
<td>-.1591</td>
<td>.046</td>
<td>.448</td>
</tr>
<tr>
<td>Masculinity</td>
<td>.002</td>
<td>.059</td>
<td>.796</td>
<td>.003</td>
</tr>
<tr>
<td>Uncertainty avoidance</td>
<td>-.006</td>
<td>-.174</td>
<td>.438</td>
<td>.030</td>
</tr>
<tr>
<td>Long term orientation</td>
<td>-.012</td>
<td>-.402</td>
<td>.064</td>
<td>.161</td>
</tr>
<tr>
<td>Long term orientation + long term orientation²</td>
<td>-.044</td>
<td>-.1434</td>
<td>.158</td>
<td>.210</td>
</tr>
<tr>
<td><strong>Government pillar</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPR quality</td>
<td>-.278</td>
<td>-.312</td>
<td>.157</td>
<td>.097</td>
</tr>
<tr>
<td><strong>Finance pillar</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loans</td>
<td>-.210</td>
<td>-.187</td>
<td>.404</td>
<td>.035</td>
</tr>
<tr>
<td>VC availability</td>
<td>-.190</td>
<td>-.134</td>
<td>.551</td>
<td>.018</td>
</tr>
<tr>
<td>Subsidies</td>
<td>-.228</td>
<td>-.153</td>
<td>.498</td>
<td>.023</td>
</tr>
<tr>
<td><strong>Other independent variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local Networks</td>
<td>.015</td>
<td>.278</td>
<td>.210</td>
<td>.077</td>
</tr>
<tr>
<td>GDP growth (% of GDP) [over a period of 5 years]</td>
<td>.038</td>
<td>.100</td>
<td>.658</td>
<td>.010</td>
</tr>
<tr>
<td>Role model firms &gt; (Logarithmic)</td>
<td>-.622</td>
<td>-.349</td>
<td>.112</td>
<td>.122</td>
</tr>
<tr>
<td>Liberal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coordinated</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mediterranean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Country groups not present in sample
Table 9: Results for the univariate regressions for strong economies

<table>
<thead>
<tr>
<th>Variables</th>
<th>B - unstandardized coefficient</th>
<th>Beta - standardized coefficient</th>
<th>Significance</th>
<th>R-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education &amp; Skills pillar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Service sector</td>
<td>.010</td>
<td>.146</td>
<td>.497</td>
<td>.021</td>
</tr>
<tr>
<td>% Purely practical tertiary</td>
<td>-.050</td>
<td>-.226</td>
<td>.311</td>
<td>.051</td>
</tr>
<tr>
<td>education &gt; (Square root)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Tertiary education</td>
<td>-.007</td>
<td>-.208</td>
<td>.342</td>
<td>.043</td>
</tr>
<tr>
<td>Education system quality</td>
<td>-.105</td>
<td>-.193</td>
<td>.365</td>
<td>.037</td>
</tr>
<tr>
<td>Economy pillar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inequality (10% - 10%) &gt; (Logarithmic)</td>
<td>.233</td>
<td>.079</td>
<td>.712</td>
<td>.006</td>
</tr>
<tr>
<td>Public sector (% of GDP)</td>
<td>.012</td>
<td>.233</td>
<td>.296</td>
<td>.054</td>
</tr>
<tr>
<td>R&amp;D expenditure &gt; (Logarithmic)</td>
<td>-.212</td>
<td>-.216</td>
<td>.312</td>
<td>.046</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>.000</td>
<td>.016</td>
<td>.941</td>
<td>.000</td>
</tr>
<tr>
<td>Culture pillar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power distance</td>
<td>-.005</td>
<td>-.182</td>
<td>.394</td>
<td>.033</td>
</tr>
<tr>
<td>Individualism</td>
<td>.000</td>
<td>.017</td>
<td>.937</td>
<td>.000</td>
</tr>
<tr>
<td>Individualism + individualism^2</td>
<td>.014</td>
<td>.608</td>
<td>.628</td>
<td>.011</td>
</tr>
<tr>
<td>Masculinity</td>
<td>-.001</td>
<td>-.068</td>
<td>.751</td>
<td>.005</td>
</tr>
<tr>
<td>Uncertainty avoidance</td>
<td>.002</td>
<td>.104</td>
<td>.628</td>
<td>.011</td>
</tr>
<tr>
<td>Long term orientation</td>
<td>-.006</td>
<td>-.239</td>
<td>.260</td>
<td>.057</td>
</tr>
<tr>
<td>Long term orientation + long term orientation^2</td>
<td>.035</td>
<td>1.433</td>
<td>.241</td>
<td>.141</td>
</tr>
<tr>
<td>Government pillar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPR quality</td>
<td>-.043</td>
<td>-.061</td>
<td>.778</td>
<td>.004</td>
</tr>
<tr>
<td>Finance pillar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loans</td>
<td>.007</td>
<td>.011</td>
<td>.959</td>
<td>.000</td>
</tr>
<tr>
<td>VC availability</td>
<td>-.032</td>
<td>-.038</td>
<td>.868</td>
<td>.001</td>
</tr>
<tr>
<td>Subsidies</td>
<td>.040</td>
<td>.033</td>
<td>.884</td>
<td>.001</td>
</tr>
<tr>
<td>Other independent variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local Networks</td>
<td>.012</td>
<td>.199</td>
<td>.350</td>
<td>.040</td>
</tr>
<tr>
<td>GDP growth (% of GDP) [over a period of 5 years]</td>
<td>-.082</td>
<td>-.205</td>
<td>.337</td>
<td>.042</td>
</tr>
<tr>
<td>Role model firms &gt; (Logarithmic)</td>
<td>.564</td>
<td>.463</td>
<td>.023</td>
<td>.214</td>
</tr>
<tr>
<td>Liberal</td>
<td>.019</td>
<td>.016</td>
<td>.941</td>
<td>.000</td>
</tr>
<tr>
<td>Coordinated</td>
<td>-.001</td>
<td>-.001</td>
<td>.998</td>
<td>.000</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>.247</td>
<td>.211</td>
<td>.321</td>
<td>.045</td>
</tr>
</tbody>
</table>
‘Strong’ economies analyses

The table showing the univariate regression results for the strong economies is shown in table 9. For the strong economies group the significance levels of all the variables have dropped a lot. None of the before found effects are present for the strong economies. Only one new effect was found, which was the variable ‘role model firms’. The effect is positive meaning that the relative amount of companies that belong to the biggest in world will result in more entrepreneurs. The hypothesis revolving around this phenomenon is linked to the role model function these firms can have for new entrepreneurs. However as stated before it is very hard to measure this process without measuring other phenomena, in this case the number of strong firms is probably linked to the total business strength of the country, although it is impossible to make this assumption as the measurement only concerns the 2000 best firms of the entire world and not just the countries used in this analysis. Lastly the VoC groups were also included in this analysis, but they turned out to have no explanatory value at all for entrepreneurship.

Implications for the earlier found results

Now that both the sets of analyses have been discussed it is important to link this with the earlier found effects. Overall the effects seem to matter mostly for weak economies as for strong economies all effects are obsolete. Mostly we have seen that for weak economies economic development hampers entrepreneurial activity whereas this is not a problem for countries which belong to stronger group. The same was found for growth potential as Long term orientation and R&D expenditure have shown a negative impact on entrepreneurship for weak economies. The variables that showed a different sign in the total analysis then was hypothesized actually showed the process of economic development for development countries. This also means that the found results on the hypotheses may not present the reality as it is and thus need to be interpreted with caution. However it is impossible to adjust these on the level of ‘weak’ and ‘strong’ economies purely because the sample size will not allow it.

The next subchapter will go more in depth on the effects of entire pillars on entrepreneurship to see if there are other effects that are present without being dominated by the economic split.
4.4.4 Regression with Hierarchical Clustering (HC) dummies

The with the HC method created dendrograms have been discussed before in comparison to the MDS method and to validate the VoC groups. In order to include the results from the dendrograms into a regression analysis it is necessary to create groups in form of dummies. Each group will then receive the value ‘1’ whereas non-group members will receive the value ‘0’, by doing this the regression analysis can correct the regression for this particular group. So let’s say there is a certain linear relation for entrepreneurship but for countries within a certain culture group the level of entrepreneurship is higher, this could be corrected by the use of these dummies. This essentially is the idea of the theory of the varieties of capitalism: each type of economy has its own institutional logic and therefore also its own level of entrepreneurship. So even though the groups found in this analysis do not match the groups found in the work of Hall and Soskice the general idea of groups of countries working in a certain way is still applicable. This analysis will be applied to the entire data set as the groups created from economic pillars will take the economic split into account. This will allow for other groups to be created formed from countries from the ‘weak’ economies group and the ‘strong’ economies group. First the dummy creation process will be explained in detail. After this is done the formed dummies will be applied to the whole data set by using linear regression.

**Dummy creation**

Now in order to create groups from the dendrograms the decision had to be made on how the countries will be split from each other. Essentially there are two options on how this can be done: the first is to choose a set amount of groups that have to be extracted from the dendrograms and the second method is to choose a set distance measure. The set amount of distance makes sure that the countries are only allowed to be part of a group if they are closer than the distance chosen. By using this method you make sure that there are no groups in which the countries are a lot further away from each other than within other groups. So this means that any connection between two countries or a country and a group has a maximum distance. Although the first method is more commonly used the second method seems to be the better way to do it as the groups will be constructed in an equal way. By virtually drawing a vertical line in the dendrograms each group will be formed solely from connections that do not pass the drawn line.

When looking at the dendrograms it seemed possible to split the countries into two big groups but when considering the possibility to split the data into more groups it soon became clear that the groups would then be constructed on entirely different distance levels. This means that the comparison of these groups would be unfair because the proximity of the countries within the groups would be far from the same. Therefore the decision was made to apply both of the types of clustering, the first method to create two big clusters for each pillar and the second method to create smaller groups on the basis of a specified distance.

For the culture pillar it is clearly the case that there is some meaning in smaller groups of this pillar. These groups are corresponding a lot with the earlier discussed Scandinavian and Anglo-Saxon groups. When comparing the distance level of these groups with the other dendrograms it was decided that the most appropriate distance level to create groups was 8. This distance of 8 can be applied to the dendrograms in appendix 8 to 12. However, the second method has as a downside that pillars will not have the same amount of groups and the groups will most likely be of very different sizes across pillars.
Outliers

Then splitting up a large set of countries into groups it is often the case that certain countries do not completely belong to a certain group. These countries or very small groups of countries that are split off do not represent a meaningful group. To make sure that there are not a lot of small groups the limitation of group numbers was set to four and more, meaning that smaller groups will also become missing values. The outliers will be treated like normal outliers and will therefore not be taken into the analysis, it is essential that with the ‘0/1’ dummies these will be marked as missing value because otherwise they will be part of the ‘0’ group. Now that these issues have been addressed the found clusters can be discussed and afterward applied in the regression analysis. In appendix 15 and 16 the groups can be found.

Applying HC dummies in linear regression

In this regression the dummies will not be used on their own like the other variables were tested in the univariate analysis. In this analysis the following variables will be in the regression besides the HC dummies:

1. Intellectual property right quality (IPR)
2. GDP growth over the past 5 years
3. Local Networks
4. Role model firms
5. Continent dummies (Europe, Asia, South America, Others)

IPR is included because it represents the governmental pillar, the continent dummies are included to take care of geographical differences and the other variables are the other independent variables that are not part of any pillar. First the linear regression will be testing the relevance of the split into two groups for each pillar. Second linear regression will be applied using the multiple groups.

In the two-groups regressions the economy dummy, finance dummy and IPR variable could only be added to the regression independently due to multicollinearity. When testing the regressions it was not possible to find a model in which multiple variables were significant. When testing the pillar groups on their own, the dummies of the education and skills pillar, economy pillar and finance pillar were found significant. These univariate results are shown in table 10. Before, in the first univariate regression on variable level, the financial variables were not significant. However in this analysis the financial dummy did turned out to be significant. Yet as can be seen in the table it is clear that this significance is just on the border of the .100 mark, and also the R-square value is low in comparison to the other effects. When comparing these results to the earlier used GDP dummy it is not surprising that the education and skills and economy groups are significant. The correlation between the GDP dummy and the dummies representing the education and skills and economy pillar is namely higher than the .700. Therefore these groups measure mostly the ‘weak’ and ‘strong’ economic split, discussed in the previously analyses.
Table 10: Two-groups HC-dummies univariate regression results

<table>
<thead>
<tr>
<th>Variables</th>
<th>B - unstandardized coefficient</th>
<th>Beta - standardized coefficient</th>
<th>Significance</th>
<th>R-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education and Skills dummy</td>
<td>-.465</td>
<td>-.395</td>
<td>.012</td>
<td>.156</td>
</tr>
<tr>
<td>Economy dummy</td>
<td>-.387</td>
<td>-.326</td>
<td>.033</td>
<td>.106</td>
</tr>
<tr>
<td>Finance dummy</td>
<td>.306</td>
<td>.256</td>
<td>.094</td>
<td>.065</td>
</tr>
</tbody>
</table>

In the multiple-groups regressions we will find if there are smaller groups for which the meaning is not dominated by the economic development. For the multiple groups the problem of correlation was also present. Some groups were correlating with the IPR variable, and some with the Asia dummy. Therefore the groups and variables were not used at the same time. For these regressions there were also no models found that used groups from multiple pillars. In the table 11 the regression results are shown using only groups from the same pillar together. The only significant models were bivariate regression models. The first model uses two groups from the education pillar and the second model uses two groups from the economy pillar. The countries in these groups were the following:

For the education and skills pillar the groups contain the following countries:
Group 3: INDIA, INDONESIA, MALAYSIA, PHILIPPINES, THAILAND
Group 5: CHILE, CHINA, PERU, VENEZUELA

For the economy pillar the groups contain the following countries:
Group 2: CHILE, CHINA, INDIA, INDONESIA, LATVIA, MALAYSIA, MEXICO, PERU, PHILIPPINES, ROMANIA, RUSSIA, THAILAND, TURKEY, URUGUAY
Group 4: ARGENTINA, BRAZIL, COLOMBIA, SOUTH AFRICA, VENEZUELA

Table 11: Multiple groups HC-dummies bivariate regression results

<table>
<thead>
<tr>
<th>Variables</th>
<th>B - unstandardized coefficient</th>
<th>Beta - standardized coefficient</th>
<th>Significance</th>
<th>R-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education and Skills pillar: Group 3</td>
<td>.607</td>
<td>.318</td>
<td>.030</td>
<td>.232</td>
</tr>
<tr>
<td>Group 5</td>
<td>.693</td>
<td>.402</td>
<td>.007</td>
<td></td>
</tr>
<tr>
<td>Economy pillar: Group 2</td>
<td>.331</td>
<td>.263</td>
<td>.088</td>
<td>.127</td>
</tr>
<tr>
<td>Group 4</td>
<td>.578</td>
<td>.313</td>
<td>.044</td>
<td></td>
</tr>
</tbody>
</table>

A large proportion of the Asian countries are represented by the two groups from the education and skills pillar. The only countries that are not in these groups are Hong Kong, Singapore and Japan. These three are also the countries that do not really belong to be amongst the others as Japan is clearly distinct due to is high growth of the past and its current industrial arrangement, and Singapore and Hong Kong are basically two enormous cities which therefore do not have the properties of a country. What is very important to see is that these groups are indeed significant whereas the dummy variable for the continent Asia, used in the first univariate regression on variable level, was not significant. Also there are some countries from South America in this cluster which are also positively linked to entrepreneurship, as we also have seen with the continent dummy in the
first univariate regression on variable level. Why these countries are linked up with China and not with other South American countries is apparently based on their educational system.

When looking closer at the groups from the economic pillar it soon becomes clear that the countries in these two groups are very much like the ‘weak economies’ group used in the split analysis. The only countries that are not present are Croatia, Hungary and Czech Republic. So again the effect of the ‘weak’ and ‘strong’ economic split is presented here. Please note that the GDP dummy represented ‘strong’ economies and therefore showed a negative effect while the group effects here are indeed positive.

As is shown here the country groups do not present any information that is not linked to the economic split. Still the groups were created based on the HC method, which showed clear differences with the MDS pictures. Therefore the finding based on these groups will be validated in the following sub-chapter.

Using factor analyses to validate HC group results

In this final sub-chapter of the analysis part the effects found in regressions using the HC dummies need to be validated. In the beginning of the analysis the HC method was compared to the MDS method to see if the visual presentation was the same using different methods. However when comparing the dendrograms with the scatter plots it turned out that there were quite some differences. Using the MDS method to validate the groups here would be most ideal but the MDS method is merely a visualization tool. Thus the results will need to be validated using a different grouping method.

The most appropriate method to reduce the dimensions of each pillar to just one or two will be factor analysis. In order to avoid unnecessary repetition these newly created factors will only be correlated to GDP dummy to see if they yield the same information. For each of the pillars all the variables were used to create new dimensions. The relations of the variables with these dimensions can be examined in the factor analyses results in appendixes 17 to 20. Only for the skills and education pillar and the culture pillar more than one dimension was necessary.

When relating these dimensions to the GDP dummy it turned out that most of these are indeed correlating above the .700 threshold. Only the second factor of the skills and education pillar and the culture pillar are not related. The second factor for the skills and education pillar is based on almost only the practical education variable, which showed to be of no value for entrepreneurship in the first univariate regression. The second factor for the culture variable is however based mostly on the long term orientation variable and the masculinity variable. The long term orientation variable did show a significant contribution to the prediction of entrepreneurship in previous regressions. However as the culture groups did not show up in the HD dummy analysis it is safe to assume that the results using the HC dummies are valid.
4.4.5 Final remarks on the analyses

Before continuing to the conclusion and discussion this subchapter will provide a final overview of the analyses and summarize the findings. In the first analysis quite some significant effects were found. Strangely enough a lot of them had a sign that opposed the hypotheses, but a few hypotheses were found to be correct. The effects of inequality and local networks were found to be correctly hypothesized but had only little explanatory value. Other variables that were found significant but showed a wrong sign all correlated to the GDP dummy splitting the data into ‘weak’ and ‘strong’ economies.

In the ‘weak’ economies analyses some of the significant effects disappeared and some remained. The effects that disappeared most likely only turned out significant for the entire data set as they yielded similar information as the GDP dummy. The effects that stayed significant can be divided into two types of effects. The first effect is based on effect from the variables: GDP per capita, % service sector, public sector and individualism. They pointed to the effects of economic development for these ‘weak’ economies. The second effect is based on R&D expenditure and long term orientation which show the importance of future gains for these countries. Future gains of relate to the economic development as it is based on investment.

For the ‘strong’ economies almost no variables were found significant. The only variable found significant was the role model firm variable which is very hard to interpret on its own especially when only significant for a portion of the countries while it is measured for the entire world. The fact that no effects were found for ‘strong’ economies shows that for these countries the effect influencing entrepreneurship are likely to be found in other areas than on this institutional level. The next step would be for further research to focus on more local effects.

The next step was to use country groups (and factor analyses on the pillars) to see if there were other effects besides the found economic development issue. However this turned out to not be the case.

So now that it is clear that most of the found significant effects are related to the economic development of ‘weak’ economies it is interesting to take a closer look at this issue before concluding about it. To investigate this even further the ‘weak’ economies are presented in a scatter plot with the dependent variable TEA (not the LN as this is not a regression) on the Y axis and GDP per capita on the X axis, shown in figure 6. From this picture it soon becomes clear that the higher the GDP level gets the lower the entrepreneurial activity gets. For the ‘strong’ economies this graph can be viewed in figure 8, in this picture there is however not any relation between GDP and TEA visible.

Although the ‘weak’ economies graph did show the negative relation between GDP and TEA for these countries there are a few countries that have a low GDP per capita but don’t have a very high TEA level. The countries in question are primarily India, China and Peru. Especially India and China are currently well known for their extreme growth which could be the reason for this discrepancy. In figure 7 the GDP growth of the past five years starting from 2007 is plotted versus the GDP per capita, from this graph it is very clear that China and India are indeed the big growers. For further research for ‘weak’ economies this could be a good starting point. Regrettably it is impossible to run a regression on the ‘weak’ economies with both GDP per capita as GDP growth in the analysis due to the small sample size.
Figure 6: Scatter plot showing the distribution of countries between the TEA and the GDP per capita for weak economies

Figure 7: Scatter plot showing the distribution of countries between the GDP growth of the past 5 years counting from 2007 and the GDP per capita for weak economies

Figure 8: Scatter plot showing the distribution of countries between the TEA and the GDP per capita for strong economies
5. Conclusion and discussion

In this study we asked the question whether institutional differences between countries can explain differences in the rate of entrepreneurship in countries. From the literature study for this paper the most important institutional pillars for entrepreneurship were established. For each of the entrepreneurship pillars groups were formed to cluster countries together that have the same type of properties. The first analysis addressed the resemblance of the Varieties of Capitalism (VoC) groups with the groups found in this data while the other analyses analyzed whether different institutional groups indeed have different rates of entrepreneurship. We focused on opportunity-driven entrepreneurship and did not take into account necessity-driven entrepreneurship as the latter is unlikely to contribute to growth and welfare.

Varieties of capitalism

The resemblance of the VoC groups and the groups created from the entrepreneurship pillar was clearly not present. Only on cultural dimensions deemed relevant for entrepreneurship, the institutional groups resembled those found by Hall and Soskice (2001) in the context of innovation. The first thing is to address the fact that the theory of the VOC only describes countries with strong economies and is mostly Eurocentric. In the comparison the weaker economies were almost in all graphs separated from the stronger economies. Besides this problem also the country groups for liberal and coordinated market economies were not far as clear as the theory states. Clearly the type of market economy does not necessarily mean that the institutional arrangement in completely the same as well. Part of the discrepancies between the VoC groups and the groups found in our study can most likely be found in the way groups are created as in this research two methods were used and the results showed differences which were big enough to allow for different groups based on each of the methods. However, the general conclusion about the grouping differences will be that the differences between the strong economies are not big enough for countries to be labeled liberal or coordinated in any given category. This practically means that the results found with the VOC groups for innovation do not necessarily work for other topics as well. To end with a final remark about Mediterranean market economies: this group was really not even near being a unity, so even though this comment is constructed from entrepreneurship data it is well advised to at least remain skeptical about the cohesiveness of this group.

Entrepreneurship

The connection of the groups representing the important pillars of entrepreneurship and entrepreneurship activity was not as straightforward as was hypothesized beforehand. First, the lack of financial effects and the cultural effects can be linked to the aggregation level. Finance of a new firm is often very dependent of the local situation, meaning that country level indicators may not have been the best option for these effects. The same goes for cultural effects as cultural effects may be greatly determined by local effects. Although all people from a certain country have an average long term orientation it could well be the case that people in that country who are poor think really different about the future than the rich. Furthermore, regional differences are likely to exist, especially in larger countries.

The hypothesized effects of the independent variables were not found in the analyses in most cases. Further analysis showed that the effects of independent variables on entrepreneurship depended in quite many instances on the level of economic development of countries. The influence of the education and skills pillar, the government pillar, cultural effects and of course the economy pillar...
were correlating with the economic development of the ‘weak’ economies. This means that the institutional focus of a country is highly relying on the development level, which is closely linked to discussions about minimum requirement of certain services. For weaker economies the levels of government quality and education quality is bound to be lower than the levels of stronger economies. Therefore the inevitable correlation between these pillars and GDP per capita emerges, for this study about entrepreneur this means that the influences of these independent variables are overshadowed by the strong effect of development levels.

A simple solution could have been to dig deeper into the analysis where the strong and weak economies were split into two separate databases. However, the sample size was clearly too small to support such analyses. In the analyses the influence of the economic development was investigated for these two groups as it turned out that the economic development lost its explanatory value for strong economies while it stayed a very strong predictor for weak market economies. The question remains why the level of economic development is negatively linked to entrepreneurship for weak economies while it is not related to strong economies. When thinking about what happened to the strong economies in their process of becoming as strong as they are now, one may draw lessons from history. Although the situation for weaker economies now is very different than in the time of the industrial revolution in Europe, there is one thing that is essentially timeless for all countries who seek capitalism: economies of scale and efficiency. In the work of Mokyr (1990) on technological creativity and economic progress it is described that there are essentially three possible dominant factors of economic growth they are called Schumpeterian growth, Solovian growth and Smithian growth each of them named after the most influential author for their topic. Schumpeterian growth is closely linked to entrepreneurship. As described in the introduction Schumpeter sees the entrepreneur as the change agent that is the cornerstone for economic growth. Solovian growth is linked to growth through capital investment. Smithian growth is focused on economies of scale and efficiency. In the work of Mokyr it turns out that after important technologies start to kick in, Smithian growth lifted the countries production to the next level. For the weaker economies Smithian growth is clearly the way to go as technologies are often available from strong economies. With the rise of large and efficient firms (and state bureaucracies) comes the end of small firms. Talent, then, will be draw to large-scale bureaucracies both in the private and the public sector. This can, for example, explain why rate of entrepreneurship are rather low in India and China that are currently rapidly developing.

**Limitations**

The first data issue is the sample size in this research. As in any other country level research, there was only a limited number of countries available. A bigger sample size is better for the quality of a research but in country research the total pool of data is limited to around a total of 190 countries in the world. As we have seen in the literature overview these are many different indicators that are found to be relevant to entrepreneurship by the researchers’ community. Some of these are very hard to obtain even for developed countries as we have seen in case of the unemployment benefits and income tax variables.

The second data issue is inaccuracy of the data. In this case the data used was on country level and collected from many different countries which make data often less reliable. Even hard data is often not gathered from every person in the country but extrapolated from a (hopefully) representative sample. The used soft data like the data gathered by the World Economic Forum is even more
receptive to inaccuracy as they represent the perception of the people involved in the surveys. The obvious reason to use extrapolated hard data and soft data is because it is often the only data that is available. Although data that is unique is very valuable, it does not in itself validate that actual accuracy of this data. Importantly, the dependent variable is also subject to this problem. The used GEM database measured the difference between opportunity and necessity driven entrepreneurship by asking the entrepreneurs for their motivation. However it remains a fact that a motivation is a very subjective variable. The distinction between opportunity-driven entrepreneurs and necessity-driven entrepreneurs is a notion that makes an entrepreneur one or the other which is not something to be taken lightly. In the actual data this will be split in a very absolute way while in reality both motivations can play a role. If for instance a normal job is available but an entrepreneurial opportunity could give a higher income the situation is clearly opportunity driven. However when the payment of the ‘normal’ job is not sufficient to a decent living standard the opportunity can soon change into a necessity. From this point of view the question would not only be opportunity or necessity driven but more what is the normal standard of living, and is this standard of living equal for all countries and if not how do you determine these standards.

Further research
The two most interesting field for further research would be to look at the connection of the determinants of entrepreneurship on a smaller geographical scale with the same economic performance and to investigate the effects of local variables. In particular, comparing regions allows one to analyze the effect of institutions on entrepreneurship for many more cases. By applying a multi-level method, one can take into account the influences of national institutions as well. For example, comparing European regions in terms of entrepreneurship could provide a better test of the Varieties of Capitalism approach. At the same time, such a study could inform the European Union about crucial variables driving entrepreneurship, which may guide the Commission in designing effective entrepreneurship policies.
6. Bibliography


7. Appendixes

Appendix 1: Data sources

The World Bank
To start with the World Bank, most of the data from the World Bank are based on demographics, which make them very quantitative and therefore they are more based on facts. Although government effectiveness is not really a demographic, therefore it is far more based on opinions than on hard facts. This is measured by the World Bank by compiling thirty different resources based on surveys, commercial business information providers, non-governmental organizations and public sector organizations. They show extensive overviews of how they obtained and compiled the information which is too specific for this report.

The World Economic Forum
The World Economic Forum (WEF) annually creates a wide range of lists based on a survey they have been conducting for over many years, they put all their gathered information together in ‘The Global Competitiveness Report’ of a certain year. They interview an average of 90 executive directors per country to create these lists. Furthermore they work together with multiple parties under which the Harvard business school to create adequate guidelines for the survey. After the data collection they also do multiple checks to make sure that the received data is in fact decent, one of them being the deletion of surveys with a completion below 50% as they see those participants as not focused enough. So overall they do survey research with a high sense of control, but besides this it still is a survey meaning that all found values are opinions based on human perception.

The Global Education Digest & the Human Development Report
The UNESCO Institute of Statistics and their Global Education Digest obtain their data from three different surveys the UIS survey, the UOE survey and the World Education Indicators (WEI) program. The UIS is the statistical institute of the United Nations Educational, Scientific and Cultural Organisation they send their survey to all the member states every year, to get the highest results they adapt the survey every year to the international standards. The UOE survey is a survey that is conducted jointly by the OECD and Eurostat (UOE), the difference with the first survey is that this survey digs a lot deeper resulting in more detailed information. The WEI is a platform for middle-income countries to develop a critical mass of policy-relevant education indicators. The Human Development Report is developed for the United Nations Procurement Division (UNPD) which is again a different branch of the whole United Nations organ. The inequality figures extracted from this report are actually developed by the World Bank but as they don’t do this yearly the UNPD have compiled all the figures into one list, this however means that not all the numbers are from the same year.

The Global Corruption Report
The Global Corruption Report compiles a lot of different resources into one factor, in total twelve different sources are used to create the figures. Both World Bank as WEF figures are in this list, as are many others which will not be discussed here, the report itself is well structured and should resolve any question and further interest concerning methodology and other corruption related information.

Geert Hofstede’s Cultural Data
The cultural data provided by Geert Hofstede’s website is based on his two of his researches and a couple of other studies concerning the ‘long-term orientation’ dimension. Hofstede build up these
dimensions out of his famous research at IBM and his studies on students through his position at an international business school in Switzerland. His research is purely based on surveys as all the dimensions are purely based in the ‘soft’ side of society. Therefore it is very hard to get any kind of factual information about these processes. Although his data is used by many other researchers and it made Hofstede one of the most famous sociologists in the world this does not take away the fact that measuring these dimensions is very hard if not impossible to do perfectly due to the many extern factors.

Forbes

The biggest 2000 firms in the world were extracted from a list made by Forbes, this is actually not an independent scientific organization at all, but they were the only one that provided this kind of information. They do not disclose their way of creating this list, which makes it very hard to judge the way in which it was created. Still although it is somewhat of a non-scientific measurement it does however link perfectly to what it needs to show, as the variable it has to fuel is based on role models. Role models do not necessarily have to be the biggest but they have to be perceived as a company on which a person from that country can be proud and which can function as a role model to many entrepreneurs in that country.
Appendix 2: Descriptive statistics for the linear regressions

Normality

Officially only the residuals need to be normally distributed not the variables themselves, however it often relates, therefore the normality of the variables was checked as a precaution. Using the Kolmogorov-Smirnov test the variables could be tested to be normal distributed. If this was not the case the histogram with a normality plot was the proper indicator for the appropriate transformation. Although most of the variables were normally distributed the most important variable, the dependent, was not. As explained before the dependent is named TEA, which stands for Total Entrepreneurial Activity, which people fit to this group and who don’t is explained in the earlier chapter about data gathering. Using the Kolmogorov-Smirnov test the TEA variable was found to have a 0.001 significance level, meaning that it is not normally distributed.

Now before the decision was made to transform the dependant variable numerous regressions were executed to see if the residuals were also not normally distributed. After checking the normality of these residuals the same conclusion had to be made very time, the dependant variable needs to be transformed. There are numerous ways get the dependent variable normally distributed but the easiest way to see the correct way is to have a look at the plot which can be viewed in figure 12A which shows the normality plot of the dependent variable TEA. From this plot it was decided that the logarithmic function would work in the best way, in this case the natural logarithmic was used. The result can be seen in figure 12B, which is a lot closer to the normality line.

Table 12A: Normality plot showing TEA

Table 12B: Normality plot showing LN (TEA)

The second Kolmogorov-Smimov test indeed showed a significance level that is indeed pasted the required 0.10. As the dependent variable is very important the distributions of the dependent variable before and after the transformation are visible in figure 9 and 10 in this appendix. The implication for this is that the regression test that will be used will look for a link between the logarithmic dependent variable and the independent variable, meaning that the link between the actual dependent variable will be with the exponential function of the independent variables. So mathematically this will look like the following:

Linear regression on this:  \[ \text{LN(TEA 18-64)} = \text{constant} + (\beta \times \text{independent}) \]

Actual effect:  \[ \text{TEA 18-64} = \text{constant} + e^{(\beta \times \text{independent})} \]
Besides the dependent variable there were a few other independent variables that were also not normally distributed. It concerns the following variables: GDP per capita, R&D expenditure, % purely practical tertiary education, Inequality (10% - 10%), and the Role model firm variable. All these variables are part of a pillar except from the role model firm variable. The pillar variables will be transformed solely for the univariate analyses to check their direct impact on entrepreneurship. These normality issues were solved in the same way as the dependent variable so only the results will be discussed. The variables ‘R&D expenditure’, ‘inequality (10% - 10%)’ and ‘role model firm’ were normally distributed when transformed with a logarithmic function and the variable ‘% purely practical tertiary education’ was normally distributed after a square root. GDP per capita will be transformed into a dummy variables grouping weak and strong economies together, a normal transformation was not possible. Furthermore the same split will be used in the second analysis, in this analysis the split of the two groups will also be explained in detail.

**Linearity**

The literature review suggested that the relationship between the independent variables and the dependent had a certain direction, which means that the hypotheses were stated in a way that more of X would result in more entrepreneurs. So in this research the choice was made to go for a linear relation. However, the analysis is not exactly linear due to the fact that the dependent variable is transformed with a logarithmic function, which means the relation will be exponential. Still it is arguable that this research is focused on trying to find out which factors matter for entrepreneurship and in which way, and therefore this is less an issue than when you try to predict a certain value with the highest possible precision which is practically impossible with a intangible subject like entrepreneurship. It must be noted that this limits this research in the way that it will only check the relations in a linear may which excludes any other relations. The one thing that is probably the most important is the notion of thresholds, as certain variables may play a role to some extent in certain situations which makes it very hard to test in a statistical way. For example if income tax will only matter in countries were the standard of living is high enough that it allows for competition between a normal job and entrepreneurship it may well be that income tax is not significant on the total data set, as the effect of income tax in not relevant for countries with a lower standard of living.

**Homoscedasticity**

After each of the different regressions the standardized residuals were plotted versus the standardized predicted values to see if these were randomly distributed. If this was not the case this would mean that there would have been some connection between the two, meaning homoscedasticity. So if homoscedasticity was the case this would mean that the standard deviation of the errors would be dependent on the predictors of the regression, which is the reason why this assumption check can only be preformed after the regression analysis. In figure 8 such a scatter plot is shown as an example, this is the scatter plot for the first significant analysis of the upcoming univariate analysis, in this picture it is very clear that everything is randomly distributed. It turned out that the other regressions showed similar pictures.
Figure 9: Scatter plot showing no homoscedasticity for the first significant regression of the upcoming univariate regressions

Figure 10: Dependent variable before transformation

Figure 11: Dependent variable after transformation
### Appendix 3: Correlation matrix for the Education and skills pillar

<table>
<thead>
<tr>
<th></th>
<th>% Service sector</th>
<th>% purely practical tertiary education</th>
<th>% tertiary education</th>
<th>Education system quality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>% Service sector</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>1</td>
<td>.075</td>
<td>.504 **</td>
<td>.315</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.630</td>
<td>.000</td>
<td>.033</td>
</tr>
<tr>
<td>N</td>
<td>46</td>
<td>44</td>
<td>44</td>
<td>46</td>
</tr>
<tr>
<td><strong>% Purely practical tertiary education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.075</td>
<td>1</td>
<td>.040</td>
<td>-.112</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.630</td>
<td>.800</td>
<td>.468</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>44</td>
<td>44</td>
<td>42</td>
<td>44</td>
</tr>
<tr>
<td><strong>% Tertiary education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.504 **</td>
<td>.040</td>
<td>1</td>
<td>.346</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.800</td>
<td>.021</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>44</td>
<td>42</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td><strong>Education system quality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.315</td>
<td>-.112</td>
<td>.346</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.033</td>
<td>.468</td>
<td>.021</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>46</td>
<td>44</td>
<td>44</td>
<td>46</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).
### Appendix 4: Correlation matrix for the Economy pillar

<table>
<thead>
<tr>
<th></th>
<th>Inequality (10% - 10%)</th>
<th>Public sector (% of GDP)</th>
<th>GDP per capita</th>
<th>R&amp;D expenditure (% of GDP)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inequality (10% - 10%)</strong></td>
<td>Pearson Correlation</td>
<td>-0.340^*</td>
<td>-0.477^^</td>
<td>-0.406^^</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.024</td>
<td>0.004</td>
<td>0.005</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>46</td>
<td>44</td>
<td>46</td>
</tr>
<tr>
<td><strong>Public sector (% of GDP)</strong></td>
<td>Pearson Correlation</td>
<td>-0.340^*</td>
<td>1</td>
<td>0.403^^</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.024</td>
<td>0.007</td>
<td>0.020</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>44</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td><strong>GDP per capita</strong></td>
<td>Pearson Correlation</td>
<td>-0.477^^</td>
<td>0.403^^</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.001</td>
<td>0.007</td>
<td>0.000</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>46</td>
<td>44</td>
<td>46</td>
</tr>
<tr>
<td><strong>R&amp;D expenditure (% of GDP)</strong></td>
<td>Pearson Correlation</td>
<td>-0.406^^</td>
<td>0.348</td>
<td>0.654^^</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.005</td>
<td>0.020</td>
<td>0.000</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>46</td>
<td>44</td>
<td>46</td>
</tr>
</tbody>
</table>

^**. Correlation is significant at the 0.01 level (2-tailed).

^*. Correlation is significant at the 0.05 level (2-tailed).
### Appendix 5: Correlation matrix for the Culture pillar

<table>
<thead>
<tr>
<th></th>
<th>Power distance</th>
<th>Individualism</th>
<th>Masculinity</th>
<th>Uncertainty avoidance</th>
<th>Long Term Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Power distance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>1</td>
<td>-.700**</td>
<td>.058</td>
<td>.179</td>
<td>.046</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.000</td>
<td>.702</td>
<td>.234</td>
<td>.760</td>
</tr>
<tr>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td><strong>Individualism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>-.700**</td>
<td>1</td>
<td>.024</td>
<td>-.221</td>
<td>.149</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.875</td>
<td>.141</td>
<td>.324</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td><strong>Masculinity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.058</td>
<td>.024</td>
<td>1</td>
<td>.045</td>
<td>.076</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.702</td>
<td>.875</td>
<td>.769</td>
<td>.616</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td><strong>Uncertainty Avoidance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.179</td>
<td>-.221</td>
<td>.045</td>
<td>1</td>
<td>-.054</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.234</td>
<td>.141</td>
<td>.769</td>
<td>.720</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td><strong>Long Term Orientation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.046</td>
<td>.149</td>
<td>.076</td>
<td>-.054</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.760</td>
<td>.324</td>
<td>.616</td>
<td>.720</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).
**Appendix 6: Correlation matrix for the Government pillar**

<table>
<thead>
<tr>
<th></th>
<th>Highest Marginal Income tax rate</th>
<th>Government effectiveness</th>
<th>IPR Quality (WEF)</th>
<th>Corruption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest Marginal Income tax rate</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>.408**</td>
<td>.441**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.005</td>
<td>.002</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Government effectiveness</td>
<td>Pearson Correlation</td>
<td>.408**</td>
<td>1</td>
<td>.940**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.005</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>IPR Quality (WEF)</td>
<td>Pearson Correlation</td>
<td>.441**</td>
<td>.940**</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.002</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Corruption</td>
<td>Pearson Correlation</td>
<td>.400**</td>
<td>.956**</td>
<td>.922**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.006</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.05 level (2-tailed).  
**. Correlation is significant at the 0.01 level (2-tailed).

**Appendix 7: Correlation matrix for the Finance pillar**

<table>
<thead>
<tr>
<th></th>
<th>VC availability</th>
<th>Subsidies for the firm</th>
<th>Loans</th>
</tr>
</thead>
<tbody>
<tr>
<td>VC availability</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>.611**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td>Subsidies for the firm</td>
<td>Pearson Correlation</td>
<td>.611**</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td>Loans</td>
<td>Pearson Correlation</td>
<td>.552**</td>
<td>.684**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>44</td>
<td>44</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).  
*. Correlation is significant at the 0.05 level (2-tailed).
Appendix 8: Dendrogram and MDS map for the Education and skills pillar
Appendix 9: Dendrogram and MDS map for the Economy pillar
Appendix 10: Dendrogram and MDS map for the Culture pillar
Appendix 11: Dendrogram and MDS map for the Government pillar (based solely on IRP quality)
Appendix 12: Dendrogram and MDS map for the Finance pillar
Appendix 13: Correlation matrix 1 for ‘weak’ economies

<table>
<thead>
<tr>
<th></th>
<th>GDP per capita</th>
<th>Public sector (% of GDP)</th>
<th>Individualism</th>
<th>% Service sector</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GDP per capita</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>1</td>
<td>.726**</td>
<td>.523*</td>
<td>.530*</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td><strong>Public sector</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(% of GDP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.726**</td>
<td>1</td>
<td>.612**</td>
<td>.586**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.002</td>
<td>.004</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td><strong>Individualism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.523*</td>
<td>.612**</td>
<td>1</td>
<td>.688**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.012</td>
<td>.002</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td><strong>% Service sector</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.530*</td>
<td>.586**</td>
<td>.688**</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.011</td>
<td>.004</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Appendix 14: Correlation matrix 1 for ‘weak’ economies

<table>
<thead>
<tr>
<th></th>
<th>R&amp;D expenditure (% of GDP)</th>
<th>Long term orientation index</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R&amp;D expenditure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(% of GDP)</td>
<td>Pearson Correlation</td>
<td>.651**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.001</td>
</tr>
<tr>
<td>N</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td><strong>Long term orientation index</strong></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.651**</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
Appendix 15: HC grouping results for two groups

<table>
<thead>
<tr>
<th>Education and skills pillar</th>
<th>Economy pillar</th>
<th>Culture pillar</th>
<th>Finance pillar</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cluster 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARGENTINA</td>
<td>AUSTRALIA</td>
<td>ARGENTINA</td>
<td>ARGENTINA</td>
</tr>
<tr>
<td>AUSTRALIA</td>
<td>AUSTRIA</td>
<td>AUSTRIA</td>
<td>BRAZIL</td>
</tr>
<tr>
<td>AUSTRIA</td>
<td>BELGIUM</td>
<td>BELGIUM</td>
<td>CHINA</td>
</tr>
<tr>
<td>CZECH REPUBLIC</td>
<td>CANADA</td>
<td>CHINA</td>
<td>COLOMBIA</td>
</tr>
<tr>
<td>DENMARK</td>
<td>CROATIA</td>
<td>COLOMBIA</td>
<td>CROATIA</td>
</tr>
<tr>
<td>FRANCE</td>
<td>CZECH REPUBLIC</td>
<td>CZECH REPUBLIC</td>
<td>GREECE</td>
</tr>
<tr>
<td>GREECE</td>
<td>DENMARK</td>
<td>FINLAND</td>
<td>HUNGARY</td>
</tr>
<tr>
<td>HUNGARY</td>
<td>FRANCE</td>
<td>FRANCE</td>
<td>INDONESIA</td>
</tr>
<tr>
<td>IRELAND</td>
<td>GERMANY</td>
<td>GERMANY</td>
<td>ITALY</td>
</tr>
<tr>
<td>ISRAEL</td>
<td>GREECE</td>
<td>GREECE</td>
<td>JAPAN</td>
</tr>
<tr>
<td>ITALY</td>
<td>HONG KONG</td>
<td>HONG KONG</td>
<td>LATVIA</td>
</tr>
<tr>
<td>JAPAN</td>
<td>HUNGARY</td>
<td>HUNGARY</td>
<td>MEXICO</td>
</tr>
<tr>
<td>LATVIA</td>
<td>IRELAND</td>
<td>INDIA</td>
<td>PERU</td>
</tr>
<tr>
<td>NETHERLANDS</td>
<td>ITALY</td>
<td>INDONESIA</td>
<td>PHILIPPINES</td>
</tr>
<tr>
<td>NORWAY</td>
<td>NETHERLANDS</td>
<td>ITALY</td>
<td>PORTUGAL</td>
</tr>
<tr>
<td>PORTUGAL</td>
<td>NORWAY</td>
<td>JAPAN</td>
<td>ROMANIA</td>
</tr>
<tr>
<td>ROMANIA</td>
<td>PORTUGAL</td>
<td>MALAYSIA</td>
<td>RUSSIA</td>
</tr>
<tr>
<td>RUSSIA</td>
<td>SINGAPORE</td>
<td>MEXICO</td>
<td>SLOVENIA</td>
</tr>
<tr>
<td>SLOVENIA</td>
<td>SLOVENIA</td>
<td>PERU</td>
<td>SOUTH AFRICA</td>
</tr>
<tr>
<td>SPAIN</td>
<td>SPAIN</td>
<td>PHILIPPINES</td>
<td>SPAIN</td>
</tr>
<tr>
<td>SWEDEN</td>
<td>SWITZERLAND</td>
<td>PORTUGAL</td>
<td>THAILAND</td>
</tr>
<tr>
<td>SWITZERLAND</td>
<td>UNITED KINGDOM</td>
<td>ROMANIA</td>
<td>TURKEY</td>
</tr>
<tr>
<td>UNITED STATES OF AMERICA</td>
<td>UNITED STATES OF AMERICA</td>
<td>SWITZERLAND</td>
<td>URUGUAY</td>
</tr>
<tr>
<td>URUGUAY</td>
<td>VENEZUELA</td>
<td>UNITED STATES OF AMERICA</td>
<td>VENEZUELA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cluster 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRAZIL</td>
<td>ARGENTINA</td>
<td>AUSTRALIA</td>
<td>AUSTRALIA</td>
</tr>
<tr>
<td>CHILE</td>
<td>BRAZIL</td>
<td>CANADA</td>
<td>AUSTRIA</td>
</tr>
<tr>
<td>CHINA</td>
<td>CHILE</td>
<td>BELGIUM</td>
<td>BELGIUM</td>
</tr>
<tr>
<td>COLOMBIA</td>
<td>CHINA</td>
<td>CANADA</td>
<td>BELGIUM</td>
</tr>
<tr>
<td>CROATIA</td>
<td>COLOMBIA</td>
<td>CHINA</td>
<td>BOLIVIA</td>
</tr>
<tr>
<td>INDIA</td>
<td>INDIA</td>
<td>COLOMBIA</td>
<td>CANADA</td>
</tr>
<tr>
<td>INDONESIA</td>
<td>LATVIA</td>
<td>DENMARK</td>
<td>CHILE</td>
</tr>
<tr>
<td>MALAYSIA</td>
<td>MALAYSIA</td>
<td>FINLAND</td>
<td>CHILE</td>
</tr>
<tr>
<td>MEXICO</td>
<td>MEXICO</td>
<td>IRELAND</td>
<td>CHILE</td>
</tr>
<tr>
<td>PERU</td>
<td>PERU</td>
<td>ISRAEL</td>
<td>CHILE</td>
</tr>
<tr>
<td>PHILIPPINES</td>
<td>PHILIPPINES</td>
<td>LATVIA</td>
<td>CHINA</td>
</tr>
<tr>
<td>THAILAND</td>
<td>ROMANIA</td>
<td>MALAYSIA</td>
<td>CMIA</td>
</tr>
<tr>
<td>TURKEY</td>
<td>RUSSIA</td>
<td>MEXICO</td>
<td>COLOMBIA</td>
</tr>
<tr>
<td>VENEZUELA</td>
<td>SOUTH AFRICA</td>
<td>PERU</td>
<td>DENMARK</td>
</tr>
<tr>
<td></td>
<td>THAILAND</td>
<td>PHILIPPINES</td>
<td>FINLAND</td>
</tr>
<tr>
<td></td>
<td>TURKEY</td>
<td>ROMANIA</td>
<td>GERMANY</td>
</tr>
<tr>
<td></td>
<td>URUGUAY</td>
<td>SINGAPORE</td>
<td>HONG KONG</td>
</tr>
<tr>
<td></td>
<td>VENEZUELA</td>
<td>SOUTH AFRICA</td>
<td>INDIA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SWEDEN</td>
<td>IRELAND</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SOUTH AFRICA</td>
<td>ISRAEL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNITED KINGDOM</td>
<td>MALAYSIA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNITED STATES OF AMERICA</td>
<td>NETHERLANDS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNITED STATES OF AMERICA</td>
<td>NORWAY</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNITED STATES OF AMERICA</td>
<td>SWITZERLAND</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNITED STATES OF AMERICA</td>
<td>UNITED KINGDOM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UNITED STATES OF AMERICA</td>
<td>UNITED STATES OF AMERICA</td>
</tr>
</tbody>
</table>

| Outliers                   |                |               |                |
| BELGIUM                    |                |               |                |
| HONG KONG                  |                |               |                |
| ISRAEL                     |                |               |                |

| Missing values             |                |               |                |
| CANADA                     |                | JAPAN         |                |
| GERMANY                    |                | SWEDEN        |                |
| SINGAPORE                  |                |                | FRANCE         |
| SOUTH AFRICA               |                |                | SWEDEN         |
### Appendix 16: HC grouping results for multiple groups

<table>
<thead>
<tr>
<th>Education and skills domain</th>
<th>Economy domain</th>
<th>Culture domain</th>
<th>Finance domain</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cluster 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUSTRALIA</td>
<td>AUSTRIA</td>
<td>ARGENTINA</td>
<td>BRAZIL</td>
</tr>
<tr>
<td>AUSTRIA</td>
<td>BELGIUM</td>
<td>BRAZIL</td>
<td>CHILE</td>
</tr>
<tr>
<td>DENMARK</td>
<td>CROATIA</td>
<td>CHILE</td>
<td>CHINA</td>
</tr>
<tr>
<td>FRANCE</td>
<td>CZECH REPUBLIC</td>
<td>CROATIA</td>
<td>CROATIA</td>
</tr>
<tr>
<td>IRELAND</td>
<td>DENMARK</td>
<td>GREECE</td>
<td>GREECE</td>
</tr>
<tr>
<td>ISRAEL</td>
<td>FINLAND</td>
<td>PERU</td>
<td>INDONESIA</td>
</tr>
<tr>
<td>JAPAN</td>
<td>FRANCE</td>
<td>PORTUGAL</td>
<td>INDONESIA</td>
</tr>
<tr>
<td>NETHERLANDS</td>
<td>GERMANY</td>
<td>ROMANIA</td>
<td>ITALY</td>
</tr>
<tr>
<td>SWEDEN</td>
<td>GREECE</td>
<td>SLOVENIA</td>
<td>JAPAN</td>
</tr>
<tr>
<td>SWITZERLAND</td>
<td>HUNGARY</td>
<td>SPAIN</td>
<td>LATVIA</td>
</tr>
<tr>
<td>UNITED KINGDOM</td>
<td>IRELAND</td>
<td>THAILAND</td>
<td>MEXICO</td>
</tr>
<tr>
<td>UNITED STATES OF AMERICA</td>
<td>ITALY</td>
<td>TURKEY</td>
<td>MALAYSIA</td>
</tr>
<tr>
<td></td>
<td>NETHERLANDS</td>
<td></td>
<td>PHILIPPINES</td>
</tr>
<tr>
<td></td>
<td>NORWAY</td>
<td></td>
<td>ROMANIA</td>
</tr>
<tr>
<td></td>
<td>PORTUGAL</td>
<td></td>
<td>RUSSIA</td>
</tr>
<tr>
<td></td>
<td>SLOVENIA</td>
<td></td>
<td>SLOVENIA</td>
</tr>
<tr>
<td></td>
<td>UNITED KINGDOM</td>
<td></td>
<td>SOUTH AFRICA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SPAIN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>THAILAND</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VENEZUELA</td>
</tr>
<tr>
<td><strong>Cluster 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARGENTINA</td>
<td>CHILE</td>
<td>DENMARK</td>
<td>AUSTRIA</td>
</tr>
<tr>
<td>CZECH REPUBLIC</td>
<td>CHINA</td>
<td>FINLAND</td>
<td>BELGIUM</td>
</tr>
<tr>
<td>HUNGARY</td>
<td>INDIA</td>
<td>LATVIA</td>
<td>CANADA</td>
</tr>
<tr>
<td>ITALY</td>
<td>INDONESIA</td>
<td>MALAYSIA</td>
<td>FINLAND</td>
</tr>
<tr>
<td>LATVIA</td>
<td>LATVIA</td>
<td>MEXICO</td>
<td>GERMANY</td>
</tr>
<tr>
<td>PORTUGAL</td>
<td>MALAYSIA</td>
<td>PERU</td>
<td>IRELAND</td>
</tr>
<tr>
<td>ROMANIA</td>
<td>MEXICO</td>
<td>PHILIPPINES</td>
<td>MALAYSIA</td>
</tr>
<tr>
<td>RUSSIA</td>
<td>PERU</td>
<td>ROMANIA</td>
<td>NETHERLANDS</td>
</tr>
<tr>
<td>SPAIN</td>
<td>PHILIPPINES</td>
<td>RUSSIA</td>
<td>NORWAY</td>
</tr>
<tr>
<td>URUGUAY</td>
<td>ROMANIA</td>
<td>THAILAND</td>
<td>SWEDEN</td>
</tr>
<tr>
<td></td>
<td>RUSSIA</td>
<td>TURKEY</td>
<td>AUSTRIA</td>
</tr>
<tr>
<td></td>
<td>SLOVENIA</td>
<td></td>
<td>BELGIUM</td>
</tr>
<tr>
<td></td>
<td>SOUTH AFRICA</td>
<td></td>
<td>CANADA</td>
</tr>
<tr>
<td></td>
<td>SPAIN</td>
<td></td>
<td>IRELAND</td>
</tr>
<tr>
<td></td>
<td>SWITZERLAND</td>
<td></td>
<td>SOUTH AFRICA</td>
</tr>
<tr>
<td></td>
<td>UNITED KINGDOM</td>
<td></td>
<td>UNITED KINGDOM</td>
</tr>
<tr>
<td></td>
<td>UNITED STATES</td>
<td></td>
<td>UNITED STATES</td>
</tr>
<tr>
<td></td>
<td>OF AMERICA</td>
<td></td>
<td>OF AMERICA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cluster 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHILE</td>
<td>AUSTRALIA</td>
<td>AUSTRALIA</td>
<td>ARGENTINA</td>
</tr>
<tr>
<td>CHINA</td>
<td>CANADA</td>
<td>CANADA</td>
<td>COLOMBIA</td>
</tr>
<tr>
<td>PERU</td>
<td>HONG KONG</td>
<td>IRELAND</td>
<td>PERU</td>
</tr>
<tr>
<td>VENEZUELA</td>
<td>SINGAPORE</td>
<td>SOUTH AFRICA</td>
<td>TURKEY</td>
</tr>
<tr>
<td></td>
<td>SPAIN</td>
<td>UNITED KINGDOM</td>
<td>URUGUAY</td>
</tr>
<tr>
<td></td>
<td>SWITZERLAND</td>
<td>UNITED STATES</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UNITED STATES</td>
<td>OF AMERICA</td>
<td></td>
</tr>
</tbody>
</table>
### Appendix 16 continued:

<table>
<thead>
<tr>
<th>Education and skills domain</th>
<th>Economy domain</th>
<th>Culture domain</th>
<th>Finance domain</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cluster 4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRAZIL</td>
<td>ARGENTINA</td>
<td>CHINA</td>
<td>AUSTRALIA</td>
</tr>
<tr>
<td>COLOMBIA</td>
<td>BRAZIL</td>
<td>HONG KONG</td>
<td>DENMARK</td>
</tr>
<tr>
<td>CROATIA</td>
<td>COLOMBIA</td>
<td>INDIA</td>
<td>HONG KONG</td>
</tr>
<tr>
<td>MEXICO</td>
<td>SOUTH AFRICA</td>
<td>INDONESIA</td>
<td>INDIA</td>
</tr>
<tr>
<td>TURKEY</td>
<td>VENEZUELA</td>
<td>SINGAPORE</td>
<td>ISRAEL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SWITZERLAND</td>
</tr>
<tr>
<td><strong>Cluster 5</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INDIA</td>
<td></td>
<td>AUSTRIA</td>
<td></td>
</tr>
<tr>
<td>INDONESIA</td>
<td></td>
<td>GERMANY</td>
<td></td>
</tr>
<tr>
<td>MALAYSIA</td>
<td></td>
<td>HUNGARY</td>
<td></td>
</tr>
<tr>
<td>PHILIPPINES</td>
<td></td>
<td>ITALY</td>
<td></td>
</tr>
<tr>
<td>THAILAND</td>
<td></td>
<td>SWITZERLAND</td>
<td></td>
</tr>
<tr>
<td><strong>Cluster 6</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BELGIUM</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CZECH REPUBLIC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>FRANCE</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RUSSIA</td>
<td></td>
</tr>
<tr>
<td><strong>Outliers</strong></td>
<td></td>
<td>ISRAEL</td>
<td>COLOMBIA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ISRAEL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>JAPAN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MALAYSIA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MEXICO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PHILIPPINES</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VENEZUELA</td>
</tr>
<tr>
<td><strong>Missing values</strong></td>
<td></td>
<td>SWEDEN</td>
<td>FRANCE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JAPAN</td>
<td>SWEDEN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Appendix 17: Factor analysis for the Education and Skills pillar

<table>
<thead>
<tr>
<th></th>
<th>% Service sector</th>
<th>% Tertiary education</th>
<th>Education system quality</th>
<th>% Purely practical tertiary education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education and skills factor</td>
<td>Component 1</td>
<td>.803</td>
<td>.820</td>
<td>.677</td>
</tr>
<tr>
<td></td>
<td>Component 2</td>
<td>.190</td>
<td>.085</td>
<td>-.372</td>
</tr>
</tbody>
</table>

### Appendix 18: Factor analysis for the Economy pillar

<table>
<thead>
<tr>
<th></th>
<th>Inequality (10%-10%)</th>
<th>Public sector (% of GDP)</th>
<th>GDP per capita</th>
<th>R&amp;D expenditure (% of GDP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economy factor</td>
<td>Component 1</td>
<td>-.704</td>
<td>.661</td>
<td>.853</td>
</tr>
</tbody>
</table>

### Appendix 19: Factor analysis for the Culture pillar

<table>
<thead>
<tr>
<th></th>
<th>Power distance</th>
<th>Individualism</th>
<th>Masculinity</th>
<th>Uncertainty avoidance</th>
<th>Long term orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture factor</td>
<td>Component 1</td>
<td>.876</td>
<td>-.904</td>
<td>.048</td>
<td>.451</td>
</tr>
<tr>
<td>Component 2</td>
<td></td>
<td>.167</td>
<td>.088</td>
<td>.706</td>
<td>.013</td>
</tr>
</tbody>
</table>

### Appendix 20: Factor analysis for the Finance pillar

<table>
<thead>
<tr>
<th></th>
<th>Venture Capital availability</th>
<th>Subsidies for the firm</th>
<th>Loans</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finance factor</td>
<td>Component 1</td>
<td>.830</td>
<td>.891</td>
</tr>
</tbody>
</table>