Knowledge creation and commercialisation:

the role of R&D experience and R&D network position

Master Thesis (45 ECTS)

Science and Innovation Management

Department of Innovation, Environmental and Energy Sciences

Faculty of Geosciences

Utrecht University

4 November 2013

Student
Ingrid van Prooijen (3275256)
ingridvanprooijen@gmail.com
Jacob Catsstraat 31Bis
3521 VE Utrecht
The Netherlands

Supervisors
Dr. Frank van Rijnsoever (Utrecht University)
Dr. Klaas van Alphen (Global CCS Institute)

Second reader
Dr. Gaston Heimeriks (Utrecht University)

Universiteit Utrecht

---This page is intentionally left blank---

Abstract

The aim of this research is to empirically examine the relationships between an organisation's prior research and development (R&D) experience, its position in multiple collaborative R&D networks and its ability to create, and commercialise knowledge. Including the notion that knowledge creation and knowledge commercialisation co-evolve and interact on a system level, we simultaneously analyse these R&D outputs on the level of the organisation.

With insights from organisational learning and social network theory we hypothesised that prior R&D experiences, the number of collaborations and the extent of clustering in R&D networks are related to knowledge creation and knowledge commercialisation. The R&D activities in which organisations can gain experience are measured as participating in publicly (co-)funded research projects, publishing scientific articles and filing patents. Through these R&D experiences organisations are inherently included in a corresponding R&D network.

We test our hypotheses using carbon dioxide (CO_2) capture technology R&D in the process of carbon capture and storage (CCS) as case. The results show that prior R&D experiences positively influence the likelihood of creating and commercialising knowledge. This is especially the case when (i) publication experience is combined with experience with other R&D activities for knowledge creation and (ii) patent experience is combined with experience with other R&D activities for knowledge commercialisation. Cross-relationships between patent experience and knowledge creation and between publishing experience and knowledge commercialisation are not found. Depending on the type of R&D network and the type of R&D output, the number of collaborations is positively or negatively related to the R&D output. The degree of clustering is negatively associated with R&D output. Based on these findings we provide practical implications and suggestions for further research.

Keywords: Knowledge creation, knowledge commercialisation, R&D experience, network position, Carbon Capture and Storage, CO_2 capture

--This page is intentionally left blank--

Table of Contents

---This page is intentionally left blank---

1 Introduction

Research and development (R&D) are important parts for innovation considering innovation relies on novel combinations of knowledge. Therefore, organisations invest in R&D as information source for their innovation process (Cassiman & Veugelers, 2006). Moreover, as innovation is an important determinant of economic development and growth, many governments support innovation by investments in R&D (Hekkert & Negro, 2009; OECD, 2009). There are two R&D outputs which form the basis of innovation; knowledge creation and knowledge commercialisation. Here, knowledge creation refers to the creation of new scientific knowledge through research and knowledge commercialisation refers to potentially marketable knowledge for the development of new technologies or products.

Organisations obtain these R&D outputs by accumulating experience with R&D activities. The combination of lessons from prior experiences with current R&D experiences, results in organisational routines which facilitate R&D output (Becker, 2004; Schulz, 2002). Organisations rely partly on their in-house R&D facilities, but a large part of R&D experience is gained through various forms of collaboration with others (Powell et al., 1996; Tijssen, 2004). Therefore, previous studies researched the influence of an organisation's network position on the R&D output of companies (Ahuja, 2000) and on the R&D output of research organisations (Debackere et al., 1996; Lee et al., 2012). In these studies R&D output is seen as a result of individual or collaborative experience with R&D activities, which is measured by either knowledge creation or commercialisation.

However, scholars also argue that the processes leading to knowledge creation and commercialisation co-evolve and interact (Etzkowitz & Leydesdorff, 2000; Hekkert et al., 2007; Nelson & Winter, 1982). Since this perspective is mainly studied on a systemic level, it asks for a simultaneous analysis of both types of R&D output on the level of the organisation. Consequently, this raises the question if an organisation's experience with R&D activities and its subsequent network position have different effects on knowledge creation than on knowledge commercialisation. To answer this question, we distinguish three main R&D activities in which organisations can gain experience i.e. participating in publically funded research projects, publishing a scientific article and filing a patent. Furthermore, we define R&D networks that correspond to these activities. By taking this approach this study is the first to focus on the influence of different types of R&D experience and network position on an organisation's R&D output in terms of *both* knowledge creation and knowledge commercialisation. The following research question is leading:

How do experiences with different types of R&D activities and R&D network positions influence an organisation's knowledge creation and knowledge commercialisation?

With insights from organisational learning and social network theory, we develop hypotheses for the influence of prior R&D experiences, the number of collaborations and the extent of clustering in R&D networks on knowledge creation and knowledge commercialisation.

Empirically, we study the case of capturing carbon dioxide (CO_2) in the process of carbon capture and storage (CCS). CCS is the process of separating CO_2 from large point sources (such as power plants) and transporting the captured CO_2 to a storage site, where it is finally long-term isolated into underground geological formations (Global CCS Institute, 2012). It is seen as an important option within the portfolio of low-carbon technologies in the mitigation of climate change. However, the deployment of CCS is currently commercially unattractive (SBC Energy Institute, 2012). CO_2 capture is responsible for roughly two-thirds of the costs of a typical CCS installation. To bring down the costs of CO_2 capture public and private parties are involved R&D collaborations all over the globe (van Alphen et al., 2010).

With a combination of social network analysis and multiple regression analysis we demonstrate that experience in multiple R&D activities as defined in our study positively influence both types of R&D output. Moreover, we find that the influence of an organisation's network position differs per type of R&D network, and also per type of R&D output. However, no cross-relationships exist between experience in filing a patent and knowledge creation or between experience in publishing a scientific article and knowledge commercialisation. This indicates that the processes leading to knowledge creation and commercialisation co-evolve to a limited extent on the organisational level.

The outcomes of this research are particularly interesting for R&D programme managers who wish to manage the type of R&D output of their R&D projects. We provide implications regarding what R&D experience to stimulate and which organisation type to include in R&D projects for knowledge creation or knowledge commercialisation.

In the remainder of this thesis, we first explain the conceptual framework of this research. Thereafter we describe the methodology. Finally, we discuss the results and we draw conclusions from this research.

2 Theory and hypotheses

For this research we develop a framework in which R&D output is explained by an organisation's prior R&D experience and R&D network position. The first two sections elaborate on these concepts. Next, we develop our set of hypotheses.

2.1 R&D output

In this study we analyse R&D output by knowledge creation and knowledge commercialisation. We define knowledge creation as the creation of new scientific knowledge through research. We define knowledge commercialisation as potentially marketable knowledge for the development of technologies or products. In innovation theory literature there are contrasting perspectives on the processes that lead to these R&D outputs.

On the one hand, there is the longstanding perspective of the innovation process which is the linear model of innovation. In this model the innovation process starts with an emphasis on knowledge creation, followed by knowledge commercialisation, the further development into practical applications and finally the diffusion of these applications (Godin, 2006). In this perspective the process resulting in created knowledge is generally associated with research organisations, while the process leading to commercialised knowledge is associated with companies.

On the other hand, there is the systemic view, which highlights the interaction amongst actors and within the processes leading to innovation (Etzkowitz & Leydesdorff, 2000; Hekkert et al., 2007). This "co-evolving perspective" acknowledges the role of knowledge creation and knowledge commercialisation for innovation, but emphasizes the importance of networks and the various feedback loops between the processes that lead to R&D output and innovation (Chaminade & Edquist, 2006). This view has led to studies analysing the role of companies in knowledge creation (Cockburn & Henderson, 1998; Zucker et al., 2002) and to studies analysing the role research organisations have in knowledge commercialisation (Etzkowitz & Leydesdorff, 2000). It is found that despite the inability of companies to fully appropriate the returns of the knowledge created, they engage in scientific research and share parts of their knowledge via scientific publications (Polidoro

Jr. & Theeke, 2012). Motivations for companies to publish their scientific research are to maximise visibility of their R&D capabilities, to attract talent, suppliers and partners and to be connected with the scientific community (Cockburn & Henderson, 1998; Gittelman & Kogut, 2003;Tijssen, 2004; Zucker et al., 2002). Moreover, it is found that research organizations become increasingly active with knowledge commercialization. There is a growing amount of technology transfer offices, knowledge valorisation centres and universities increasingly patent their findings (Agrawal & Henderson, 2002; Cohen et al., 2002; Siegel et al., 2003). These trends are mainly supported by the obligation of public research organisations to diversify their sources of finance, the encouragement of governments to commercialise scientific knowledge and changes in regulation such as the Bayh-Dole Act in the U.S.[1] (Cohen et al., 2002; Geuna & Nesta, 2006; Mowery et al., 2001).

In sum, studies taking a co-evolving and systemic perspective on the processes that lead to knowledge creation and knowledge commercialisation have found evidence that both types of R&D output are important for commercial and research organizations. However, until now no efforts are made to analyse factors that influence *both* R&D outputs of these co-evolving processes.

2.2 R&D experience and R&D network position

With perspectives from organisational learning and social network theory we formulate hypotheses regarding an organisation's R&D experience and position in R&D networks and its influence on R&D output. To be able to do this we analyse R&D experience and the corresponding network position that originate from three main types of R&D activity which are explained in this section.

First, we look at the participation in public (co-)funded R&D projects as an activity in which an organisation can gain R&D experience. This is a relevant R&D activity to include because governments often financially contribute to R&D projects stimulating knowledge sharing and collaboration between the participating public and private organisations (OECD, 2009). Moreover, public R&D spending is used as a measure of investment in innovations with societal relevance, such as CCS (OECD, 2009). Also the activities of publishing of a scientific article and filing a patent are included as R&D activities in which organisations can gain experience. These two R&D activities are key indicators of the processes that lead to respectively knowledge creation and knowledge commercialisation (Verbeek et al., 2002). In sum, R&D experience is expressed in the following activities:

- participating in a public (co-)funded R&D project;
- publication of a scientific article;
- filing of a patent.

Through an organisation's prior individual or collective R&D experience with these R&D activities, it is inherently included in an R&D network: a heterogeneous network of organisations collaborating or individually working on R&D (Tijssen, 1998). Following the activities that determine R&D experience, i.e. participating in publically funded research projects, publishing and patenting, an organisations network position is examined in three types of R&D networks:

- funding network, consisting of organisations that received public funding for R&D projects;
- publication network, consisting of organisations that published scientific articles;
- patent network, consisting of organisations that filed patents.

[1] The Bayh-Dole Act changed the U.S. legislation by allowing organisations that received governmental funding for their research to file patents on the results of this research (Mowery et al., 2001).

In the remainder of this chapter, we formulate the hypotheses central in this study. The research model is presented in figure 1.

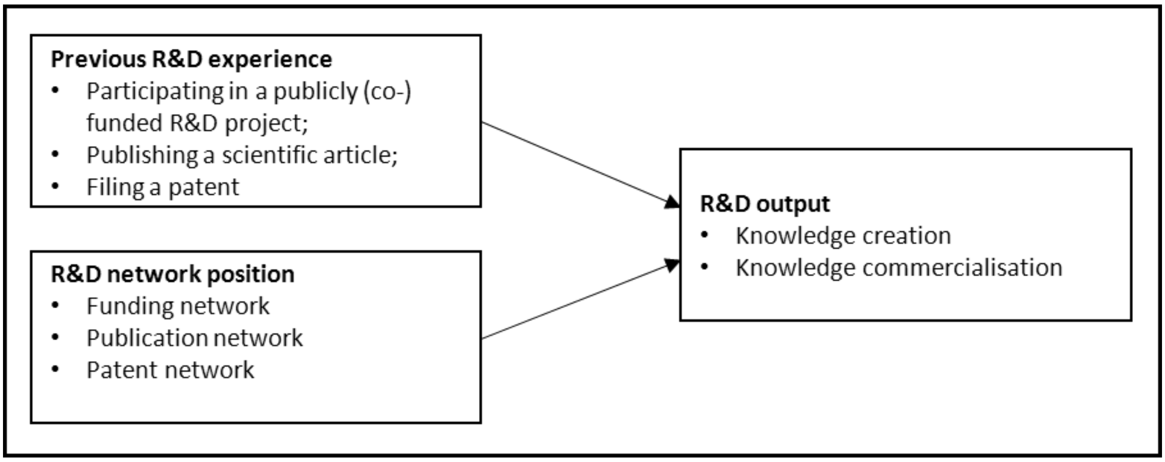

Figure 1: Visualisation of the main concepts and relations of this study

2.3 R&D experience and R&D output

The organisational learning mechanism of "learning by doing" builds on direct prior experiences of the organisation (Huber, 1991; Levitt & March, 1988). To capture lessons from prior experiences it is imperative to execute the experienced activities repeatedly (Levitt & March, 1988). By building on these lessons, successful activities evolve in organisational routines (Nelson & Winter, 1982). With the right routines in place, routines can contribute to stability, trust and the reduction of risk, supporting the chance of successful R&D output (Becker, 2004; Tidd & Bessant, 2011). Drawing this organisational learning perspective on R&D experience, it is implied that organisations with experience in R&D generate a higher R&D output as these organisations have developed R&D facilitating routines from lessons of prior experiences (Powell et al., 1996; Zollo & Winter, 2002).

The following hypotheses capture the effect of prior experience as such as well as accumulated experiences in relation to knowledge creation and knowledge commercialisation. As the activity of publishing a scientific article is a key indicator of knowledge creation we expect a positive relation here. More specifically, it is expected that organisations which have (accumulated) experience with the R&D activity publishing a scientific article to have learned from this experience and are therefore able to create more knowledge in the following years. Similarly, the activity of filing a patent is a key indicator of knowledge commercialisation. Therefore, it is expected that organisations which have (accumulated) experience with the R&D activity commercialising knowledge to accumulate this capability which results in more commercialised knowledge in the following years.

Hypothesis 1a: Organisations with (accumulated) prior experience with the R&D activity of publishing a scientific article create more knowledge in the following years, than organisations without any prior R&D experience.

Hypothesis 1b: Organisations with (accumulated) prior experience with the R&D activity of filing a patent commercialise more knowledge in the following years, than organisations without any prior R&D experience.

As this research includes the notion that the processes leading to knowledge creation and knowledge commercialisation co-evolve, we also formulate hypotheses on the cross-relationship between experience in R&D activities and R&D output. From the co-evolving perspective it can be argued that the R&D activity of publishing a scientific article influences knowledge commercialisation. Correspondingly, the R&D activity of filing a patent can influence knowledge creation. When we combine this insight with organisational learning we expect that an organisation's experience with publishing a scientific article has a positive influence on its ability to commercialise knowledge. Similarly, we expect that experience with the activity filing a patent influences knowledge creation positively. Moreover, the activity of participating in public (co-)funded R&D projects can lead to both knowledge creation and knowledge commercialisation. As the processes leading to these R&D outputs co-evolve and experience accumulates capability, we expect that experience in participating in public (co-)funded R&D projects influences both knowledge creation and knowledge commercialisation. This results in the following hypotheses.

> *Hypothesis 2a: Organisations with prior experience with the R&D activity of filing a patent create more knowledge in the following years, than organisations without any prior R&D experience.*

> *Hypothesis 2b: Organisations with prior experience with the R&D activity of participating in public (co-)funded R&D projects create more knowledge in the following years, than organisations without any prior R&D experience.*

> *Hypothesis 2c: Organisations with prior experience with the R&D activity of publishing a scientific article commercialise more knowledge in the following years, than organisations without any prior R&D experience.*

> *Hypothesis 2d: Organisations with prior experience with the R&D activity of participating in public (co-)funded R&D projects commercialise more knowledge in the following years, than organisations without any prior R&D experience.*

Finally, it is also possible that organisations build experience in more than one of the R&D activities defined in this study. For example, an organisation can participate in a publicly (co-)funded R&D project while also publishing a scientific article or filing a patent. Experience in a combination of R&D activities can influence both knowledge creation and knowledge commercialisation. We argue that gaining experience in multiple R&D activities gives organizations the opportunity of accessing diverse information sources (Powell et al., 1996). These information sources can be other organisations through collaboration. This way gaining experience in multiple R&D activities ties, for example, a research community focussing on knowledge creation together with a research community working on knowledge commercialisation. This broadens the knowledge on state-of-the-art R&D and an organisation's awareness on additional research opportunities (Powell et al., 1996). Moreover, gaining experience in multiple R&D activities can facilitate the co-evolving process between knowledge creation and knowledge commercialisation. All and all, we expect that experience in multiple R&D activities influences R&D output positively. This results in the following hypothesis:

> *Hypothesis 3a: Organisations that have experience with a combination of R&D activities create more knowledge per year than organisations that have experience with one R&D activity or that have no experience at all.*

Hypothesis 3b: Organisations that have experience with a combination of R&D activities commercialise more knowledge than organisations that have experience with one R&D activity or that have no experience at all.

2.4 R&D network position and R&D output

From a social network theory perspective, the organisation's network position is an important determinant of the organisation's R&D output (Owen-Smith & Powell, 2004; Powell et al., 1999). The network position describes an organisation's location in the network which is established by its direct connections as well as by the connections of the organisations around it (Tsai, 2001). This is expressed by the following two factors:

1. the number collaborations an organisation has;
2. the degree organisations cluster together.

The paragraphs below elaborate on the relation between these factors and an organisation's R&D output.

2.4.1 Number of collaborations

An organisation's direct connections with other organisations enable complementation of competences and the generation of economies of scale for investment in R&D (Ahuja, 2000; Gemünden et al., 1996). Accordingly, an organisation's direct connections reduce uncertainty and risks involved in for instance R&D projects (Gargiulo, 2000; Powell, Koput, & Smith-doerr, 2013). Besides the importance of being connected in general, it is also theoretically argued that the number of distinct organisations to which an organisation is connected is relevant. As a connection comes with access to knowledge, more connections increases an organisation's exposure to knowledge sources (Gulati & Gargiulo, 1999; Owen-Smith & Powell, 2004). Organisations which are able to recognize the value of knowledge and subsequently can combine the new information received through their connections, experience a positive effect on creativity and innovation (Cohen & Levinthal, 1990; Uzzi & Spiro, 2005). Furthermore, the positive relation between the number of collaborations and knowledge creation (Debackere et al., 1996) and knowledge commercialisation is also found empirically in areas in which R&D is of great importance to maintain or gain a competitive advantage, such as chemical industry (Ahuja, 2000), the biotechnology industry (Debackere et al., 1996; Shan et al., 1994) and the biomedical industry (McFadyen et al., 2009).

Altogether, linkages and the extent of interaction amongst organisations engaging in R&D are seen as important factors influencing R&D output. However, the question can be raised if the effect on R&D output is the same for all R&D networks. Following the co-evolving perspective, organisations can adapt knowledge accessed from collaboration in patent activities as input for publishing activities or vice versa. Therefore, we expect that the number of collaborations positively influence knowledge creation and knowledge commercialisation for each R&D network. However, to be able to empirically test this expectation we analyse the hypotheses for each R&D network separately.

Hypothesis 4a: Organisations with a higher number of collaborations create more knowledge than organisations with a low number of collaborations.

Hypothesis 4b: Organisations with a higher number of collaborations commercialise more knowledge than organisations with a low number of collaborations.

2.4.2 Degree of clustering

The degree of clustering indicates the extent to which the connections of an organisation are connected with each other. In network theory, the situation in which all organisations are directly connected with each other is called network closure (Giuliani, 2013). In this situation there is a high degree of clustering. In case an organisation is connected with organisations, which are otherwise not or weakly connected, it is said that this organisation is spanning a structural hole (Burt, 2001). Here, an organisation has a low degree of clustering. The relation between these extremes in clustering and R&D output is debated in various scientific publications. A more elaborate explanation of this debate can be found in Appendix A.

On the one hand, scholars argue the benefit of networks with closure i.e. high degree of clustering. According to Coleman (1988) a situation in which all organisations are connected leads to the emergence of effective behavioural norms. As information on deviant behaviour would disseminate quickly and the behaviour would be sanctioned, a closed network constraints organisations to behave different from expectation e.g. to behave opportunistically (Walker et al., 1997). This constraint enables organisations in the network to trust each other (Coleman, 1988), which allows organisations to invest more time and capital in research collaboration without the threat of opportunism. Furthermore, as the connections between organisations intensify, knowledge-sharing routines are developed, facilitating R&D output (Ahuja, 2000; Zollo et al., 2002).

On the other hand, Burt (2001) argues the importance of spanning structural holes for organisations i.e. low degree of clustering. This allows the organisation to function as broker of information between organisations that are not directly connected with each other (Provan et al., 2007). As information is known to circulate within sets of connected organisations before it circulates between sets of organisations, the 'broker' can control the information flows between the sets of organisations. Moreover, it receives the information from both sets of organisations enabling the broker to combine information from the different sources. From this perspective, gaining access to multiple knowledge sources through spanning structural holes i.e. a low degree of clustering, stimulates creativity which increases R&D output (Burt, 2004).

In relation with knowledge creation we follow the perspective of Powell et al. (1996): "*knowledge creation occurs in the context of a community, one that is fluid and evolving rather than bound or static*". This view stresses the relevance of a low degree of clustering since connections of organisations with a lower network density offer more flexibility and opportunities for creativity, than the connections in a network where behaviour becomes more routinized (McFadyen et al., 2009). On the other hand, we argue that the benefit of a high degree of clustering is more likely to hold for knowledge commercialisation than the benefit of a low degree of clustering. As collaboration for knowledge commercialisation includes high levels of secrecy and investments, trust amongst organisations is an important factor for successful collaboration (Mora-Valentin et al., 2004; Powell et al., 1996; Zaheer et al., 1998). Following the co-evolving perspective, we do not expect differences between the R&D networks in relation to R&D output. However, we do test this relation for each R&D network separately. This results in the final hypotheses of this study:

> *Hypothesis 5a: Organisations with a low degree of clustering create more knowledge than organisations with a high degree of clustering.*
>
> *Hypothesis 5b: Organisations with a high degree of clustering commercialise more knowledge than organisations with a low degree of clustering.*

2.5 Control variable: Organisation type

In the relation between prior R&D experience, R&D network position and R&D output, we control for the type of organisation. This can be a research organisation, a multinational company, a national company, or an "other" organisation. Even though knowledge creation and knowledge commercialisation are R&D outputs for organisation active in R&D, the influence of the core activities of an organisation must be taken into account. Therefore, we define research organisations as organisations that have R&D as their primary task. Examples of research organisations are universities and national applied science organisations such the Dutch Organisation for Applied Scientific Research (TNO). Organisations that conduct R&D as input for their innovation process, but that have other core activities which are often profit-related are defined a company. In case a company has offices in multiple countries it is considered a multinational company, otherwise it is defined as a national company. This way we indirectly control for company size. Also public agencies, intermediaries and interest groups are active in R&D. These types of organisations are grouped as "other" organisation.

3 Methodology

3.1 Research case

The hypotheses are tested in the context of the field of Carbon Capture and Storage (CCS). This bundle of technologies is a viable option for reducing greenhouse gas emissions from large scale emitters. Therefore, it is considered a key solution in the mitigation of climate change (Global CCS Institute, 2012; IEA, 2009). As CCS is seen as a technology that benefits the public good, both governments and private parties are greatly investing in its development; the total amount of funds that has been allocated towards CCS R&D and demonstration around the world is estimated at $28 billion (Global CCS Institute 2012).

Within the field of CCS this study focuses on CO_2 capture which is currently the economical bottleneck for the large scale deployment of CCS. Therefore, most of the funding allocated until to date is targeted at R&D for more cost effective CO_2 capture technologies (SBC Energy Institute, 2012). In exact science fields, such as capturing CO_2, knowledge creation and commercialisation are important R&D outputs (Ahuja, 2000; Debackere et al., 1996). More specifically, within CO_2 capture R&D the exploration of more cost effective capture processes resembles the process that leads to knowledge creation. On the other hand, knowledge commercialisation is relevant to be able to develop more cost effective capture processes for commercial purposes in CCS.

Within CO_2 capture technologies several processes can be distinguished. Following the categorisation of the European Patent Office this research studies CO_2 capture by absorption, adsorption, chemical separation, biological separation, membranes, diffusion, rectification and condensation (Veefkind & Hurtado-Albir, 2010). An introduction to the different CO_2 capture technologies can be found in Appendix B. The development of CO_2 capture technologies is studied over the period 2002 till 2010 within North America and Europe. The period of analysis is set between 2002 and 2010, because CO_2 capture R&D activity was limited before 2002 and the CO_2 capture patent data for this study is only reliable up to 2010. North America and Europe form the geographical delineation for this research as these continents are responsible for 85% of the total financial contribution to CCS R&D projects (SBC Energy Institute, 2012).

3.2 Data collection

Data was collected in a dataset for each type of R&D activity. The following information was collected for each dataset:

1. the title of the publicly (co-)funded R&D project, publication or patent;
2. the start year for R&D projects, publication year for publications, and application year for patents;
3. the organisations that participated in the R&D project, published a scientific article or filed a patent.

To gather this information multiple data sources were used. A more detailed description of the data collection and preparation can be found in Appendix C. This appendix also includes an overview of the number of R&D projects, publication, patents and organisations included in this research.

Firstly, data is collected on publicly (co-)funded R&D projects on CO_2 capture and corresponding organisations resulting in the funding dataset. For this dataset information was gathered from the National Energy Technology Laboratory (NETL), Natural Resources Canada, CORDIS and European grant databases of national research programs (CORDIS, 2012; Legg & Campbell, 2006 (NRCan); NETL, 2002-2012). NETL is part of the U.S. Department of Energy national laboratory system and implements research and development programs in the energy field. Natural Resources Canada is the ministry of the government of Canada which is amongst others responsible for energy related matters. CORDIS is the database of the European Union (EU) providing information on all EU funded initiatives. Finally, examples of sources used to retrieve R&D project data (co-)funded by European member states include the IEA Greenhouse Gas R&D programme and the UK EPSRC databases (EPSRC, 2013; IEAGHG, 2013). Together, the funding dataset consists of 253 research projects executed by 430 organisations.

Secondly, Web of Science is used for collecting data on CO_2 capture related articles and corresponding organisations for the publication dataset. Web of Science is an online research platform which provides access to citation databases including multidisciplinary content from "*over 12,000 of the highest impact journals worldwide*" (Web of Science, 2013). The data for the publication dataset was collected using a series of queries that was iteratively established using CCS reports and publications as input. The final list of queries was checked and validated by a number of field experts. A total of 976 articles were included in the dataset. These articles were (co-)authored by 454 organisations.

Thirdly, for the construction of the patent dataset, the PATSTAT database is used. This is the patent statistical database of the European Patent Office including worldwide patents. After limiting the dataset to the CO_2 capture categories given in the PATSTAT database and to the other demarcations of this research, 1379 patents and 446 organisations are included in the patent dataset.

In total 1069 organisations are included in the research. The website of each organisation is visited to determine the type of organisation i.e. research organisation, multinational company, national company or "other" organisation.

3.3 R&D network database construction

From each dataset we constructed a R&D network of collaborating organisations and organisations that worked individually on R&D. In the construction of networks it is difficult to determine for what period connections of organisations are maintained or valuable for organisations. Due to a lack of available data, previous studies did not make a distinction between the moment a connection is formed and the period a connection is maintained (Ahuja, 2000). However, this could result in excluding findings with regards to the network position which are influenced by time. For instance, Soda et al. (2004) found that past network closure rather than current network closure influences the output of organisations in the network.

To include the possible influence of time on the network position in relation to an organisation's R&D output, we composed three different network databases for each R&D network. The difference between the three network databases is the number of years an organisation and its connections are included in the network after the year they participated in a R&D project, published a scientific article or applied for a patent. In other words, the R&D network databases are composed of organisations and their collaborations of a certain year of collaboration (t) combined with organisations and their connections one year before that collaboration year (t-1), two years before that collaboration year (t-1+t-2) or three years before that collaboration year (t-1+t-2+t-3). These databases are named respectively network database 1 Year, 2 Years and 3 Years. An illustrative table of the network database construction can be found in Appendix D.

3.4 Measurement

Knowledge creation is measured by the number of scientific publications per year and knowledge commercialisation is measured by the number of patents per year. Publications are used as proxy for knowledge creation as it reflects the codification of the new scientific knowledge which is created (McFadyen et al., 2009). The number of patents per year is used as a proxy for knowledge commercialisation, because potentially marketable knowledge is generally encoded in patents, a mode of protection to appropriate the benefits from its potential (Partha & David, 1994; Teece, 1986). This is especially the case in exact science fields, such as CO_2 capture R&D (Ahuja, 2000; Debackere et al., 1996).

An organisation's prior R&D experience is measured by the variables; experience, accumulated publication experience and accumulated patent experience. The nominal variable of experience measures if an organisation previously participated in a publicly (co-)funded R&D project, or if it published a scientific article, or if it filed a patent, or if it engaged in a combination of these activities or if it does not has prior R&D experience. The variables of accumulated publication experience and accumulated patent experience measure respectively the cumulative number of publications and the cumulative number patents over the previous years.

The R&D network position of an organisation is determined for each R&D network separately. The number of collaborations is measured by the social network measure of degree centrality: the amount of *direct* connections the organisation has (Wasserman, 1994). The degree of clustering is measured by the local clustering coefficient. This measure is calculated by dividing the number of actual collaborations between the connections of an organisation with the number of possible connections that could exist between them (Watts & Strogatz, 1998). To calculate these social network measures the R&D networks are analysed using the igraph package in the software program R (Csardi & Nepusz, 2006; R Core Team, 2013). The complete R-script of the social network analysis can be found in Appendix E. Furthermore, the descriptive statistics and correlation matrices for the variables in the three network databases can be found in Appendix F.

To remove the problem that an organisation's publication and patent experience and its corresponding network position are determined by the R&D output of the same year, the independent variables are lagged with one year. In other words, if R&D output is measured in year t, the independent variables are measured for year t-1.

3.5 Data analysis

All hypotheses are tested using a generalized linear mixed model with a random intercept dependent on the organisation. The random intercept takes into account that organisations were included in the analysis for multiple years. The analysis is done using the lme4-package in the software program R (Bates et al., 2009; R Core Team, 2013).

As the dependent variables in this study are count data, the hypotheses are tested against a Poisson distribution. However, a common effect of the Poisson distribution is the occurrence of over-dispersion when the variance does not equal the mean. Therefore, a model with Poisson distribution is tested against a model with Poisson distribution which corrects for over-dispersion, by including a second random intercept for the number of observations (Bolker & Stainbrook, 2010). This way the model is translated to a lognormal-Poisson model, which allows both fixed and random effects, functioning as a model with a quasi-Poisson distribution (Elston et al., 2001). The results indicate that the second model gives consistently a better fit than an ordinary Poisson model (knowledge creation $p<0,01$, knowledge commercialisation $p<0,001$[2]). Consequently the second model is selected as final model.

Since this is a longitudinal study, time dummies are included for each year (2002 – 2010). Moreover, we ran the analysis with a subset of the data which included organisations that currently gain R&D experience and organisations that through prior R&D experience are part of a R&D network. This subset is made, because including all organisations for the years without current or prior experience in the analysis would give a false picture of what is happening. However, since all organisations included in the analysis gain R&D experience at some point, this subset emphasizes the moment that an organisation creates knowledge or commercialises knowledge for the first time in our database. To correct for this effect a variable for the first time an organisation publishes a publication or files a patent is added. Finally, the McFadden R-square is calculated based on the log-likelihoods (McFadden, 1973). The complete R-script of the statistical analysis for testing the hypotheses can be found in Appendix G.

4 Results

Table 1 and table 2 show the results of the generalized linear mixed model of respectively knowledge creation and knowledge commercialisation. The estimators in the tables are unstandardized and the tables include the results for the three R&D network databases. The McFadden R-square is good for both models: between 0.43 and 0.50 for knowledge creation and between 0.39 and 0.43 for knowledge commercialisation. The three network databases show essentially similar results, which indicate the robustness of the findings. It also indicates that time has a relatively small effect on the influence of network position on R&D output, which is in contrast to the findings of Zaheer (2004). This can be explained by the time period over which the analysis is done. Zaheer (2004) studied the network position and its influence on R&D output over 7 years while in this research the network position is combined with the network up to 3 years before.

[2] Knowledge creation: network database 1 Year and 2 Years $p<0,001$; 3 Years $p<0,01$.
Knowledge commercialisation: network database 1 Year, 2 Years and 3 Years $p<0,001$.

Table 1: Results of multiple regression analysis with the dependent variable knowledge creation.

*p<0,1, **p<0,05, ***p<0,01

Knowledge Creation	Variable	Specification	Model - 1 Year -	Model - 2 Years -	Model - 3 Years -
Intercept	Fixed		-3,37***	-4,21***	-4,99***
	Variance observations		0,22	0,16	0,10
	Variance organisations		0,08	0,12	0,15
Control variables	Year	2004	0,61**	0,59**	0,60**
		2005	0,60**	0,55**	0,55**
		2006	0,10	0,10	0,04
		2007	0,44*	0,36	0,33
		2008	0,47**	0,37	0,33
		2009	0,59**	0,55**	0,46**
		2010	0,60**	0,55**	0,51**
	First publication		3,05***	3,89***	4,71***
	First patent		-0,10	0,12	0,25
	Organisation type	Research organisation	Ref	Ref	Ref
		Multinational company	-0,74***	-0,71***	-0,70***
		National company	-1,04***	-0,92***	-0,82***
		Other organisation	-0,66*	-0,48	-0,36
Independent variables	Experience	Null	Ref	Ref	Ref
		Gov. Funded R&D	-0,32	-0,20	-0,34*
		Publication	1,92***	2,74***	3,50***
		Patent	-0,82***	-0,32	-0,09
		Publication & Patent	2,07***	3,02***	3,74***
		P. Funded R&D & Publication	2,32***	3,08***	3,79***
		P. Funded R&D & Patent	0,24	0,37	0,19
		P. Funded R&D, Publication & Patent	2,21***	3,17***	3,86***
	Cumulative experience	Cumulative publications	0,07***	0,06***	0,05***
		Cumulative patents	0,01	0,00	0,01
	Number of collaborations	Funding	-0,01*	-0,01**	-0,01**
		Publication	-0,06*	-0,03	-0,02
		Patent	-0,11	0,19	0,13
	Collaboration with Research organisation	Funding	0,46*	0,47*	0,54**
		Publication	0,10	0,08	0,11
		Patent	0,61	0,02	0,05
	Collaboration with Multinational Company	Funding	0,21	0,38*	0,41**
		Publication	0,57***	0,50***	0,53***
		Patent	0,09	-0,05	-0,19
	Collaboration with National Company	Funding	0,45***	0,39**	0,36**
		Publication	0,04	0,11	0,09
		Patent	0,06	-0,30	0,04
	Collaboration with Other organisation	Funding	0,14	0,16	0,17
		Publication	-0,68**	-0,97***	-0,90***
		Patent	0,03	0,09	-0,54
	Degree of clustering	Funding	-0,82***	-0,89***	-0,82***
		Publication	-0,24*	-0,36***	-0,33***
		Patent	0,11	-0,58	-0,26
Model indicators	Number of Obs.		3784	4080	4270
	LogLikelihood		-939,1	-905,6	-867,1
	McFadden R-square		0,43	0,47	0,50

Table 2: Results of multiple regression analysis with the dependent variable knowledge commercialisation.

*p<0,1, **p<0,05, ***p<0,01

Knowledge Commercial.	Variable	Specification	Model - 1 Year -	Model - 2 Years -	Model - 3 Years -
Intercept	Fixed		-2,79***	-3,65***	-4,12***
	Variance observations		0,34	0,31	0,28
	Variance organisations		0,30	0,31	0,33
Control variables	Year	2004	-0,44**	-0,45**	-0,46**
		2005	-0,28	-0,45***	-0,49***
		2006	0,06	-0,05	-0,22
		2007	-0,14	-0,24	-0,36**
		2008	-0,20	-0,28*	-0,37**
		2009	-0,34**	-0,44***	-0,53***
		2010	-1,22***	-1,41***	-1,54***
	First publication		-0,19	0,06	0,19
	First patent		2,93***	3,89***	4,46***
	Organisation type	Research organisation	Ref	Ref	Ref
		Multinational company	0,64***	0,56***	0,50***
		National company	0,03	-0,01	-0,02
		Other organisation	-0,37	-0,30	-0,21
Independent variables	Experience	Null	Ref	Ref	Ref
		Gov. Funded R&D	-1,22***	-0,68**	-0,40
		Publication	-0,65*	-0,26	-0,35
		Patent	1,08***	1,99***	2,49***
		Publication & Patent	1,42***	2,79***	3,27***
		Gov. Funded R&D & Publication	-0,08	0,10	0,11
		Gov. Funded R&D & Patent	1,62***	2,72***	3,40***
		Gov. Funded R&D, Publication & Patent	1,56***	2,96***	3,83***
	Cumulative experience	Cumulative publications	0,02	0,01	0,00
		Cumulative patents	0,05***	0,04***	0,03***
	Number of collaborations	Funding	0,02***	0,02***	0,02***
		Publication	-0,07	-0,02	-0,04
		Patent	-1,31***	-0,35*	-0,23
	Collaboration with Research organisation	Funding	0,13	-0,09	-0,19
		Publication	-0,08	-0,25	-0,13
		Patent	1,70***	0,70**	0,71***
	Collaboration with Multinational Company	Funding	0,12	0,02	-0,08
		Publication	0,18	0,17	0,49
		Patent	1,42***	0,38	0,17
	Collaboration with National Company	Funding	-0,25	-0,23	-0,28
		Publication	0,02	-0,09	-0,08
		Patent	1,66***	0,76**	0,64**
	Collaboration with Other organisation	Funding	-0,59***	-0,76***	-0,79***
		Publication	0,17	-0,10	0,04
		Patent	-12,8	-14,0	-13,8
	Degree of clustering	Funding	-0,08	0,18	0,33
		Publication	-0,17	-0,17	-0,34
		Patent	1,09*	-0,13	-0,63
Model indicators	Number of Obs.		3784	4080	4270
	LogLikelihood		-939,1	-936	-926,2
	McFadden R-square		0,39	0,41	0,43

4.1 R&D Experience and R&D output

The results in table 1 show that there is positive effect of experience with the R&D activity of publishing a scientific article on knowledge creation. Also the accumulated experiences with publishing scientific articles influence knowledge creation positively. These results together, support hypothesis 1a. Experience with the R&D activity of filing a patent shows a positive effect on knowledge commercialisation (table 2). A positive relation is also found between accumulated experiences with filing patents and knowledge commercialisation positively. With these findings hypothesis 1b is supported.

Moreover our findings show that there are no significant relationships found in the relationships between experience with the R&D activities of filing a patent and participating in (co-) funded R&D projects and knowledge creation (table 1). With these results, hypothesis 2a and hypothesis 2b are not supported. In table 2, the results show that the relation between publishing a scientific article and knowledge commercialisation is negative, but not significant. Therefore these results are insufficient to consider hypothesis 2c contradicted. A significant negative relation is found between participating in (co-)funded R&D projects and knowledge commercialisation, contradicting hypothesis 2d.

Furthermore, the results in table 1 show that experience with a combination of R&D activities is positively related to knowledge creation, which supports hypothesis 3a. An important note here is that this relation is only found for organisations that also have experience with publishing a publication. In other words, gaining experience with R&D activities benefits knowledge creation in case the activities are at least combined with publication experience. The same type of effect is found for the relation between prior experience in a combination of R&D activities and knowledge commercialisation (table 2). Here, gaining experience with multiple R&D activities also shows a positive effect, which supports hypothesis 3b. However, the experience with multiple R&D activities is only beneficial for knowledge commercialisation when combined with patent experience. For both knowledge creation and knowledge commercialisation the effect of experience with multiple R&D activities is stronger than the individual effect of respectively experience with publishing or patenting.

4.2 R&D network position and R&D output

4.2.1 Number of collaborations

The results in table 1 show a negative relationship between the number of collaborations in the funding network and knowledge creation. The number of collaborations in the publication network also indicates a negative effect. However, this relation is only significant in network database 1 Year ($p<0,1$). In the patent network no significant relation is found. Overall, hypothesis 4a is contradicted for the funding network. The results for the publication network and patent network are found insufficient to consider hypothesis 4a supported.

The relation between the number of collaborations and knowledge commercialisation gives also conflicting results between the R&D networks (table 2). In the funding network the number of collaborations is positively related to knowledge commercialisation. This is in contrast to the patent network, where the number of collaborations has a negative effect. Hence, hypothesis 3b is supported for the funding network, but it is contradicted for the patent network. The relation in the publication network lacks significance; hence hypothesis 3b is not supported for this R&D network.

To get a better understanding of the influence of collaborations on R&D output, the relation between collaborating with a certain organisation type and its effect on the R&D output is further analysed. This is done by adding binary variables to the model for collaboration with a research organisation, multinational company, national company and other organisation per R&D network. These binary variables indicate what the influence is on R&D output for collaborating with an organisation type, relative to not collaborating with that organisation type. The addition of the variables to the final model improves the model significantly (knowledge creation $p<0,01$; knowledge commercialisation $p<0,01$). The model is also run with an interaction on the control variable of organisation type with the variables of collaboration with a certain organisation type. Since this did not improve the model, this interaction is left out of the analysis.

The analyses show that collaborating with research organisations, multinational companies and national companies in the funding network has a positive effect on knowledge creation (table 1). As for the publication network, it is beneficial for an organisation's knowledge creation to work together with multinational companies. The results show a negative effect in the publication network for collaborating with "other" organisations such as government agencies or intermediaries. Besides, with the control variable on organisation type it is found that research organisations have a significant higher probability of creating knowledge than multinational or national companies. Thus, although it is relevant to collaborate with multinational and national companies for knowledge creation, research organisations are the most important source for knowledge creation for innovation.

In the patent network collaborating with research organisations, national companies and multinational companies has a positive effect on knowledge commercialisation (table 2). In the funding network, working together with "other" organisations influences knowledge commercialisation negatively. Finally, with the control variable on organisation type it is found that multinational companies have a greater probability to commercialise knowledge than research organisations. This indicates that in this case companies are the dominant organisation type with respect to knowledge commercialisation.

4.2.2 Degree of clustering

The results show that there is a negative influence of the degree of clustering in the publication network and the funding network on knowledge creation (table 1). In other words, in case the connections of an organisation to a low extent also collaborate with each other, the organisation is more likely to create knowledge. This supports hypotheses 5a for the funding and publication network. However, no relationship is found for the degree of clustering in the patent network and knowledge creation. The results in table 2 show that there is no significant relation between the influence of the degree of clustering on knowledge commercialisation. Therefore, in this case the data does not support hypothesis 4b.

5 Conclusions and discussion

5.1 Conclusions

In this paper we studied the following research question: *"How do experiences with different types of R&D activities and R&D network positions influence and organisation's knowledge creation and knowledge commercialisation?"* Insights from organisational learning and social network theory are combined with the co-evolving perspective; processes leading to knowledge creation and knowledge commercialisation co-evolve and interact. This combination of perspectives is used to explain the R&D output of organisations active in R&D. We present our findings in light of the co-evolving perspective.

Our results do not show cross-relationships between experience in the R&D activity of filing a patent and knowledge creation and between experience in the R&D activity of publishing a scientific article and knowledge commercialisation. However, an organisation's experience with the R&D activity of publishing does benefit knowledge creation. Similarly, experience with the R&D activity of filing a patent positively affects an organisation's knowledge commercialisation.

Looking further at the cross-relationships in the R&D networks, we find that there is no effect of the number of collaborations in the patent network on knowledge creation. Whereas a negative relation is found between the number of collaborations in the patent network and knowledge commercialisation. Notable here is that the results do not show a relationship between the number of collaborations in the publication network and knowledge creation or knowledge commercialisation. In the cross-relationship for the degree of clustering, we do not find a relationship in the publication network on knowledge commercialisation. However, we do find support for the negative relation between the degree of clustering in the publication network and knowledge creation. In contrast, no relationships are found between the degree of clustering in the patent network and knowledge creation or knowledge commercialisation. Altogether, the relationship between network position and the R&D output in terms of knowledge creation and knowledge commercialisation differs per network and per R&D output. However, the results show that there are no cross-relationships between an organisation's position in the patent network and knowledge creation and between an organisation's position in the publication network and knowledge commercialisation.

Nevertheless, our findings do show effects on knowledge creation and knowledge commercialisation for organisations that have experience with participating in in publicly (co-) funded R&D projects. This holds specifically for organisations that combine the experience of participating in publicly (co-)funded R&D projects with the experiences of publishing a scientific article or of filing a patent; this gives a positive influence on respectively knowledge creation and knowledge commercialisation. Also from a network perspective the organisations participating in publicly (co-)funded R&D projects tie the processes leading to knowledge creation and knowledge commercialisation together. It is found that the number of collaborations in the funding network negatively influences knowledge creation, but positively influences knowledge commercialisation. Furthermore, support is found that the degree of clustering in the funding network has a negative relationship with knowledge creation. However, no relationship is found between the degree of clustering in the funding network and knowledge commercialisation.

Overall, our results imply that the realms of creating knowledge and commercialising knowledge co-evolve to a limited extent on the level of the organisation. This is also supported by the findings that research organisations are the dominant organisation type for knowledge commercialisation and (multinational) companies are more likely to commercialise knowledge; in CO_2 capture R&D a distinction between knowledge creation and knowledge commercialisation for organisation is still visible. However, for organisations that have experience in publicly (co-) funded R&D projects, which combine this experience with other R&D experiences, there are opportunities for knowledge creation and knowledge commercialisation to co-evolve within the organisation.

5.2 Discussion

There are two main limitations and opportunities for further research that warrant further discussion. We conclude with practical implications for R&D programme managers and organisations active in R&D.

First, there is no rule of thumb for the period prior experiences or connections between organisations stay valuable. Therefore, it is difficult to determine for what period organisations and their connections should be included in a network and what effect this has on R&D output. We dealt with this issue by creating three network databases that included organisations and their connections over different timespans. A more precise exercise would be to research the value of prior experiences and continuation of connections qualitatively. However, it would be difficult to uncover for each and every organisation and connection and therefore very time-consuming (Ahuja, 2000). Moreover, as the method used in this study resulted in consistent results over the different network databases, we expect to have sufficiently controlled for this limitation.

Secondly, in this study the participation and collaboration of publicly funded R&D projects are directly included in this research. Data on privately funded R&D projects was unavailable. Therefore, these efforts are indirectly included through the publication and patent network. With this limitation it is possible that this study excluded organisations and collaborations that did not produce an R&D output in terms of a publication or a patent. However, we expect that this is only a small percentage. Since the CO_2 capture processes are not yet developed in to commercial technologies, it is expected that the largest part of R&D efforts in CO_2 capture are facilitated by publicly (co)-funded R&D projects. Future research studying a further developed industry should aim to replicate our findings and if possible include participating in privately funded R&D projects directly, besides the R&D activities which are currently included in this research.

Moreover, as this research gains insights in the co-evolvement of the processes leading to knowledge creation and knowledge commercialisation on the level of the organisation, new opportunities for research arise. Our findings imply that there are organisations in which the processes of knowledge creation and knowledge commercialisation co-evolve. However, there is also still a clear distinction between the type of organisation creating knowledge and commercialising knowledge. Where this study focused on the quantity of knowledge creation and knowledge commercialisation, this opens a research opportunity for the quality of knowledge creation and knowledge commercialisation. The question can be raised if the quality of knowledge created and commercialised differs for organisations in which the processes co-evolve internally from organisations in which the processes co-evolve externally i.e. within the system of innovation. This is especially interesting in the light of innovation policy, as currently scholars are arguing for innovation policy stimulating the co-evolvement of these processes on organisational level (Etzkowitz & Leydesdorff, 2000). Our results also show that the network position in (different) R&D networks can influence the R&D output differently. Further research can explore explanations for these differences.

Finally, this research has five practical implications for policy on publicly (co-)funded R&D:

1. R&D programme managers should try to include organisations in publicly (co-)funded R&D projects that have not previously collaborated in this type of R&D projects. This way the diversity of R&D activities is directly supported which relates to our results as they show that experience in a combination of R&D activities benefits R&D output. Our results also show that a low degree of clustering positively influences knowledge creation. By including organisations that have not previously participated in R&D projects the lower degree of clustering is indirectly facilitated.
2. When managers of publicly (co-)funded R&D programmes aim at knowledge creation, they should try to keep the R&D project group small as the results indicate that a lower number of collaboration is beneficial for knowledge creation.

3. Our results show a positive influence of organisations with (cumulative) experience on knowledge creation. Also the organisation type national companies and research organisations are positively related with knowledge creation. Therefore, R&D programme managers should aim to include these types of organisations in publicly (co-)funded R&D projects in case knowledge creation is the goal.
4. Our findings also show that organisations with (cumulative) patenting experience benefit knowledge commercialisation. Therefore R&D programme managers should aim to include this type of organisation if the R&D project is aiming at knowledge commercialisation.
5. In case knowledge commercialisation is the goal of the R&D project R&D managers should aim at keeping the role of other organisations such as government agencies and intermediaries low. This relates to the results that other organisations have a negative influence on knowledge commercialisation.

This study also has practical implications for organisations active in R&D. First of all, the findings show that it is beneficial for an organisation aiming at creating knowledge to work together with multinational companies. On the contrary these organisations should not collaborate with other organisations, such as public agencies and intermediaries, as our results show that this type of collaboration influence knowledge creation negatively. Finally, it is generally not recommended to collaborate in patenting activities as the results indicate that the number of collaborations in the patent network has a negative influence on knowledge commercialisation. However, in case an organisation does collaborate in patenting activities, our results show that working together with national companies and research organisations influences knowledge commercialisation positively.

Acknowledgements

My thanks go to Frank van Rijnsoever, who was my supervisor from the Utrecht University for this research. I conducted this research for the largest part during a research internship at the Global CCS Institute in Canberra. Therefore, feedback and discussions about the direction of this research mainly took place over e-mail or Skype. Although these sessions were very valuable for the research and my learning process, I would like to acknowledge the contribution of the sessions with Frank in Sydney and Canberra. My thanks are also due to the project team of the Global CCS Institute for welcoming me as a part of the team and providing me with expert views on CCS related subjects. Specifically, I would like to thank Klaas van Alphen, for providing me with feedback and for being a sparring partner during my research, even once I got back to the Netherlands. Besides, I thank my family Willem, Monique, Tanja and Karin van Prooijen for their support and encouraging words when needed. Finally, I am grateful for the support of Henk Steinz, who supported me with every step forwards (and backwards) in this research process.

References

Agrawal, A. & R. Henderson. (2002), Putting patents in context: Exploring knowledge transfer from MIT. Management Science 48(1), pp. 44-60.

Ahuja, G. (2000), Collaboration networks, structural holes, and innovation: A longitudinal study. Administrative Science Quarterly 45(3), pp. 425-455.

Ames, E. (1961), Research, invention, development and innovation. The American Economic Review 51(3), pp. 370-381.

Becker, M.C. (2004), Organizational routines: a review of the literature. Industrial and corporate change 13(4), pp. 643-678.

Burt, R.S. (2001), Structural holes versus network closure as social capital. Social capital: Theory and research, pp.31-56.

Bush, V. (1945), Science: The endless frontier. Transactions of the Kansas Academy of Science (1903-) 48(3), pp.231-264.

Cassiman, B. & R. Veugelers. (2006), In search of complementarity in innovation strategy: Internal R & D and external knowledge acquisition. Management Science 52(1), pp. 68-82.

Chaminade, C. & C. Edquist. (2006), From theory to practice: The use of the systems of innovation approach in innovation policy. Innovation, Science and Institutional Change, Oxford University Press, Oxford

Cockburn, I.M. & R.M. Henderson. (1998), Absorptive capacity, co-authoring behavior, and the organization of research in drug discovery. Journal of Industrial Economics 46(2), pp. 157-182.

Cohen, W.M. & D.A. Levinthal. (1990), Absorptive capacity: a new perspective on learning and innovation. Administrative Science Quarterly, pp. 128-152.

Cohen, W.M., R.R. Nelson & J.P. Walsh. (2002), Links and impacts: The influence of public research on industrial R&D. Management Science 48(1), pp.1-23.

Coleman, J.S. (1988), Social capital in the creation of human capital. American journal of sociology, pp. S95-S120.

CORDIS. (2012). Search Cordis Europe [online]., 2013. Available on the world wide web: <http://cordis.europa.eu/newsearch/index.cfm?page=advSearch&js=1&language=en>.

Csardi G. & T. Nepusz. (2006). The igraph software package for complex network research

Debackere, K., B. Clarysse & M.A. Rappa. (1996), Dismantling the ivory tower: The influence of networks on innovative output in emerging technologies. Technological forecasting and social change 53(2), pp. 139-154.

EPSRC. (2013). EPSRC support by Research Area in Carbon Capture and Storage [online]., 2013. Available on the world wide web: <http://gow.epsrc.ac.uk/NGBOChooseTTS.aspx?Mode=ResearchArea&ItemDesc=Carbon+Capture+and+Storage>.

Etzkowitz, H. & L. Leydesdorff. (2000), The dynamics of innovation: From National Systems and "mode 2" to a Triple Helix of university-industry-government relations. Research Policy 29(2), pp. 109-123.

Geuna, A. & L.J.J. Nesta. (2006), University patenting and its effects on academic research: The emerging European evidence. Research Policy 35(6), pp. 790-807.

Gittelman, M. & B. Kogut. (2003), Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. Management Science 49(4), pp. 366-382.

Giuliani, E. (2013), Clusters, Networks and Firms' Product Success, an Empirical Study. Management Decision 51(6), pp. 2-2.

Godin, B. (2006), The linear model of innovation: The historical construction of an analytical framework. Science Technology and Human Values 31(6), pp. 639-667.

Gulati, R. & M. Gargiulo. (1999), Where do inter-organizational networks come from? 1. American journal of sociology 104(5), pp. 1439-1493.

Hekkert, M.P. & S.O. Negro. (2009), Functions of innovation systems as a framework to understand sustainable technological change: Empirical evidence for earlier claims. Technological Forecasting and Social Change 76(4), pp. 584-594.

Hekkert, M.P., R.A. Suurs, S.O. Negro et al. (2007), Functions of innovation systems: A new approach for analysing technological change. Technological Forecasting and Social Change 74(4), pp. 413-432.

Huber, G.P. (1991), Organizational learning: The contributing processes and the literatures. Organization science 2(1), pp. 88-115.

IEAGHG. (2013). IEA Greenhouse Gas R&D Programme - RD&D Database [online]., 2013. Available on the world wide web: <http://ieaghg.org/ccs-resources/rd-database>.

Legg J. F. & F.R. Campbell. (2006). Carbon Dioxide Capture and Storage: A Compendium of Canada's Participation. NRCan.

Levitt, B. & J.G. March. (1988), Organizational learning. Annual review of sociology, pp. 319-340.

McFadyen, M.A., M. Semadeni & A.A. Cannella Jr. (2009), Value of strong ties to disconnected others: Examining knowledge creation in biomedicine. Organization Science 20(3), pp. 552-564.

Mora-Valentin, E.M., A. Montoro-Sanchez & L.A. Guerras-Martin. (2004), Determining factors in the success of R&D cooperative agreements between firms and research organizations. Research Policy 33(1), pp. 17-40.

Mowery, D.C., R.R. Nelson, B.N. Sampat et al. (2001), The growth of patenting and licensing by U.S. universities: An assessment of the effects of the Bayh-Dole act of 1980. Research Policy 30(1), pp. 99-119.

Nelson, R. & Winter, S. (1982). An evolutionary theory of economic change. Harvard University Press.

NETL. (2002-2012). Carbon Capture and Storage Project Factsheet [online]., 2013. Available on the world wide web: <http://www.netl.doe.gov/publications/factsheets/proj_toc.html>.

NETL. (2011-2012). Industrial Carbon Capture and Sequestration (ICCS) Project Factsheet [online]., 2013. Available on the world wide web:
<http://www.netl.doe.gov/publications/factsheets/proj_toc.html>.

OECD. (2009). Policy Responses to the Economic Crisis: Investing in Innovation for Long-Term Growth. OECD Publishing.

Owen-Smith, J. & W.W. Powell. (2004), Knowledge networks as channels and conduits: The effects of spillovers in the Boston biotechnology community. Organization science 15(1), pp. 5-21.

Partha, D. & P.A. David. (1994), Toward a new economics of science. Research policy 23(5), pp. 487-521.

Polidoro Jr., F. & M. Theeke. (2012), Getting competition down to a science: The effects of technological competition on firms' scientific publications. Organization Science 23(4), pp. 1135-1153.

Powell, W.W., K.W. Koput, L. Smith-Doerr et al. (1999), Network position and firm performance: Organizational returns to collaboration in the biotechnology industry. Research in the Sociology of Organizations 16(1), pp. 129-159.

Powell, W., K. Koput & L. Smith-Doerr. (1996), Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. Administrative Science Quarterly 41(1), pp. 116-145.

R Core Team. (2013). R: A language and environment for statistical computing

Shan, W.J., G. Walker & B. Kogut. (1994), Interfirm Cooperation and Startup Innovation in the Biotechnology Industry. Strategic Management Journal 15(5), pp. 387-394.

Siegel, D.S., D. Waldman & A. Link. (2003), Assessing the impact of organizational practices on the relative productivity of university technology transfer offices: An exploratory study. Research Policy 32(1), pp. 27-48.

Soda, G., A. Usai & A. Zaheer. (2004), Network memory: The influence of past and current networks on performance. Academy of Management Journal 47(6), pp. 893-906.

Teece, D.J. (1986), Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. Research policy 15(6), pp. 285-305.

Tidd, J. & Bessant, J. (2011). Managing innovation: integrating technological, market and organizational changeWiley. com.

Tijssen, R.J. (2004), Is the commercialisation of scientific research affecting the production of public knowledge? Global trends in the output of corporate research articles. Research Policy 33(5), pp. 709-733.

Tijssen, R. (1998), Quantitative assessment of large heterogeneous R&D networks: The case of process engineering in the Netherlands. Research Policy 26(7-8), pp. 791-809.

Tsai, W. (2001), Knowledge transfer in intra-organizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. Academy of management journal 44(5), pp. 996-1004.

Uzzi, B. & J. Spiro. (2005), Collaboration and Creativity: The Small World Problem1. American Journal of Sociology 111(2), pp. 447-504.

Veefkind V. & J. Hurtado-Albir. (2010). Clean Energy and Patents. Munich, Germany: European Patent Office.

Verbeek, A., K. Debackere, M. Luwel et al. (2002), Linking science to technology: Using bibliographic references in patents to build linkage schemes. Scientometrics 54(3), pp. 399-420.

Wasserman, S. (1994). Social network analysis: Methods and applications. Cambridge university press.

Zaheer, A. & G.G. Bell. (2005), Benefiting from network position: firm capabilities, structural holes, and performance. Strategic Management Journal 26(9), pp. 809-825.

Zaheer, A., B. McEvily & V. Perrone. (1998), Does trust matter? Exploring the effects of inter-organizational and interpersonal trust on performance. Organization science 9(2), pp. 141-159.

Zollo, M., J.J. Reuer & H. Singh. (2002), Inter-organizational routines and performance in strategic alliances. Organization Science 13(6), pp. 701-713.

Zollo, M. & S.G. Winter. (2002), Deliberate learning and the evolution of dynamic capabilities. Organization Science 13(3), pp. 339-351.

Zucker, L.G., M.R. Darby & J.S. Armstrong. (2002), Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology. Management Science 48(1), pp. 138-153.

Appendix A Elaborating on the degree of clustering

In this appendix the discussion between the benefits of a high degree of clustering and the benefits of a low degree of clustering for an organisation is described and visualised.

On the one hand, there are scholars that argue the benefit of networks with closure i.e. a high degree of clustering (Coleman, 1988). A situation with a high degree of clustering for organisation D, and its set of organisations, is visualized in Figure 2. As the organisations in this set are all connected with each, information on deviant behaviour quickly disseminates to all organisations, organisations feel the necessity to behave according to the shared norms and values. This results in high trust amongst the organisation which facilitates knowledge sharing and stimulates innovative output. Therefore, it is possible that the set of organisations of organisation D have a high R&D. As organisation C has a low degree of clustering, it is possible that this organisation will behave opportunistic. After all, organisation C can say one thing to Set 1 and another thing to Set 2. Therefore the trust between organisation C and other organisations is low and there is a possibility that this hampers its R&D output.

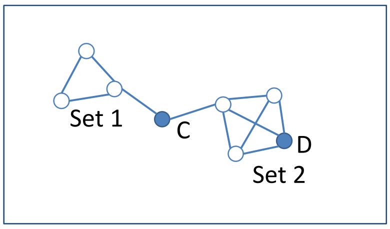

Figure 2: The high network closure of D and the low network closure of C.

On the other hand, there are scholars that argue the benefit of spanning a structural hole i.e. a low degree of clustering. As information is known to circulate within sets of connected organisations before it circulates between sets of organisations, the 'broker' can control the information flows between the sets of organisations. In Figure 2 this is visualised by organisation C as Information between Set 1 and Set 2 set must first go through organisation C. Moreover, organisation C gains access to the information from both sets of organisations enabling organisation C – the broker – to combine information from the different sources. From this perspective, gaining access to multiple knowledge sources through spanning structural holes i.e. a low degree of clustering, increases the likelihood of positive R&D output (Burt, 2004).

Appendix B CO_2 capture for Carbon Capture and Storage

Table 3 provides an introduction to the CO_2 capture processes which are the subject of the government funded R&D projects, publications and patents included in this research.

Table 3: Introduction of the main capture processes for this study.

Capture technology	**Description**
Absorption	This process is based on dissolving CO_2 into liquid using solvents.
Adsorption	In this process CO_2 is collected at the surface of a solid.
Chemical separation	This process separates CO_2 by means of chemical processes such as chemical looping where a reactive material circulates to drive the chemical process of CO_2 capture.
Biological separation	CO_2 is captured by with biological processes such as by algae or photovoltaic-like processes.
Membranes or diffusion	This process separates CO_2 from the flue gases by using of membranes.
Rectification and condensation	This process captures nearly pure stream of CO_2 by cooling the flue gas, separating solid CO_2 that forms during cooling which is than compressed.
Combined processes	Processes that can be categorized from combinations of the processes described above.

Appendix C Detailed description of the data collection process

C.1 Funding dataset

The funding dataset is a combination of an existing dataset from Utrecht University which is completed with data from NETL, Natural Resources Canada, CORDIS and European grant databases of national research programs (CORDIS, 2012; Legg & Campbell (NRCan), 2006; NETL, 2002-2012). Both the project data for the United States as for Europe needed additional R&D project information for 2009 and 2010. The Canadian database needed complementing information for 2006 until 2010.

From each data source the data collection method was different. First of all for the data collection on the CO_2 capture R&D projects from NETL, the project factsheet of Carbon Capture and Storage as well as Industrial Carbon Capture and Sequestration was used. The projects were scanned on their titles and project description to check whether the R&D project focused on CO_2 capture. In total 17 projects from the United States were added to the funding dataset. Secondly, we had contact with the National Resources Canada to request the draft version of the CCS Compendium to complement the period of 2006 till 2010. Fortunately, the National Resources Canada was willing to share the draft version with the Global Carbon Capture and Storage Institute, which gave us access to the CO_2 R&D project information from Canada. A total of 23 Canadian R&D projects were added to the dataset. Finally, for R&D project information on European level, we searched for "Projects" on amongst others "CO_2 capture" in the CORDIS database, the IEA GHG database and national grant schemes such as the UK EPSRC. In total 19 European R&D projects on CO_2 capture were included in the dataset.

C.2 Publication dataset

The results of the search queries that were iteratively established were analysed for their content. During this analysis 1511 articles were included in the final data set as they included for instance "carbon dioxide absorption". All other 963 articles were analysed individually based on their title and abstract to determine whether the publication focused on CO_2 in relation to CCS. From this analysis a total of 587 publications were excluded from the data set, leaving 1972 publications and 753 organisations for the period between 2002 and 2012. Unfortunately, we needed to exclude the publication and organisations active in 2011 and 2012 as the patent data was only reliable till 2010. After limiting the publication dataset to the time period 2002 till 2010, 976 articles and 454 organisations remained to form the publication dataset of this research.

C.3 Patent dataset

The PATSTAT database used for this research includes patents that were published before October 2012. All patents that were filed before October 2012, but that were not officially published are not included in the PATSTAT database. As the time period between application and publication can be up to 18 month, the most reliable patent data in this database is till 2010. Therefore, this research limited its timespan between 2002 and 2010.

The data from PATSTAT is derived through the following steps. Firstly, the database was limited for the period from 2002 till 2010 based on the application date of the patent. Secondly, the geographical delineation is made on both the country of the inventor, independent of the country of the patent applicant, and country of the patent applicant, independent of the inventor's country. This way least patents were excluded by incomplete information on either the applicant's country or the inventor's country. Furthermore, as the database included CCS categories per patent it was possible to only include the categories related to CO_2 capture. After collecting the required data from the database it was necessary to clean the data based on their patent families. A patent family represents a patent which is filed at multiple patent offices. These patents are included in the database with their own application number, but based on their patent family number it is possible to identify these patents as one patent. The patent which was applied first within the patent family was included in the final data set. To keep information on the applicants all applicants who filed the patent, regardless or the patent office, are included in the final dataset.

C.4 An overview of the datasets

Table 4 gives an overview of the content and main sources of each dataset in this research. A number of organisations in the three datasets overlap. Therefore in total 1069 organisations are included in this research.

Table 4: An overview of the R&D projects, publications, patents and organisations

Dataset	Content	Source
Funding dataset	253 R&D projects 430 organisations	NETL Project Fact Sheet Canadian CCS Compendia European CORDIS IEA GHG RD&D database UK EPSRC database
Publication dataset	976 scientific articles 454 organisation	Web of Science
Patent dataset	1379 patents 446 organisations	PATSTAT database

Appendix D Illustrative table of network database construction

Table 5 provides an illustration of how the network databases are constructed. Year(t) refers to the collaboration year; the start year for R&D projects, the publication year for publications, and the application year for patents. The headings 1 Year, 2 Years and 3 Years refer to the three different databases including organisations and their connections of the collaboration year and respectively the organisation and connections from 1 year before the collaboration year, 2 years before the collaboration year and three years before the collaboration year. Table 4 helps clarifying the structure of the network databases.

Table 5: Network construction for social network analysis

Year(t)	1 Year (t + t-1)	2 Years (t + t-1+t-2)	3 Years (t + t-1+t-2+t-3)
2002	2002	2002	2002
2003	2003+2002	2003+2002	2003+2002
2004	2004+2003	2004+2003+2002	2004+2003+2002
2005	2005+2004	2005+2004+2003	2005+2004+2003+2002
2006	2006+2005	2006+2005+2004	2006+2005+2004+2003
Etcetera.			

Appendix E R script social network analysis

This script is an example of the final script. The script here only shows the analysis of the data for 2002 whilst the same analysis is run for each year included in this research (2002-2010).

```
#Load the required packages and set working directory
library(igraph)
library(tcltk)
setwd("setwd")

#Read edgelist per year in R
Data_02<-read.csv("Data_02.csv", sep=";", header=T)

#Read vertice attributes per year
a02<-read.csv("Vertice_attr_02.csv", sep=";", header=T)

#Make an igraph object of the data per year
graph02<-graph.data.frame(Data_02, directed=FALSE, vertices=NULL)

#Add vertice attributes to the organisations in the graph
V(graph02)$Year=as.character(a02$Year[match (V(graph02)$name,a02$Organisation)])
V(graph02)$Type=as.character(a02$Organisation.type[match
(V(graph02)$name,a02$Organisation)])
V(graph02)$Type_specified=as.character(a02$Organisation_specified[match
        (V(graph02)$name,a02$Organisation)])
V(graph02)$Type_Other=as.numeric(a02$Type_Other[match
(V(graph02)$name,a02$Organisation)])
V(graph02)$Type_Research=as.numeric(a02$Type_Research[match
(V(graph02)$name,a02$Organisation)])
V(graph02)$Type_Multinational=as.numeric(a02$Type_Multinational[match
        (V(graph02)$name,a02$Organisation)])
V(graph02)$Type_National=as.numeric(a02$Type_National[match
(V(graph02)$name,a02$Organisation)])
V(graph02)$Continent=as.character(a02$Continent[match (V(graph02)$name,a02$Organisation)])

#Remove all ties that are loops to organisations themselves (loops of isolates)
graph02S<-simplify(graph02, remove.multiple = TRUE, remove.loops = TRUE)

#Make a list of all graphs to be able to run analyses on multiple graphs simultaneously
graphlist<-list(graph02S, graph03S, graph04S, graph05S, graph06S, graph07S, graph08S, graph09S,
graph10S)
names(graphlist)<-c("Year2002", "Year2003", "Year2004", "Year2005", "Year2006", "Year2007",
"Year2008", "Year2009", "Year2010")

#For all network descriptives:
#Find the number of organisations per year in the network
vcounts<-lapply(graphlist, vcount)

#Average path length in the network
a_path_length<-lapply(graphlist, average.path.length, unconnected=TRUE)
```

#Mean degree network level
meandegree<-lapply(graphlist, function(x)mean(degree(x)))

#Extract the amount of clusters in the network, CL=cluster
graph02SCL<-clusters(graph02S)

#Make a list of all clusters to be able to run analyses on multiple clusters simultaneously
clusterlist<-list(graph02SCL, graph03SCL, graph04SCL, graph05SCL, graph06SCL, graph07SCL,
    graph08SCL, graph09SCL, graph10SCL)
names(clusterlist)<-c("Year2002", "Year2003", "Year2004", "Year2005", "Year2006", "Year2007",
    "Year2008", "Year2009", "Year2010")

#Number of clusters
no_clusters<-lapply(clusterlist, function(x)x$no)

#Frequency of the sizes of clusters in the network
size_clusters<-lapply(clusterlist, function(x)x$csize)

#Measure the cluster coefficient on network level by calculating transitivity.
network_cc<-lapply(graphlist, transitivity, type=c("global"))

#Isolates
isolates<-lapply(graphlist, function(x)sum(degree(x)==0))

#Merge the network descriptives
networkdes<-mapply(c, vcounts, a_path_length, no_clusters, isolates, meandegree, network_cc)
write.table(networkdes, "Location and name of output.txt", sep=",")

#For all cluster descriptives
#Extract largest cluster, LC = largest cluster
graph02SLC<-induced.subgraph(graph02S,which(graph02SCL$membership ==
which.max(graph02SCL$csize)))

#Make a list of all largest clusters to be able to run analyses on multiple largest clusters simultaneously
cluster_large_list<-list(graph02SLC, graph03SLC, graph04SLC, graph05SLC, graph06SLC, graph07SLC,
    graph08SLC, graph09SLC, graph10SLC)
names(cluster_large_list)<-c("Year2002", "Year2003", "Year2004", "Year2005", "Year2006",
    "Year2007", "Year2008", "Year2009", "Year2010")

#Number of organisations within largest cluster
vcounts_cluster<-lapply(cluster_large_list, vcount)

#Average path length within largest cluster
a_path_length_cluster<-lapply(cluster_large_list, average.path.length, unconnected=TRUE)

#Measure cluster coefficient on cluster level
cluster_cc<-lapply(cluster_large_list, transitivity, type=c("global"))

#Merge cluster descriptives
clusterdes<-mapply(c, vcounts_cluster, a_path_length_cluster,cluster_cc)
write.table(clusterdes, "Location and name of output.txt", sep=",")

#Measures of organisations on network level

```
dataframe_network<-lapply(graphlist, function(x)get.data.frame(x, what=c("vertices")))
degree_network<-lapply(graphlist, function(x)degree(x, v=V(x), mode=c("all")))
clusterC_network<-lapply(graphlist, transitivity, type=c("local"),isolates=c("zero"))

Percentage_multinational_network<-lapply(graphlist, function(x){unlist(lapply(get.adjlist(x,
        mode="all"), function (y) { sum(V(x)[y]$Type_Multinational, na.rm=T)} )) / degree(x,v=V(x),
        mode="all")})

Percentage_national_network<-lapply(graphlist, function(x){unlist(lapply(get.adjlist(x, mode="all"),
        function (y) { sum(V(x)[y]$Type_National, na.rm=T) } )) / degree(x,v=V(x), mode="all")})

Percentage_research_network<-lapply(graphlist, function(x){unlist(lapply(get.adjlist(x, mode="all"),
        function (y) { sum(V(x)[y]$Type_Research, na.rm=T) } )) / degree(x,v=V(x), mode="all")})


Percentage_other_network<-lapply(graphlist, function(x){unlist(lapply(get.adjlist(x, mode="all"),
        function (y) { sum(V(x)[y]$Type_Other, na.rm=T) } )) / degree(x,v=V(x), mode="all")})
```

#Network 02

```
network02<-mapply(c, degree_network$Year2002, Percentage_multinational_network$Year2002,
        Percentage_national_network$Year2002, Percentage_research_network$Year2002,
        Percentage_other_network$Year2002, clusterC_network$Year2002)
matrix02<-matrix(unlist(network02), nrow=vcount(graph02S), ncol=6, byrow=TRUE)
        colnames(matrix02)<-c("Degree", "Percentage_Multinational", "Percentage_National",
        "Percentage_Research", "Percentage_Other", "Cluster_coefficient")
rownames(matrix02)<-dataframe_network$Year2002$name
write.table(matrix02,"Location and name output.txt", sep=",")
write.table(dataframe_network$Year2002,"Location and name output.txt", sep=",")
-------------------------------------------------------------------------------------------------------------------------------
```

Appendix F Descriptive statistics and correlation matrices

F.1 Descriptive statistics

Table 6 displays the descriptive statistics for the (numeric) variables in this study.

Table 6: Mean and standard deviation of the (numeric) variables

Variable	**Specification**	**Network database 1 Year**		**Network database 2 Years**		**Network database 3 Years**	
		Mean	*Standard deviation*	*Mean*	*Standard deviation*	*Mean*	*Standard deviation*
Number of publications per year		0,39	1,07	0,36	1,03	0,34	1,01
Number of patents per year		0,36	1,27	0,33	1,23	0,32	1,20
Cumulative publications		0,77	2,64	0,74	2,55	0,72	2,50
Cumulative patents		1,17	4,15	1,14	4,01	1,14	3,92
Number of collaborations	Funding	7,37	13,89	7,05	13,71	6,85	13,56
	Publication	0,40	1,15	0,47	1,30	0,50	1,40
	Patent	0,11	0,45	0,14	0,52	0,16	0,58
Degree of clustering	Funding	0,34	0,44	0,32	0,44	0,31	0,43
	Publication	0,07	0,25	0,08	0,26	0,08	0,26
	Patent	0,01	0,11	0,02	0,13	0,02	0,13

F.2 Correlation matrix network database 1 Year

Network database 1 Year	First pub	First pat	Cumul. Pub.	Cumul. Pat.	Degree - Funding	Degree - Publication	Degree - Patent	Collab. Research Fun.	Collab. Research Pub.	Collab. Research Pat.	Collab. M-comp. Fun	Collab. M-comp. Pub	Collab. M-comp. Pat	Collab. N-comp. Fun	Collab. N-comp. Pub	Collab. N-comp. Pat	Collab. Other. Fun	Collab. Other. Pub	Collab. Other. Pat	Clustering Fun	Clustering Pub
First pat	-0,01																				
Cumul. Pub.	-0,05	-0,03																			
Cumul. Pat.	0,01	-0,04	0,10																		
Degree - Funding	0,00	-0,03	0,33	0,30																	
Degree - Publication	-0,05	-0,03	0,55	0,15	0,31																
Degree - Patent	0,00	-0,03	0,07	0,24	0,14	0,11															
Collab. Research Fun.	0,00	-0,04	0,18	0,14	0,69	0,18	0,04														
Collab. Research Pub.	-0,07	-0,03	0,46	0,11	0,20	0,76	0,05	0,15													
Collab. Research Pat.	-0,01	-0,02	0,06	0,21	0,11	0,11	0,68	0,02	0,05												
Collab. M-comp. Fun	0,00	-0,04	0,19	0,14	0,73	0,18	0,05	0,88	0,15	0,03											
Collab. M-comp. Pub	-0,03	-0,01	0,33	0,14	0,25	0,53	0,08	0,13	0,31	0,05	0,14										
Collab. M-comp. Pat	0,00	-0,02	0,01	0,23	0,10	0,08	0,59	0,04	0,03	0,14	0,04	0,09									
Collab. N-comp. Fun	0,00	-0,04	0,20	0,13	0,74	0,18	0,04	0,82	0,14	0,02	0,84	0,14	0,03								
Collab. N-comp. Pub	-0,02	-0,01	0,20	0,03	0,09	0,27	0,02	0,06	0,22	0,01	0,05	0,13	0,02	0,07							
Collab. N-comp. Pat	-0,01	-0,02	0,03	0,07	0,04	0,00	0,52	0,01	0,01	0,07	0,00	0,01	0,11	0,01	0,02						
Collab. Other. Fun	-0,01	-0,02	0,20	0,13	0,60	0,18	0,04	0,52	0,13	0,03	0,58	0,15	0,04	0,57	0,08	0,01					
Collab. Other. Pub	-0,01	-0,01	0,12	0,00	0,05	0,31	0,05	0,03	0,19	0,05	0,04	0,04	0,03	0,02	0,06	-0,01	0,04				
Collab. Other. Pat	0,03	-0,01	0,00	0,00	-0,01	0,00	0,14	-0,01	0,01	0,14	-0,01	0,00	0,00	0,00	0,00	0,00	0,01	0,00			
Clustering Fun	0,00	-0,05	0,09	0,07	0,54	0,09	0,00	0,91	0,09	0,00	0,86	0,07	0,02	0,79	0,02	-0,01	0,47	0,02	0,00		
Clustering Pub	-0,05	-0,03	0,20	0,07	0,13	0,66	0,02	0,09	0,65	0,02	0,09	0,32	0,02	0,09	0,28	-0,01	0,07	0,22	-0,01	0,06	
Clustering Pat	0,01	-0,01	0,06	0,02	0,04	0,05	0,69	0,01	0,03	0,50	0,01	0,03	0,14	0,01	-0,01	0,25	0,00	0,04	0,23	0,00	0,00

F.3 Correlation matrix network database 2 Years

Network database 2 Years	First pub	First pat	Cumul. Pub.	Cumul. Pat.	Degree - Funding	Degree - Publication	Degree - Patent	Collab. Research Fun.	Collab. Research Pub.	Collab. Research Pat.	Collab. M-comp. Fun	Collab. M-comp. Pub	Collab. M-comp. Pat	Collab. N-comp. Fun	Collab. N-comp. Pub	Collab. N-comp. Pat	Collab. Other. Fun	Collab. Other. Pub	Collab. Other. Pat	Clustering Fun	Clustering Pub
First pat	-0,01																				
Cumul. Pub.	-0,05	-0,03																			
Cumul. Pat.	0,01	-0,04	0,10																		
Degree - Funding	0,00	-0,03	0,35	0,31																	
Degree - Publication	-0,06	-0,03	0,62	0,15	0,36																
Degree - Patent	-0,01	-0,04	0,10	0,27	0,17	0,14															
Collab. Research Fun.	0,00	-0,05	0,18	0,13	0,68	0,19	0,04														
Collab. Research Pub.	-0,08	-0,04	0,46	0,10	0,23	0,74	0,06	0,15													
Collab. Research Pat.	-0,01	-0,03	0,08	0,22	0,11	0,11	0,68	0,02	0,06												
Collab. M-comp. Fun	0,00	-0,04	0,19	0,14	0,72	0,20	0,05	0,87	0,16	0,02											
Collab. M-comp. Pub	-0,03	-0,01	0,38	0,13	0,29	0,55	0,09	0,14	0,31	0,05	0,15										
Collab. M-comp. Pat	0,00	-0,03	0,01	0,26	0,11	0,08	0,58	0,04	0,02	0,15	0,05	0,08									
Collab. N-comp. Fun	0,00	-0,04	0,20	0,13	0,73	0,20	0,05	0,81	0,15	0,02	0,84	0,15	0,04								
Collab. N-comp. Pub	-0,02	-0,01	0,22	0,04	0,13	0,32	0,03	0,07	0,24	0,00	0,07	0,16	0,03	0,09							
Collab. N-comp. Pat	-0,02	-0,02	0,04	0,09	0,06	0,01	0,52	0,00	0,01	0,08	0,00	0,02	0,12	0,01	0,03						
Collab. Other. Fun	-0,01	-0,02	0,21	0,13	0,61	0,22	0,06	0,52	0,15	0,03	0,59	0,18	0,04	0,58	0,09	0,02					
Collab. Other. Pub	-0,02	-0,02	0,16	0,00	0,06	0,34	0,06	0,04	0,20	0,05	0,04	0,05	0,03	0,03	0,06	-0,01	0,05				
Collab. Other. Pat	0,03	-0,01	0,00	0,01	-0,01	0,01	0,14	-0,02	0,02	0,14	-0,01	0,00	0,00	0,01	0,00	0,00	0,02	0,00			
Clustering Fun	-0,01	-0,05	0,09	0,06	0,53	0,09	0,00	0,91	0,09	-0,01	0,85	0,07	0,02	0,79	0,03	-0,02	0,48	0,02	0,00		
Clustering Pub	-0,05	-0,03	0,20	0,06	0,15	0,62	0,02	0,09	0,65	0,02	0,09	0,32	0,01	0,10	0,29	-0,01	0,09	0,25	0,01	0,05	
Clustering Pat	0,00	-0,02	0,07	0,03	0,05	0,06	0,69	0,01	0,03	0,51	0,01	0,03	0,15	0,01	-0,01	0,26	0,00	0,04	0,24	-0,01	0,00

F.4 Correlation matrix network database 3 Years

Network database 3 years	First pub	First pat	Cumul. Pub.	Cumul. Pat.	Degree - Funding	Degree - Publication	Degree - Patent	Collab. Research Fun.	Collab. Research Pub.	Collab. Research Pat.	Collab. M-comp. Fun	Collab. M-comp. Pub	Collab. M-comp. Pat	Collab. N-comp. Fun	Collab. N-comp. Pub	Collab. N-comp. Pat	Collab. Other. Fun	Collab. Other. Pub	Collab. Other. Pat	Clustering Fun	Clustering Pub
First pat	-0,01																				
Cumul. Pub.	-0,05	-0,03																			
Cumul. Pat.	0,01	-0,04	0,10																		
Degree - Funding	0,00	-0,03	0,35	0,31																	
Degree - Publication	-0,06	-0,03	0,66	0,16	0,39																
Degree - Patent	-0,01	-0,04	0,13	0,30	0,19	0,17															
Collab. Research Fun.	0,00	-0,05	0,18	0,13	0,68	0,20	0,05														
Collab. Research Pub.	-0,08	-0,04	0,47	0,11	0,24	0,73	0,07	0,16													
Collab. Research Pat.	-0,01	-0,03	0,09	0,22	0,11	0,12	0,68	0,01	0,06												
Collab. M-comp. Fun	-0,01	-0,04	0,19	0,14	0,72	0,21	0,06	0,87	0,16	0,02											
Collab. M-comp. Pub	-0,03	-0,02	0,40	0,13	0,30	0,55	0,09	0,14	0,32	0,05	0,16										
Collab. M-comp. Pat	0,00	-0,03	0,01	0,29	0,12	0,08	0,58	0,05	0,02	0,16	0,06	0,08									
Collab. N-comp. Fun	0,00	-0,04	0,20	0,13	0,73	0,21	0,05	0,81	0,16	0,01	0,84	0,15	0,05								
Collab. N-comp. Pub	-0,03	-0,01	0,23	0,06	0,15	0,34	0,04	0,08	0,24	0,01	0,08	0,17	0,03	0,10							
Collab. N-comp. Pat	-0,02	-0,03	0,04	0,11	0,06	0,00	0,52	0,00	0,01	0,09	0,00	0,01	0,12	0,00	0,02						
Collab. Other. Fun	-0,01	-0,02	0,21	0,13	0,60	0,24	0,06	0,52	0,16	0,03	0,59	0,18	0,04	0,58	0,11	0,02					
Collab. Other. Pub	-0,02	-0,02	0,20	0,00	0,08	0,37	0,06	0,04	0,22	0,05	0,05	0,07	0,03	0,03	0,07	-0,01	0,06				
Collab. Other. Pat	0,02	-0,01	0,00	0,01	-0,01	0,02	0,13	-0,02	0,03	0,15	-0,02	-0,01	-0,01	0,01	0,00	0,00	0,02	0,00			
Clustering Fun	-0,01	-0,05	0,08	0,06	0,53	0,09	0,00	0,90	0,08	-0,01	0,85	0,06	0,02	0,79	0,03	-0,03	0,48	0,02	0,01		
Clustering Pub	-0,05	-0,03	0,20	0,07	0,16	0,60	0,02	0,09	0,65	0,02	0,10	0,32	0,01	0,10	0,29	-0,02	0,09	0,26	0,03	0,05	
Clustering Pat	-0,01	-0,02	0,09	0,04	0,05	0,06	0,69	0,01	0,03	0,52	0,01	0,02	0,16	0,01	-0,01	0,28	0,00	0,04	0,25	-0,01	0,00

Appendix G R script statistical analysis for hypotheses testing

This script is an example of the final script for hypotheses testing. The analysis shown here only includes the analyses over the network database 1 Year, whilst the analyses are also run for Network 2 Years and Network 3 Years.

#Load the required packages and set working directory

```
library(lattice)
library(Matrix)
library(lme4)
library(foreign)
library(memisc)

setwd("setwd")
```

#Read the network databases

```
data_d1<-read.spss("1 yr t-1 spss.sav",
use.value.labels=TRUE,to.data.frame=TRUE,max.value.labels=Inf,trim.factor.names=FALSE)
```

#Develop a new variable for experience with the R&D activities

```
data_d1$Presence<-(data_d1$Presence_fun)+(data_d1$Presence_pub)+(data_d1$Presence_pat)+
data_d1$Publications_peryear+data_d1$Patents_patperyear
```

#Make a nominal experience variable

```
data_d1$presencefact <-
as.factor(1000+100*data_d1$Presence_fun+10*data_d1$Presence_pub+data_d1$Presence_pat)
```

#Add observations variable

```
nrow(data_d1)
data_d1$Observations<-1:10690
```

#Model 0: Only random effects to calculate mcFadden R-square

```
model0_pubd1<-glmer(Publications_peryear~1+(1|Observations)+(1|name), data=data_d1,
family=poisson, subset=(Presence) >= 1)
summary(model0_pubd1)
```

#Model1: Years

```
model1_pubd1<-glmer(Publications_peryear~1+(1|Observations)+(1|name)+as.factor(Year),
data=data_d1, family=poisson, subset=(Presence) >= 1)
summary(model1_pubd1)
```

#Model 2: Years, Presence interaction, controls for first publication and patent
model2_pubd1<-
glmer(Publications_peryear~1+(1|Observations)+(1|name)+as.factor(Year)+presencefact+firstpub+firstpat, data=data_d1, family=poisson, subset=(Presence) >= 1)
summary(model2_pubd1)

#Model 3: Years, Presence interaction, controls for first publication and patent, Organisation type
model3_pubd1<-
glmer(Publications_peryear~1+(1|Observations)+(1|name)+as.factor(Year)+presencefact+firstpub+firstpat+Organisationtype, data=data_d1, family=poisson, subset=(Presence) >= 1)
summary(model3_pubd1)

#Model 4: Years, Presence interaction, controls for first publication and patent, Cumulative publications and Cumulative patents, Degrees, Organisation type
model4_pubd1<-
glmer(Publications_peryear~1+(1|Observations)+(1|name)+as.factor(Year)+presencefact+firstpub+firstpat+lagcumpub+lagcumpat+Degree_fundynamic+Degree_pubdynamic+Degree_patdynamic+Organisationtype, data=data_d1, family=poisson, subset=(Presence) >= 1)
summary(model4_pubd1)

#Model 5: Years, Presence interaction, controls for first publication and patent, Cumulative publications and Cumulative patents, Degrees, Clustering coefficient, Organisation type
model5_pubd1<-
glmer(Publications_peryear~1+(1|Observations)+(1|name)+as.factor(Year)+presencefact+firstpub+firstpat+lagcumpub+lagcumpat+Degree_fundynamic+Degree_pubdynamic+Degree_patdynamic+Cluster_coefficient_fundynamic+Cluster_coefficient_pubdynamic+Cluster_coefficient_patdynamic+Organisationtype, data=data_d1, family=poisson, subset=(Presence) >= 1)
summary(model5_pubd1)

#Final model including all determinants
model6_pubd1<-
glmer(Publications_peryear~1+(1|Observations)+(1|name)+as.factor(Year)+presencefact+firstpub+firstpat+lagcumpub+lagcumpat+Degree_fundynamic+Degree_pubdynamic+Degree_patdynamic+Cluster_coefficient_fundynamic+Cluster_coefficient_pubdynamic+Cluster_coefficient_patdynamic+Degree_Multinational_fun+Degree_Multinational_pub+Degree_Multinational_pat+Degree_National_fun+Degree_National_pub+Degree_National_pat+Degree_Research_fun+Degree_Research_pub+Degree_Research_pat+Degree_Other_fun+Degree_Other_pub+Degree_Other_pat+Organisationtype, data=data_d1, family=poisson, subset=(Presence) >= 1)
summary(model6_pubd1)

#Extract the results from R
write.csv(getSummary.mer(model6_pubd1)$coef,"model6_pubd1.csv")
