

Geenhuizen, Marina Van; Stek, Pieter

**Conference Paper**

## Mapping innovation in the global photovoltaic industry: a bibliometric approach to cluster identification and analysis

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Geenhuizen, Marina Van; Stek, Pieter (2015) : Mapping innovation in the global photovoltaic industry: a bibliometric approach to cluster identification and analysis, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124683>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Mapping Innovation in the Global Photovoltaic Industry: A Bibliometric Approach to Cluster Identification and Analysis

Pieter E. Stek and Marina van Geenhuizen

Faculty of Technology, Policy and Management, Delft University of Technology,  
Jaffalaan 5, 2628 BX Delft, The Netherlands. Email: p.e.stek@tudelft.nl

## Abstract

The innovation output of the photovoltaics sector has rapidly expanded during the past decade-and-a-half as investment have increased in an attempt to increase solar energy in the energy mix. Using a quantitative model based on bibliometric data, this study identifies clusters and explores the influence of global knowledge networks on innovation performance. Up to 129 clusters are identified around the world, of which less than 50 can be considered as major clusters. By extracting innovation input and output indicators from the bibliometric data (patent grants and scientific publications) and various forms of co-authorship or assignee-inventor relationships, a model of cluster innovation performance is constructed that incorporates input, output and knowledge network indicators. The estimation results of this model suggest that both agglomeration effects and relational proximity influence innovation performance, and that this occurs in a non-linear way. Furthermore, certain kinds of knowledge relations appear to reduce innovation performance while the presence of headquarters within a cluster appears to be beneficial. Although second-degree network effects are also explored in this study, their effect on innovation performance is not clear. A number of practical conclusions can also be drawn based on the results.

## 1 Introduction

This paper presents a quantitative assessment of the influence of global knowledge networks on the innovation of industry clusters in the photovoltaics sector. Photovoltaics is a knowledge intensive and globally distributed sector that incorporates multiple technological domains, including electronics and materials science, like nanotechnology. The photovoltaic sector has also received large inflows of public and private research investment during recent decades as part of a global push towards green growth. Photovoltaics by means of using solar energy are expected to contribute to the reduction of greenhouse gas emissions and the diversification of energy supply (Breyer, Birkner, Meiss, Goldschmidt, & Riede, 2013).

It is therefore not surprising that the innovation performance of the sector has been the subject of significant attention, both in the academic community and among policy makers. Academic research has generally focused on the photovoltaic sector in one or a small number of countries in the context of industrial policy, international technology transfer and international competition (de la Tour, Glachant, & Meniere, 2011; Grau, Huo, & Neuhoff, 2012; Kim & Kim, 2015; Klitkou & Godoe, 2013; Lo, Wang, & Huang, 2013; Vidican, McElvaney, Samulewicz, & Al-Saleh, 2012; Wu, 2014; Zheng & Kammen,

2014). There are also numerous studies of specific photovoltaic industry clusters which address similar aspects (Dewald & Fromhold-Eisebith, 2015; Klitkou & Coenen, 2013; Luo, Lovely, & Popp, 2013; West, 2014). And there is a small number of global studies (Breyer et al., 2013; Leydesdorff & Alkemade, 2014) which explore the size, growth and geographic distribution of innovation output in the photovoltaic sector.

This study focuses on the role of global knowledge networks in the photovoltaics sector. Global knowledge networks are networks of personal and institutional relationships that enable the transfer and transformation of knowledge on a global scale. Examples include the global research collaboration network (De Prato & Nepelski, 2014) and the global academic network (Barnett, Park, Jiang, Tang, & Aguillo, 2014). Other salient examples include the cross-border knowledge transfers that take place within multinational corporations (MNCs) which are enabled by MNCs private network of branch laboratories located in industry clusters worldwide (Castellani, Jimenez, & Zanfei, 2013).

While it is clear that international research activities are rapidly growing, especially in terms of their global distribution and the increasing importance of global knowledge networks (Audretsch, Lehmann, & Wright, 2014; Locke & Wellhausen, 2014), there are also indications that internationalization may weaken or limit the development of internal knowledge networks (Kwon, Park, So, & Leydesdorff, 2012; Van Geenhuizen & Nijkamp, 2012; Ye, Yu, & Leydesdorff, 2013), and lower the innovation performance of some knowledge intensive industries. However it is unclear what the effects of a shift from local towards global knowledge networks is on the innovation performance of industry clusters, including in the photovoltaics sector.

To address this knowledge gap a quantitative model is proposed to assess the influence of multiple (local and global) knowledge networks on the innovation performance of industry clusters. This multi-scalar approach is based on recent re-evaluation of the influence of spatial proximity and of the influence of relational proximity in innovation, and on a multidimensional understanding of knowledge networks as consisting of different types of relationships. In this study four types of knowledge network are investigated: personal collaboration networks, institutional collaboration networks (Dodgson, 1992), Triple Helix (university-government-industry relations) networks (Etzkowitz & Leydesdorff, 2000) and MNC branch networks. Together these networks provide a broad and differentiated perspective on relational proximity.

The study makes various contributions to the understanding of innovation. It makes a theoretical contribution by exploring how relational proximity (global networks) influences spatially concentrated knowledge creation in industry clusters, including the possibility of knowledge flows mainly in one direction, which might raise the innovation performance of the receiving industry cluster at the expense of the sending cluster. This latter option is considered to be a possibility for the Triple Helix and MNC branch networks, which are conceptualized as directed networks.

The study makes a methodological contribution by using only bibliographic data to develop a model of innovation performance, combining spatial analysis, scientometric network analysis and the knowledge production function. This approach involves the identification of industry clusters (independent of administrative boundaries), the extraction of cluster innovation input and output indicators and the construction of multiple inter-cluster knowledge networks from which network indicators are extracted. These innovation input, output and network indicators are then used to estimate a quantitative model of innovation performance. Because only bibliographic data are used, the method has the potential to be

applied in many other knowledge-intensive sectors, especially when innovation data in terms of production for the market are not available. And the study also provides practical results that can help regional policy makers and cluster managers optimize the external relations policy of their clusters to enhance innovation performance.

The study is centered on the following research questions:

1. What is the pattern of global development of the photovoltaic sector in terms of growth in innovation output, geographic distribution and the prevalence of knowledge networks during the past decade?
2. What is the influence of global knowledge networks (personal collaboration, institutional collaboration, Triple Helix, MNC branch) on the innovation performance of industry clusters in the photovoltaic sector?

These research questions are first discussed on the basis of the current literature (section 2), which forms the basis for the research model and hypotheses (section 3). This is followed by a description of the data and methodology (section 4) which precedes the results and analysis (section 5). The final section (6) of the paper contains a discussion and conclusion.

## **2 Literature review**

This review touches on the significant theoretical and empirical advances that have recently been made in relation to global knowledge networks and innovation performance. The section begins by discussing the idea of spatial and relational proximity and theory on agglomeration economies and regional innovation. This is followed by three more empirically-focused sections on the differentiation in knowledge relations, the mapping and analysis of knowledge networks and incorporating network indicators into cluster models.

### *Spatial and relational proximity*

Regional approaches to international research interaction have been dominated for years by a rather one-dimensional approach in which spatial proximity enhances knowledge spillovers (Acs, Audretsch, & Feldman, 1994; Audretsch & Feldman, 1996). In the course of the past decades, an increased attention has been developed for other types of proximity (Boschma, 2005; Breschi & Lissoni, 2001) and these include among others networks of cooperation in knowledge production (relational proximity) (Asheim & Isaksen, 2002; Bathelt, Malmberg, & Maskell, 2004). Relational proximity may be defined as the capability of regions or clusters (and their organizations and firms) to learn through cooperation with other regions or clusters (Camagni & Capello, 2002). Relational proximity is seen as being facilitated by socio/cultural similarity that enables the absorption of knowledge spillovers from other regional contexts, through a common set of values and beliefs (Fazio & Lavecchia, 2013). In addition, an approach which puts an emphasis on synergies between different proximities, like between spatial and relational proximity has also been developed (Ponds, Van Oort, & Frenken, 2007; Ponds, van Oort, & Frenken, 2010).

### *Agglomeration economies*

The influence of spatial proximity on innovation performance under the label of agglomeration economies, has been studied theoretically and empirically for decades (Capello, 2009). The first source of these localized advantages originates from being spatially close in larger cities or clusters, including lower transport and transmission costs, proximity to final markets (for firms) or test/launching markets (for innovations), a larger chance for meeting of two agents eventually leading to serendipity, and

easier exchange of creative ideas (Morgan, 2004). The second source of advantages draws on productivity increases due to cost reductions (scale effect) and localized accumulation of production skills (labor market). A third source of agglomeration advantages originates from synergy and refers to the rise of a set of common values and beliefs which in fact act as the economic rationale for the reduction of transaction costs (Williamson, 1981). In so-far the advantages deal with knowledge, the appropriate concept is localized knowledge spillovers, and invention and innovative activity at universities, research institutes and companies in cities or clusters are regarded as benefiting from them (Acs et al., 1994; Anselin, Varga, & Acs, 1997; Audretsch & Feldman, 1996; Autant-Bernard, 2001; Jaffe, 1989). However, various doubt have been casted on the condition of spatial proximity in productive inter-organizational learning, summarized in the assumption that spatial proximity is neither a necessary nor a sufficient condition for creative learning and innovation. Rather, it merely facilitates (Boschma, 2005; D'Este, Guy, & Iammarino, 2012; Karlsson, 2010).

### *Relational economies*

The idea that 'advantages of spatial proximity' also work on a distance and in a similar manner compared to the ones that are localized, has emerged in the literature since the early 2000s and has been increasingly elaborated ever since (Ertur & Koch, 2011). Thus, Breschi and Lissoni (2003) argue that collaborative networks are channels for knowledge spillovers that are not limited to local environments, instead, they can span long physical distances (Maggioni, Nosvelli, & Uberti, 2007; Maggioni & Uberti, 2009; Ponds et al., 2007). According to this line of thinking, the study of regional invention and innovation has shifted from a focus on close territorial relationships towards an emphasis on technological collaboration that increasingly occurs between cities or clusters as widely spread network-based systems through which knowledge circulates, is transformed and enriched (Cohendet & Amin, 2006). And in the past years, it has been increasingly recognized that relational proximity between the organizations involved (whether local or global) is key in productive collaborative learning (e.g. Bathelt et al., 2004; Cook, 2007).

Often such a situation has been viewed as merely positive in enhancing innovation. Particularly in high-tech sectors, research collaboration through global networks has been regarded as crucial for corporate innovative performance, like in the biotechnology industry (Cooke, 2007; Gertler & Levitte, 2005). What however might occur, if local firms are strongly collaborating with Multinational Corporations (MNCs) from elsewhere or if they are established or acquired by such companies, is that these local firms develop knowledge strategies depending on their role in the production organization of the (parent) MNC. Particularly, the role of producing knowledge for the MNC means that MNCs learn from their foreign subsidiaries, which is named 'reverse' knowledge transfer in some studies (Ambos, Ambos, & Schlegelmilch, 2006; Awate, Larsen, & Mudambi, 2014; Castellani et al., 2013; Dunning, 2000; Frost & Zhou, 2005; Frost, 2001). Under these conditions, foreign MNCs extract knowledge from the cluster and if this is based on exclusivity, the cluster might weaken instead of grow due to global research interaction.

Firms within a cluster can leverage their relationships with partners inside and outside of the cluster. In smaller clusters which lack the local networking advantages of agglomeration economies, it is likely that knowledge relations with partners outside of the cluster play a more prominent role (Tödtling & Trippel, 2005). At the same time, the strong relational proximity of a cluster to other clusters may be strengthening the agglomeration effect, as large clusters tend to be important nodes in national and global knowledge networks (Bathelt et al., 2004), for example in biotechnology (Huallacháin & Lee, 2014). However, as the example of knowledge relations within MNCs shows, the type of knowledge

and the type of knowledge relationship influences a cluster's innovation performance.

#### *Differentiation in knowledge and knowledge relations*

In addition to considering the spatial dimension (local relationship inside the cluster and relationships between organizations in different clusters), there are significant differences in type of knowledge and knowledge relations. Most prominent is perhaps the 'classic' distinction between tacit and codified knowledge (Polanyi, 1958) and its empirical application (e.g. Gertler, 2003; Simmie, 2003), both of which play an important role in the innovation process. Because codified knowledge is easier to communicate, it is understandable that formal research collaboration between clusters has been identified as a factor that enhances regional innovation performance (Huallacháin & Lee, 2014), but this is not evident from formal within-cluster research collaboration (Fritsch, 2004). Within-cluster collaboration may involve primarily tacit exchanges of knowledge, which are not captured by formal research collaboration indicators such as co-invented patents.

In fact research collaboration, which has often been measured through co-authorship only accounts for a small share of the total knowledge transfer that takes place between institutions, including between university and industry (Bekkers & Bodas Freitas, 2008; Bukvova, 2010). In addition to this, research collaboration is sensitive to 'power realities' between the research partners (Hervas-Oliver & Albers-Garrigos, 2013; Van Geenhuizen & Nijkamp, 2012), with a stronger partner often exerting more influence over how the relationship is conducted and how potential benefits are appropriated. Even within MNCs, larger labs in more prominent clusters tend to have significant autonomy over how and what kind of research they conduct (Mudambi & Navarra, 2004).

Patterns of knowledge relations may also vary depending on the industry sector (Iammarino & McCann, 2006; Jensen, Johnson, Lorenz, & Lundvall, 2007), the knowledge resource base and the social capital of the region (Masciarelli, Laursen, & Prencipe, 2010; Tödtling & Trippl, 2005). Finally the type of actors involved in the collaboration, such as universities, government or industry, and user groups (Etzkowitz & Leydesdorff, 2000; Ranga & Etzkowitz, 2011).

In the context of the spatial and relational proximity discussed earlier, and in the context of the knowledge networks discussed below, this suggests that a plurality of relations and networks co-exist which may influence innovation performance in different ways.

#### *Mapping and analysis of knowledge networks*

The diffusion of computing power and the internet, and easier access to databases of scientific documents, have made it possible to map the geographic distribution and network relations of scientific output on a worldwide basis (Leydesdorff & Alkemade, 2014; Leydesdorff & Persson, 2010). In addition to mapping at the micro-level (i.e. individual addresses or institutions), relational knowledge networks such as co-invention networks have also been mapped and analyzed at higher geographic agglomeration levels, such as at the level of urban areas (Huallacháin & Lee, 2014) or countries (De Prato & Nepelski, 2014). These studies reveal that network indicators such as degree centrality (Wasserman & Faust, 1994) correlate closely to the size of the cluster or of a country's innovation output.

Weblinks between university websites have also been used as an indicator of relational knowledge networks (Barnett et al., 2014), and network indicators such as degree centrality show a statistically significant correlation to the size and academic reputation of the institution. Interestingly, physical

distance is not identified as a statistically significant factor in any of the studies mentioned in this paragraph, thus reaffirming the importance of relational proximity in knowledge networks. However there appears to be a hierarchy in knowledge networks, with national sub-networks in which a few dominant national institutions maintain most international relations (Ortega & Aguillo, 2009).

Triple Helix relations between university, industry and government actors (Etzkowitz & Leydesdorff, 2000) have also been analyzed in a number of studies, providing an evolutionary and comparative perspective on (national) innovation systems. Of note is the finding that international collaboration, as measured by co-authorship of scientific papers, tends to weaken domestic Triple Helix relations (Kwon et al., 2012; Leydesdorff & Sun, 2009; Ye et al., 2013), although the effect of this collaboration shift on innovation growth and performance is unclear.

#### *Incorporating network indicators into a cluster model*

Given the predictive potential shown by network indicators, incorporating them into a model that predicts innovation performance is a logical next step. Several models for understanding innovation systems have been proposed at the national scale (Edquist, 1997; Freeman, 1995; Lundvall, 1992; Nelson, 1993), regional scale (Cooke, 2001; Porter, 1998; Tödting & Trippel, 2005) and at the sectoral scale (Malerba & Orsenigo, 1997; Porter, 1998). All of these models incorporate the three core Triple Helix actor groups of university, industry and government (Etzkowitz & Leydesdorff, 2000), in addition to many other actor groups such as knowledge intensive business services that can be viewed as intermediaries (Hertog, 2000; Schlierf & Meyer, 2013).

In addition to intermediaries, there is a growing literature exploring the influence of factors such as entrepreneurship (Autio, Kenney, Mustar, Siegel, & Wright, 2014) and the contribution of universities to entrepreneurship (Van Looy et al., 2011; Perkmann et al., 2013) on cluster innovation performance, to name but a few recent studies. However, to include network and other indicators in the model and to then estimate it, quantitative innovation performance models tend to be reduced to a knowledge production function in which the influence of a smaller number of factors is studied. This approach has led to insights such as: the institutional and policy factors that influence national innovation performance (de Rassenfosse & van Pottelsberghe de la Potterie, 2009; Furman, Porter, & Stern, 2013), the geographic distribution of university-industry knowledge spillovers within a country (Ponds et al., 2010) and the low importance of formal research collaboration within clusters to explain cluster innovation performance (Fritsch, 2004). So although the knowledge production function typically represents a simplification, it can be considered as a well-accepted methodology.

When certain factors influence innovation at different spatial levels (e.g. national vs. local), then the estimation of the knowledge production function (or similarly expressed models) benefits from multilevel regression analysis, which allows these effects to be incorporated into the model (Fischer & Getis, 2010; Fotheringham, Brunson, & Charlton, 2003; Raspe & van Oort, 2009; Raspe & van Oort, 2011).

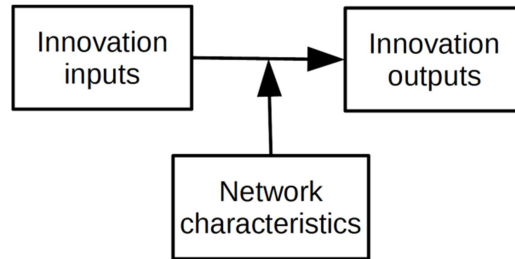
### **3 Research model and hypotheses**

In this section the research model is presented followed by the formulation of a number of hypotheses.

#### *Research Model*

The model used in this study is described in figure 1. Innovation outputs are the dependent variable, while innovation inputs are the independent variables, and network characteristics are regarded as

moderating variables. The respective innovation and network indicators are described in more detail in section 3.



**Figure 1:** Research model

The model is mathematically expressed as a knowledge production function (Jones, 1995; Romer, 1990). The basic form of the knowledge production function is given in equation 1.

$$P_i = \delta L_i^\lambda \tag{1}$$

Here  $P_i$  is the innovation output in cluster  $i$ ,  $L_i$  is the innovation input in cluster  $i$ ,  $\delta$  is the innovation propensity and  $\lambda$  the research productivity. Innovation propensity is understood to be determined by the national science and technology policy environment (de Rassenfosse & van Pottelsberghe de la Potterie, 2009) as well as the specific technological and competitive features of the sector (Arundel & Kabla, 1998; Kleinknecht, Montfort & Brouwer, 2002), which have an influence of the production of intellectual property, specifically patents. Patent-based indicators are used to measure innovation output in this study. Because this study considers only a single sector, innovation propensity is assumed to vary only by country.

If researchers innovate equally well, then there should be no differences in research productivity between clusters. However the literature cited in the previous section suggests that researcher's social network and various institutional factors do influence research productivity. Equation 1 can be transformed as equation 2. Here  $\varepsilon_i$  is an error term for each cluster  $i$ .

$$\log P_i = \log \delta + \lambda \log L_i + \varepsilon_i \tag{2}$$

Equation 2 can be re-written as equation 3 by taking into account the following. Innovation propensity varies by country  $m$ . Innovation output  $P_i$  can be decomposed into the number of claims in patent grants ( $CLM$ ) and the number of patent citations ( $PCT$ ). Innovation input  $L_i$  can be decomposed into the number of scientific publication citations ( $ACT$ ) and the number of active researchers ( $RES$ ). Network indicators ( $NET$ ) act as moderating variables and various different network indicators are to be tested.  $\beta$  represent the model parameters that are to be estimated.

$$\log(CLM_i * PCT_i) = \beta_m + \beta_1 NET_i * (\log(RES_i * ACT_i)) \tag{3}$$



### *Hypotheses*

The hypotheses presented here are based on results of the literature review (section 2). The first hypothesis presented here aims to evaluate evidence of an agglomeration economy in the photovoltaic sector:

*H1* Cluster size correlates positively to cluster innovation performance

The other hypotheses consider relational economies by addressing the correlation between cluster centrality and innovation performance. If relational economies are conceived as derived from networks, then clusters can be seen as vertices, and the relationships maintained by researchers within those clusters with researchers from other clusters as the networks' edges. There are various measures of centrality in networks, the simplest of which is degree centrality: the number of direct connections a cluster (vertex) has with other clusters (Wasserman & Faust, 1994). Other measures include betweenness centrality, which is the number of shortest paths between all vertices in a network that pass through a particular vertex; closeness centrality, which is an inverse measure of the total number of shortest paths between a vertex and all other vertices in the network; and eigenvector centrality, which is a measure of how closely connected a vertex is to other well-connected vertices (Wasserman & Faust, 1994).

Therefore two pairs of hypotheses about network centrality are formulated, the first pair, H2a and H2b, test the influence of direct network links on innovation performance. These direct links are typically the focus of the literature on inter-organizational learning (Dodgson, 1992, 1993).

*H2a* Network degree centrality correlates positively to cluster innovation performance

*H2b* Network degree centrality does not correlate to cluster innovation performance

The second pair of hypotheses, H3a and H3b, considers the fact that the existence of links between two clusters can affect the establishment or existence of links between another pair of clusters. This is often the case if the network has a hierarchical structure, meaning that some vertices are significantly better connected than others. A hierarchical network structure has been found in international co-invention networks (De Prato & Nepelski, 2014) and the academic weblink network (Barnett et al., 2014).

*H3a* Network betweenness, closeness or eigenvector centrality correlate positively to cluster innovation performance

*H3b* Network betweenness, closeness or eigenvector centrality do not correlate to cluster innovation performance

The hypotheses concerning network centrality are intentionally ambiguous with respect to the different types of networks. In this study we consider the following network types:

1. Inter-personal innovation networks based on patent co-invention
2. Inter-organizational innovation networks based on patent co-assignment
3. Research networks based on scientific paper co-affiliation
4. Triple Helix networks of university, industry and government relations based on patent co-assignment and scientific paper co-affiliation (directed network with knowledge flowing from university/government to industry)

5. Branch networks between inventors and firms headquartered outside the cluster (directed network with knowledge flowing from inventors to headquarters outside the cluster)

In addition to these networks, the possibility of weighted and non-weighted networks is considered. For weighted networks the “thickness” of the relationships are also explored (i.e. how frequently do knowledge interactions occur?), whereas in non-weighted networks only the number of cluster connections are considered. In effect, the four hypotheses about the relational economy are 40 sub-hypotheses if one considers the various network types. As with other innovation indicators (Hagedoorn & Cloudt, 2003), we expect a high degree of correlation between the various knowledge networks and their respective indicators.

#### 4 Data, methodology and indicators

This study considers the 2000-2009 period, which is the most recent decade for which detailed bibliometric data is available (see below). In this section the data, methodology, indicator development and the model estimation strategy are described.

##### *Data*

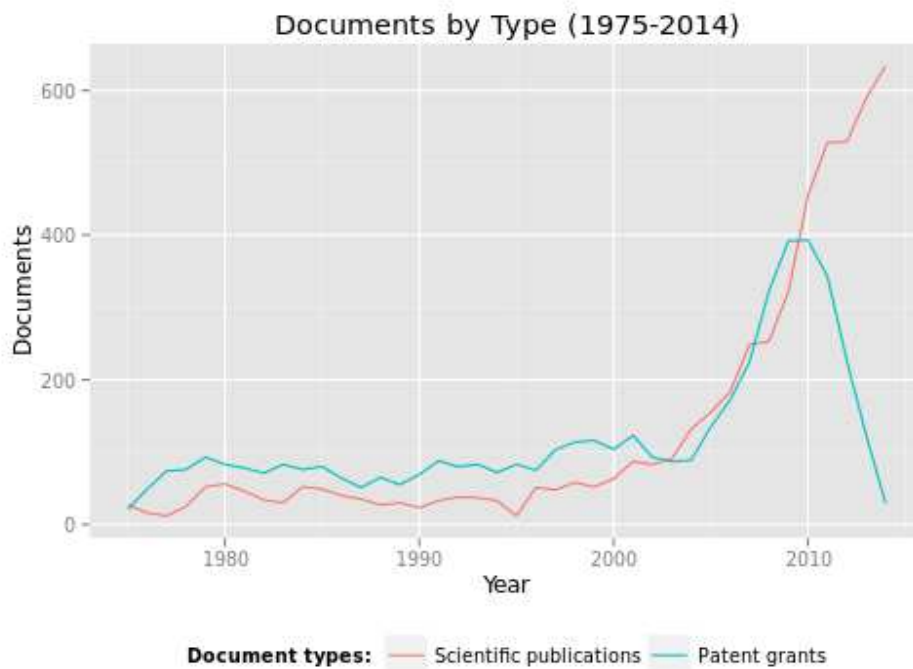
This study is based exclusively on bibliographic sources (patents and scientific publications) which enable the observation of changes over longer time periods, while offering global coverage at a local scale by using the address information contained in the bibliographic sources. These bibliographic sources form a “paper trail” of innovation activity (Jaffe, Trajtenberg, & Henderson, 1993), the data from which can be used to estimate the knowledge production function. On the one hand the use of bibliometric data as an innovation indicator has disadvantages, including variations in patenting propensity between sectors (Kleinknecht et al., 2002). On the other hand bibliometric indicators such as patent counts and citation counts in high-technology industries tend to show close statistical overlap with other innovation indicators such as R&D inputs and new product announcements, which are also used to measure innovation performance (Hagedoorn & Cloudt, 2003). A summary of the data sources used in this study is provided in table 1.

<b>Features</b>	<b>Patent grants</b>	<b>Scientific publications</b>
Home address	Available	None
Institutional address	Available	Available
Inbound citations data	Available	Available
Publication lag	up to 5 years, sometimes more	None
Most active institutions	Industry	University
Data source (for this study)	United States Patent & Trademark Office	Scopus® by Elsevier
Selection criterion	Y02E/50 (Cooperative Patent Classification)	“photovoltaic cell” (keyword search of abstracts)
Documents selected (this study)	5,524	5,270

**Table 1:** Patent grants and scientific publication data used in the study

For this study, bibliometric data is obtained from patent grants from the United States Patent & Trademark Office (USPTO) and scientific publications are downloaded from the *Scopus*® database, which is maintained by Elsevier, an academic publisher. The USPTO data was accessed through its website<sup>1</sup> on 17 July 2015 using automated means. Elsevier has an application programming interface (API) for automated access to its *Scopus* database. All data processing, as well as all model estimation is carried out using R (R Core Team, 2015), using *rvest* (Wickham, 2015) to download to bulk data and communicate with the *Scopus* API and *RMySQL* (Ooms, James, DebRoy, Wickham, & Horner, 2015) to communicate with and populate a MySQL database of bibliographic data. All R scripts in this study are available for download<sup>2</sup>.

An overview of the number of patent grants and scientific publications during the 1975-2014 period is given in figure 2. This clearly shows the rapid increase in innovation output during the 2000-2009 period, which is being studied. Note that the “decline” in patent grants is due to their publication time-lag: many patent applications filed after 2009 are still pending. Next, one of the challenges of this research is the delineation of both the sector (technological) and cluster domains.



**Figure 2:** Documents in data set, by type

### Sector Identification

For the delineation of the sector (technology) domain we rely on the new Y02E cooperative patent classification (CPC) from the USPTO and the European Patent Office (EPO) which contains renewable energy technologies, including a subcategory for photovoltaics (Y02E/50). For scientific publications

<sup>1</sup> Available at: <http://patft.uspto.gov/netahtml/PTO/search-adv.htm>

<sup>2</sup> Available at: <http://stek.in/papers/tha2015bj.zip>

we are not aware of such a classification, and so rely on a keyword search of journal paper abstracts which contain the phrases “photovoltaic cell”. Inclusion of the word “cell” is important to distinguish from papers about astronomy, which are not directly related to the knowledge base of the industry.

#### *Cluster identification*

“Clusters are geographic concentrations of industries related by knowledge, skills, inputs, demand, and/or other linkages” (Delgado, Porter, & Stern, 2014). Given the bibliometric data in this study and a focus on innovation performance, which is just one cluster activity, a cluster identification strategy that focuses on innovation output concentration is used.

Patents and scientific publications contain (partial) addresses of the affiliated institutions (including assignees) and addresses of inventors. These addresses can be assigned coordinates using commercial mapping applications, such as the Bing Maps API, which is used for the initial screening, and the website [findlatitudedeandlongitude.com](http://findlatitudedeandlongitude.com) for the second screening in case of irregularities. This approach is not original and has been used previously, although with a smaller dataset (Leydesdorff & Alkemade, 2014; Leydesdorff & Persson, 2010). As these mapping applications and the bibliographic data are not perfect, some semi-manual corrections need to be made. These include removing or re-searching addresses that only yield country coordinates (only acceptable for small countries and territories such as Singapore, Bermuda, Hong Kong or Luxembourg), and dealing with clearly erroneous and frequently occurring classifications, such as “Yokohama, Japan”, where the Yokohama in Aomori prefecture (a town with less than 5,000 inhabitants) is frequently identified instead of Yokohama in Kanagawa prefecture (a city with 3.7 million inhabitants). In total 97% of addresses are identified using the Bing Maps API. But in actual use the accuracy is higher because more documents carry addresses of large cities, which have a higher chance of being correctly identified than, say, a small village.

The conversion of addresses into coordinates allows the locations of document authors to be plotted on a map. Based on a plot of these document-authors, clusters are identified by using the standard “heat map” algorithm, formally known as kernel density estimation (Parzen, 1962; Rosenblatt, 1956). In this study a quartic (biweight) kernel shape is used and a cell size of approximately 36 km<sup>2</sup>. One concern of any clustering methodology is robustness and therefore we identify two sets of clusters. The first set are “cities” and are identified by making a heat map with a radius of 50 km and a density threshold of 7.5 document authors per cell. This yields 129 clusters which produced both scientific publications and patent grants during the study period. These 129 clusters hold 93% of all document-authors. The second set of clusters are “regions” and are identified using the same process, but with a radius of 100 km and density threshold of 15. Using this method 69 clusters are identified which hold 92% of all document-authors. We consider both the “city” and “regional” cluster sets, to allow for comparisons.

#### *Indicator development*

Finally, we discuss the indicators, which are summarized in table 2. The innovation indicators are largely based on previous studies. Patent claims (*CLS*) and patent citations (*PCT*), along with the number of countries in which a patent has been applied for ('family size') have been used as indicators of patent quality and are therefore preferable as knowledge output indicators than simple patent counts (Lanjouw & Schankerman, 2004). The number of researchers (*RES*) has been used as a knowledge input indicator in various studies, while academic paper citation counts (*ACT*) are widely used in various university rankings<sup>3</sup> to measure the quality and impact of scientific output.

---

<sup>3</sup> Examples include: QS University Rankings (UK), Times Higher Education University Rankings (UK),

Indicator	Description and measurement	“City clusters” (n = 129)	“Regional clusters” (n = 69)
<i>RES</i>	Unique researchers listed on patent applications	minimum = 1 median = 8 mean = 50.54 maximum = 1,119	1 20 96.74 1,854
<i>ACT</i>	Scientific publication citations	minimum = 0 median = 191 mean = 1,086 maximum = 25,892	0 422 1921 26,938
<i>PCT</i>	Patent citations derived from USPTO full patent grant web application	minimum = 0 median = 49 mean = 923 maximum = 29,364	0 130 1,756 57,830
<i>CLM</i>	Patent claims derived from patent grant documents	minimum = 1 median = 240 mean = 1,564 maximum = 33,904	3 680 3'045 37,612
<i>NET<sub>IP</sub></i>	Inter-personal innovation network derived from the bidirectional co-invention network as extracted from patent grants	vertices = 76 edges = 294 (w) 147 (n) density = 0.10 (w) 0.05 (n)	vertices = 47 edges = 261 (w) 106 (n) density = 0.24 (w) 0.10 (n)
<i>NET<sub>IO</sub></i>	Inter-organizational innovation network derived from the bidirectional co-assignment network as extracted from patent grants	vertices = 19 edges = 20 (w) 15 (n) density = 0.12 (w) 0.09 (n)	vertices = 20 edges = 22 (w) 17 (n) density = 0.12 (w) 0.09 (n)
<i>NET<sub>RE</sub></i>	Research network derived from the bidirectional co-affiliation network as extracted from scientific publications	vertices = 73 edges = 108 (w) 82(n) density = 0.04 (w) 0.03 (n)	vertices = 45 edges = 88 (w) 66 (n) density = 0.09 (w) 0.07 (n)
<i>NET<sub>TH</sub></i>	Triple Helix network derived from the directional university/government → industry network as extracted from patent grants	vertices = 17 edges = 13 (w) 11 (n) density = 0.05 (w) 0.04 (n)	vertices = 17 edges = 16 (w) 13 (n) density = 0.06 (w) 0.05 (n)
<i>NET<sub>BR</sub></i>	The branch network is derived from the directional inventor → assignee network as extracted from patent grants	vertices = 88 edges = 536 (w) 197 (n) density = 0.07 (w) 0.03 (n)	vertices = 53 edges = 484 (w) 152 (n) density = 0.18 (w) 0.06 (n)

**Table 2:** Model indicators and statistical summaries, w = weighted network, n = non-weighted network

Further, network indicators are derived for the five networks that are analyzed in this study (inter-personal, inter-organizational, research, Triple Helix and branches), in both weighted and non-weighted form. Degree, closeness, betweenness and eigenvector centrality are calculated; for directed networks in-degree, out-degree, in-closeness and out-closeness centralities are also considered (Wasserman & Faust, 1994). All network indicators are calculated using the igraph package in R (Csardi & Nepusz, 2006).

A pairwise correlation analysis is conducted between all indicators which is provided in table 7 in the appendix. The main findings based on this correlation analysis and the statistical summary in table 2, are as follows:

1. Weighted and non-weighted network indicators tend to have a strong positive correlation, so the use of separate indicators is not necessary (in the remaining paper non-weighted network indicators are used).
2. The inter-organizational ( $NET_{IO}$ ) and Triple Helix ( $NET_{TH}$ ) networks are too small, and so cannot be used for any generalizations. They are therefore excluded from further analysis.
3. 42 city clusters and 32 regional clusters are part of the three remaining large networks ( $NET_{IP}$ ,  $NET_{RE}$  and  $NET_{BR}$ ) and also have a complete nonzero set of innovation input and output indicators. The correlation between the above three networks' centrality indicators also varies. There is especially significant variation in correlation between closeness centrality and degree centrality (including in/out-degree and in/out-closeness centralities). Other centrality indicators such as betweenness and eigenvector centrality tend to be positively correlated to degree centrality.

#### *Model estimation strategy*

The model estimation is attempted using a multilevel regression analysis (also known as mixed effect modeling). As with other parts of the methodology, this process is carried out in R, with the aid of the lme4 package (Bates, Machler, Bolker & Walker, 2014). Because multilevel regression analysis is being used, a step-wise estimation is conducted in which new models are compared using analysis of variance (ANOVA). Models that show a statistically significant reduction in variance are subsequently adopted as the core model. The two levels considered in this multilevel regression analysis are the country level and the cluster level.

## **5 Results and analysis**

The results of the model estimation, accompanied by analysis, are presented here in two parts, the model estimation results and the evaluation of the hypotheses.

#### *Model estimation results*

The model estimation follows a step-wise selection process, which is described in more detail in table 8. The first estimation round is used to assess whether the model will benefit from a multi-level regression approach instead of simple regression analysis. The results of the first round suggest that there the model improvement of a multi-level regression model are negligible, both for the city and regional cluster sets. This outcome is significant in itself because it suggests that national factors do not have a strong influence on the innovation performance of photovoltaic clusters. However the analysis nevertheless continues with multi-level regression because of the sound theoretical reasons for including national factors in the model.

In the next estimation round 10 knowledge network indicators are tested, along with the agglomeration effect, where the number of researchers in the cluster, *RES*, serves as the indicator. The knowledge network indicators selected are degree and closeness centrality for the three largest networks (inter-personal, branch and research). This selection is made based on the earlier pair-wise correlation analysis (see table 7 in the appendix). Based on the statistical significance of the second-round indicators, combinations of indicators are tested in the third round. Using ANOVA analysis a “winning” model is then chosen.

For the city cluster set, the ANOVA analysis of the third round reveals three “winning” models (MC3.1, MC3.5 and MC3.7). The same situation occurs for the regional cluster set, also with three “winning” models (MR3.3, MR3.5 and MR3.7). Of these models MC3.7 and MR3.7 incorporate all three knowledge networks and both degree centrality and closeness centrality indicators, while the others include only one (inter-personal network) or two (researcher and branch) knowledge networks.

To further explore this outcome, a fourth estimation round is undertaken by which the “winning” models of the third round are compared to a model in which only direct network effects are included (degree centrality), a model in which only secondary network effects are included (closeness centrality) and a model where the degree centrality and closeness centrality are inverses compared to the third-round models. The fourth round ANOVA analysis is inconclusive for the city cluster set, with the model incorporating only direct network effects (MC4.2) also comparing favorably. For the regional cluster set the ANOVA yields a clear “winner”, which is model MR4.2, which like MC4.2 only incorporates direct network effects. To better understand this estimation outcome, the four “best models” for the city and the regional cluster sets respectively are presented in table 3. If among these a “general model” had to be chosen it should be MC4.2/MR4.2, which have the best fit for both the city and the regional cluster sets.

<b>MC3.1<sup>^</sup></b>		<b>MC3.5<sup>^</sup></b>		<b>MC3.7<sup>^</sup></b>		<b>MC4.2<sup>^</sup></b>	
Indicators	Coefficients	Indicators	Coefficients	Indicators	Coefficients	Indicators	Coefficients
intercept	6.97	intercept	8.38	intercept	7.21	intercept	9.30
RES	0.127	RES	0.123	RES	0.128	RES	0.126
IP-DEG	-0.529	IP-DEG	-0.683	IP-DEG	-0.975	IP-DEG	-0.870
				BR-ICLS	3,267	BR-IDEG	0.596
		RE-CLS	-1,958	RE-CLS	-972	RE-DEG	-1.430
<b>MR3.3</b>		<b>MR3.5</b>		<b>MR3.7</b>		<b>MR4.2<sup>^</sup></b>	
Indicators	Coefficients	Indicators	Coefficients	Indicators	Coefficients	Indicators	Coefficients
intercept	9.96	intercept	15.4	intercept	16.0	intercept	8.29
RES	0.0841	RES	0.0858	RES	0.0825	RES	0.0607
		IP-CLS	-681	IP-CLS	-881	IP-DEG	0.487
				BR-ICLS	278	BR-IDEG	0.236
RE-DEG	-2.31	RE-DEG	-2.10	RE-DEG	-1.84	RE-DEG	-1.97

Acronyms: degree centrality (DEG), closeness centrality (CLS), in-degree centrality (IDEG).

<sup>^</sup> marks model(s) of best fit.

**Table 3:** Selected estimation results

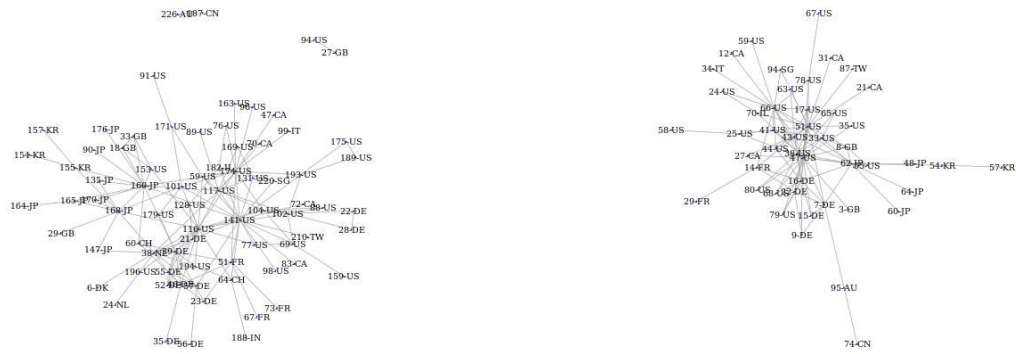
Although similar in many ways, the models also differ in terms of the factors included and, in some cases, the signs and value of the estimated coefficients. Models MC3.1, MC3.5 and MC3.7 incorporate the degree centrality of the inter-personal network and the closeness centrality of the research network. But for MR 3.3, MR3.5 and MR3.7 the opposite is true, the closeness centrality of the inter-personal network and the degree centrality of the research network are incorporated. Since degree centrality accounts for direct (first-order) network effects and closeness centrality for indirect (second-order) network effects, this change depending on the network scale is notable. However it should not be over-interpreted given that simpler models that only include direct network effects (MC4.2 and MR4.2) also have a strong fit.

The signs and values of the coefficients are broadly similar across the MC- and MR3-series models in table 3. This provides a degree of validation as to the influence of the different network indicators on innovation performance. However for MR4.2 the sign of the inter-personal degree centrality (IP-DEG) coefficient changes. Focusing on the interpretation of the estimated coefficients, it is clear that branch in-degree centrality (BR-IDEG) has a positive effect on cluster innovation performance. Thus the hosting of headquarters raises cluster innovation performance, which is not surprising given the existing literature. Although the absence of branch out-degree centrality in the model suggests that there is no evidence to suggest that the branches themselves also support cluster innovation performance.

Furthermore, the presence of negative coefficients for inter-personal network centrality and research network centrality indicators may seem surprising because there is significant theoretical evidence to support the benefits of relational proximity. However, both factors can perhaps be better understood when considering the fact that there is a consistent positive relationship between cluster scale (in terms of the number of researchers) and innovation performance, and when viewing the structures of the respective networks. It is also important to note that smaller clusters tend to have a higher degree centrality relative to their size. The correlation between size (*RES*) and inter-personal network degree centrality *per researcher* is -0.428 and for research network degree centrality *per researcher* the correlation is -0.368 in the city cluster set.

Starting with the network structure, the inter-personal knowledge network (see figure 3) at the “city scale” (50 km radius) has several “main” clusters with greater network centrality, which are surrounded by smaller “satellite” clusters with low network centrality, which are typically only connected to the main cluster. The smaller satellite clusters may depend on the main clusters for innovation inputs, or may be “suppliers” of innovation inputs to the main cluster, and therefore the negative sign of the network indicators can be interpreted as a response to small scale, i.e. “in small clusters there are fewer researchers, innovation performance is lower, and knowledge networking with other clusters plays a more important role”. However on the “regional scale” (100 km radius, see figure 3 again), the satellite clusters appear to have been absorbed by the main clusters, and in this case the presence of a strong inter-personal knowledge network (see MR4.2) *does* appear to support innovation performance, and thus supports the relational proximity thesis.

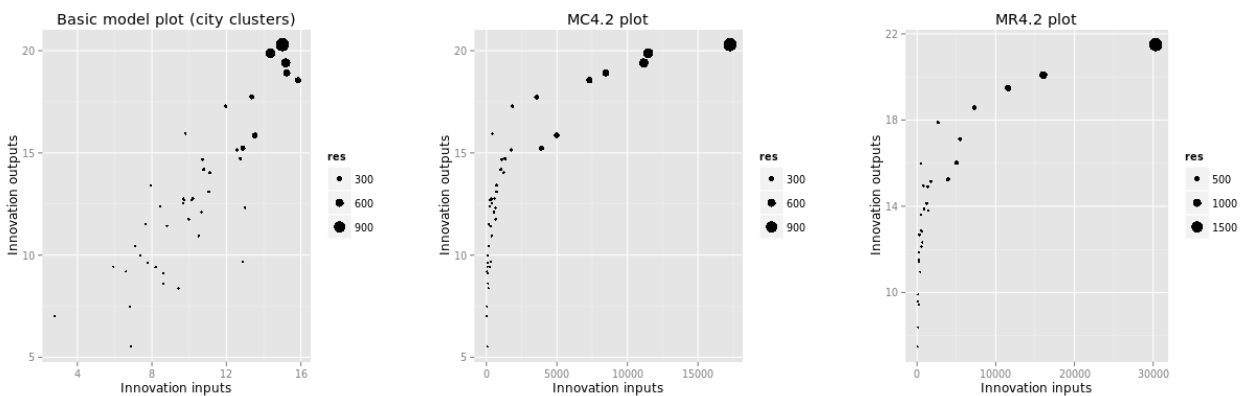




**Figure 3:** Inter-personal knowledge network for city clusters (left) and regional clusters (right)

The network structure of the city cluster set has an interesting core-periphery structure with national sub-networks, which is also observed in the university weblinks network identified by Barnett et al. (2014). In terms of structure, the city cluster set also has commonalities with the global scientific collaboration network (Leydesdorff, Wagner, & Adams, 2013), in which there is also a core-periphery structure. However the regional cluster set lacks such a core-periphery division, its network is instead more distributed with most clusters connected to more than one other clusters. This suggests that the regional cluster scale of 100 km fails to capture peripheral clusters and that, provided sufficient data is available, a smaller spatial scale adds value to the analysis, also within countries.

Referring once more to the agglomeration effect, it is worthwhile to note that all estimated models have logarithmic qualities, which appears to account for an important part of their improvement of fit compared to the baseline knowledge production function (without agglomeration and network indicators), see figure 4.



**Figure 4:** Model plots of basic model (left), MC4.2 (center) and MR4.2 (right)

The logarithmic pattern suggests that, although cluster size raises innovation performance, its effects start to diminish as a cluster grows beyond a certain size. The reasons for such an “excess agglomeration penalty” are not immediately clear from the dataset, but it is not an uncommon situation in urban economics witnessing increasing land prices and wages, as dimensions of limited accessibility and congestion (Richardson, 1995; Tabuchi, 1998; Torre & Wallet, 2014).

*Evaluation of hypotheses*

The results described above provide sufficient ground to evaluate the hypotheses formulated in section 3. Although not all knowledge networks were evaluated (the inter-organizational and branch networks were too small), all hypotheses find some degree of support in the results, as summarized in table 4.

<b>Hypothesis</b>	<b>Content</b>	<b>Evaluation</b>
H1	Agglomeration effect	Accept
H2a & H2b	Degree centrality (first-degree effect)	Accept H2a, reject H2b
H3a & H3b	Closeness centrality (second-degree effect)	Accept H3a with caution for research and branch networks, accept H3b

**Table 4:** Evaluation of hypotheses

While the agglomeration effect is clearly supported by the results (H1 accepted), the logarithmic relationship between innovation inputs and innovation outputs suggests that for very large clusters the increasing benefits of agglomeration may not always be out-weighed by rising costs. There is clear evidence of direct (first-degree) network effects across all knowledge networks, supporting the concept of relational proximity (H2a is accepted, H2b rejected). The results also show some evidence that second-degree network effects may play a role, as is the case with closeness centrality in research, inter-personal and branch networks in some of the models (H3a accepted with caution), but it is not included in all models, and a second-degree network effect-only model has not been accepted based on ANOVA (H3b also accepted). In the next section the significance of these findings are further discussed.

**6 Discussion and conclusion**

The purpose of this paper is to present a quantitative assessment of the influence of global knowledge networks on the innovation performance of industry clusters in the photovoltaics sector. This study has provided a number of insights which will be highlighted here, including: which factors are and are not significant to cluster innovation performance and how this relates to the existing theory. This is followed by a discussion of the practical implications of the results for cluster managers, the limitations of the study and the conclusion.

The results suggest that the spatial scale of a cluster, the network type and the network structure all influence innovation performance in subtly different ways. If clusters are identified on a smaller scale, then the agglomeration effect appears to be stronger, and the influence of inter-personal and researcher networks tends to be negative. However, on a larger scale the agglomeration effect is weaker and the influence of relational proximity (network indicators) is mixed.

The mixed or negative outcomes of network indicators are difficult to understand unless one considers the possibility of a power imbalance (or a “benefits imbalance”) between the partners. In the case of the inbound branch network, having high network in-degree centrality is clearly beneficial to a cluster. A

high in-degree centrality signifies the presence of headquarters within the cluster. The benefits of headquarters are many, including having access to knowledge in other clusters through “knowledge pipelines” (Bathelt et al., 2004; Castellani et al., 2013; Gertler, 2003). The cluster may also benefit from closer interaction with top decision-makers, which can improve innovation performance by better allocating resources to the most promising areas of research, an effect that has previously been noted in large cities (Sassen, 2002).

Having noted previously the negative correlation between cluster size and network degree centrality *per researcher*, it is likely that participants in a research project from a smaller cluster are frequently invited to participate in research, but are not part of the core group of researchers driving an innovation process. If they were part of the core group, it would make more sense for them to relocate to the main cluster. Therefore, while they contribute to a research project that is mainly taking place in a core cluster, the overall benefits, tangible and intangible, of such research, accrue unequally to their home cluster. At a larger scale this “small cluster effect” seems to dissipate as smaller clusters are either not detected or incorporated into a larger cluster. Hence for the regional cluster set, inter-personal network degree centrality contributes positively to cluster innovation performance, as is evident from model MR4.2. This positive contribution fits within the broader theoretical framework of relational proximity.

In addition to the above it is noteworthy that the country-level tend not to influence cluster innovation performance even though the data set includes clusters in advanced economies such as the United States and a major emerging economy, China. The model intercept (which is how the country-level is incorporated into the models) are all very close to the mean. For example, in the MR4.2 model, the mean is 8.29, with China scoring 8.39 and the United States 8.82. So while China may have fewer clusters than the United States, the innovation performance of these clusters is at a similar level. Admittedly this conclusion is partly influenced by the data sources, which are international. Poorer quality research and researchers from China may simply fail to appear in the data set. Yet the lack of variation between countries in the model suggests that cluster factors such as network centrality influence a cluster's development more strongly than the national innovation system.

The study provides a number of practical insights for cluster managers. First is the trend of clear benefits of having headquarters (i.e. commercial control) located inside clusters. While this does not confirm the presence of a “reverse knowledge flow” from branches to headquarters, the benefits of hosting a headquarters are clearly shown, while there are no significant benefits of hosting the branches. Furthermore, there is close correlation between the network indicators of weighted and non-weighted knowledge networks. This suggests that successful research collaborations (which result in the granting of a patent or scientific publication) tend to occur repeatedly with partners from the same cluster. So although there are certainly benefits from having many weak links (Granovetter, 1973), clusters tend to sustain only a limited number of deep inter-cluster relationships, with small clusters sustaining relationships with just one or two other clusters.

And finally, the results suggest that there is a strong and positive agglomeration effect. This suggests that a city or region's economic policy should be somewhat focused, instead of making small investments in many different sectors. But there is a limit to the agglomeration effect, beyond around 500 researchers (in the bibliometric sense, and not necessarily equivalent to “real-life” researchers), the agglomeration effect dissipates. Naturally the above is purely from the perspective of innovation performance; the economic performance and growth of clusters often does not depend on innovation performance alone.

Although this study shows that a bibliometric model of industry clusters and knowledge networks is feasible, and holds great promise for use in other studies in which bibliometrics is one or the only consistent source of data, the study also shows the limitations of this methodology. First of all some knowledge networks, such as the inter-organizational and Triple Helix networks, only appear in very large clusters and so there is a limitation in the study of smaller clusters. Second, this study has only addressed the photovoltaics cluster which were identified at two different geographic scales. Although a simple and transparent method for cluster identification, the use of bibliometrics means that only one dimension of the sector and one dimension of innovation particularly the close relation to invention activity, are measured (Carlino and Kerr, 2014). It is therefore likely that the method is “blind” to clusters which may have significance in terms of production or production-related R&D but not in terms of more general R&D that is eventually published as patents or peer-reviewed scientific publications. Since only photovoltaics was explored, different sectors may display significantly different agglomeration and network effects.

In conclusion, this study has shown that the innovation output of the photovoltaics sector has been growing rapidly during the past decades and that photovoltaics research is concentrated in less than 50 clusters that are distributed around the world, and that these clusters are connected through multiple non-identical knowledge networks. A quantitative model of innovation performance based on bibliometric data (patent grants and scientific publications) suggests that both agglomeration and relational proximity affect cluster innovation performance. Relational proximity is found in different types of knowledge networks (inter-personal, research, branch), and incorporating multiple knowledge networks into the model allows for the development of a more precise model.

## Appendix

### “City cluster” (50 km) innovation indicators

Cluster	Country (ISO-3166)	<i>RES</i>	<i>ACT</i>	<i>PCT</i>	<i>CLM</i>
101-US	US	142	1,108	6,779	4,748
102-US	US	249	2,536	5,021	10,016
116-US	US	502	8,278	7,407	22,278
117-US	US	121	2,396	1,170	3,229
128-US	US	91	500	1,259	1,893
131-US	US	15	79	55	626
135-JP	JP	20	108	513	197
141-US	US	691	5,729	7,929	33,904
155-KR	KR	357	2,098	622	12,434
159-US	US	39	464	2,820	2,978
160-JP	JP	1,119	2,939	29,364	22,303
165-JP	JP	9	621	194	47
168-JP	JP	773	2,220	28,297	15,244
169-US	US	7	792	41	135
170-JP	JP	32	506	305	1,149
174-US	US	425	17,886	9,708	11,807
175-US	US	52	519	677	520
182-IL	IL	43	997	81	2,228
187-CN	CN	30	1,221	28	2,002
193-US	US	23	291	97	943
196-US	US	24	195	521	459
21-DE	DE	35	453	325	870
210-TW	TW	297	1,329	466	8,900
220-SG	SG	13	952	17	255
226-AU	AU	19	905	473	680
24-NL	NL	2	8	53	21
29-GB	GB	5	148	31	317
35-DE	DE	20	181	15	819
38-NL	NL	89	3,872	882	2,813
39-DE	DE	68	41	413	1,627
46-DE	DE	50	431	108	1,181
51-FR	FR	87	763	907	1,365
55-DE	DE	80	600	1,283	1,141
56-DE	DE	5	191	2	125
59-US	US	40	11,030	434	515
60-CH	CH	7	54	97	130
64-CH	CH	15	25,892	31	524
70-CA	CA	7	335	156	97
72-CA	CA	7	224	94	231
73-FR	FR	63	993	291	1,682
98-US	US	25	1,007	421	780
99-IT	IT	4	228	11	162

**“Regional cluster” (100 km) innovation indicators**

Cluster	Country (ISO-3166)	<i>RES</i>	<i>ACT</i>	<i>PCT</i>	<i>CLM</i>
14-FR	FR	93	763	942	1,480
15-DE	DE	114	600	1,441	2,054
16-DE	DE	342	7,943	2,720	10,013
17-US	US	41	11,030	438	516
22-DE	DE	87	26,938	382	2,574
24-US	US	45	591	854	948
27-CA	CA	21	335	230	627
34-IT	IT	4	274	11	162
35-US	US	26	1,007	421	783
38-US	US	195	1,108	6,898	8,527
41-US	US	73	238	1,800	1,727
43-US	US	121	2,397	1,170	3,229
47-US	US	934	11,396	14,548	36,479
48-JP	JP	20	108	513	197
51-US	US	716	5,808	8,296	34,865
54-KR	KR	363	2,098	723	12,561
58-US	US	46	1,225	2,903	3,055
60-JP	JP	15	621	207	70
62-JP	JP	1,854	5,159	57,830	37,612
63-US	US	10	792	93	217
64-JP	JP	42	518	330	1,210
66-US	US	428	17,886	9,817	11,852
67-US	US	53	1,107	706	521
7-DE	DE	71	599	477	2,232
70-IL	IL	44	1,092	81	2,284
74-CN	CN	32	1,221	28	2,044
78-US	US	23	291	97	943
8-GB	GB	85	4,138	504	2,720
80-US	US	26	195	580	539
87-TW	TW	299	1,336	466	8,938
9-DE	DE	21	181	15	831
94-SG	SG	13	952	17	255
95-AU	AU	19	929	473	680

**Table 5:** Cluster innovation indicators

**“City cluster” (50 km) network indicators\***

Cluster	BR- DEG	BR- IDEG	BR- ODEG	BR- CLS	BR- ICLS	BR- OCLS	BR- BTW	BR- EVC	IP- DEG	IP- CLS	IP- BTW	IP- EVC	RE- DEG	RE- CLS	RE- BTW	RE- EVC
101-US	8	4	4	0.02%	0.11%	0.02%	87	38.0%	6	0.22%	120	41.0%	3	0.09%	69	15.4%
102-US	15	8	7	0.02%	0.11%	0.02%	389	54.3%	8	0.21%	193	34.4%	1	0.02%	0	0.0%
116-US	38	26	12	0.02%	0.12%	0.02%	1,282	100%	19	0.23%	567	100%	8	0.09%	440	76.7%
117-US	9	5	4	0.02%	0.11%	0.02%	307	30.3%	12	0.22%	326	67.2%	3	0.08%	32	16.9%
128-US	12	7	5	0.02%	0.11%	0.02%	227	43.8%	4	0.20%	12	21.0%	1	0.08%	0	18.8%
131-US	3	1	2	0.02%	0.10%	0.02%	0	15.2%	3	0.20%	8	20.7%	1	0.02%	0	0.0%
135-JP	4	2	2	0.02%	0.10%	0.02%	0	13.5%	2	0.20%	0	12.2%	1	0.08%	0	6.4%
141-US	33	18	15	0.02%	0.11%	0.02%	1,088	82.6%	20	0.23%	721	86.8%	6	0.09%	356	59.6%
155-KR	6	3	3	0.02%	0.10%	0.02%	61	19.7%	4	0.20%	139	12.6%	4	0.09%	81	43.8%
159-US	2	1	1	0.02%	0.11%	0.02%	0	13.1%	1	0.18%	0	3.0%	1	0.02%	0	0.0%
160-JP	15	9	6	0.02%	0.11%	0.02%	446	44.1%	15	0.22%	610	57.4%	8	0.09%	524	58.8%
165-JP	2	0	2	0.02%	0.01%	0.02%	0	6.8%	3	0.20%	70	12.4%	1	0.08%	0	14.4%
168-JP	15	11	4	0.02%	0.11%	0.02%	298	35.2%	10	0.21%	242	39.8%	6	0.09%	315	59.2%
169-US	3	2	1	0.02%	0.11%	0.02%	129	11.3%	2	0.19%	0	11.0%	2	0.08%	33	19.2%
170-JP	2	0	2	0.02%	0.01%	0.02%	0	6.8%	2	0.20%	0	12.2%	2	0.08%	58	15.4%
174-US	24	13	11	0.02%	0.11%	0.02%	886	78.4%	16	0.23%	496	74.3%	7	0.09%	257	69.5%
175-US	4	1	3	0.02%	0.10%	0.02%	116	16.4%	2	0.18%	0	3.6%	1	0.09%	0	24.5%
182-IL	4	3	1	0.02%	0.11%	0.02%	41	17.7%	4	0.20%	15	24.4%	1	0.02%	0	0.0%
187-CN	2	0	2	0.03%	0.01%	0.03%	0	8.6%	1	0.02%	0	0.0%	4	0.09%	151	64.2%
193-US	5	2	3	0.02%	0.10%	0.02%	47	21.7%	5	0.20%	151	25.4%	2	0.08%	2	8.9%
196-US	4	4	0	0.01%	0.12%	0.01%	0	13.1%	1	0.20%	0	12.5%	1	0.08%	0	18.8%
21-DE	5	1	4	0.02%	0.10%	0.02%	8	26.0%	4	0.21%	18	27.7%	1	0.08%	0	0.9%
210-TW	4	3	1	0.02%	0.10%	0.02%	118	11.3%	1	0.20%	0	10.9%	2	0.08%	23	18.4%
220-SG	2	1	1	0.02%	0.10%	0.02%	0	14.1%	2	0.20%	0	20.2%	1	0.08%	0	10.2%
226-AU	1	1	0	0.01%	0.01%	0.01%	0	0.7%	1	0.02%	0	0.0%	2	0.08%	0	1.7%
24-NL	1	0	1	0.02%	0.01%	0.02%	0	2.3%	1	0.18%	0	6.7%	1	0.08%	0	7.4%
29-GB	1	0	1	0.02%	0.01%	0.02%	0	3.0%	1	0.18%	0	5.0%	1	0.02%	0	0.0%
35-DE	2	0	2	0.03%	0.01%	0.03%	0	12.1%	2	0.18%	4	6.1%	4	0.09%	182	49.7%
38-NL	6	2	4	0.02%	0.11%	0.02%	51	27.4%	11	0.21%	224	53.1%	2	0.08%	58	3.8%
39-DE	13	7	6	0.02%	0.11%	0.02%	226	40.4%	12	0.21%	161	59.9%	2	0.02%	1	0.0%
46-DE	11	5	6	0.02%	0.11%	0.02%	273	46.8%	9	0.21%	99	43.8%	1	0.02%	0	0.0%
51-FR	14	11	3	0.02%	0.11%	0.02%	352	41.5%	7	0.21%	181	36.6%	4	0.09%	225	30.1%
55-DE	13	8	5	0.02%	0.11%	0.02%	143	42.0%	6	0.21%	6	40.7%	1	0.02%	0	0.0%
56-DE	1	0	1	0.02%	0.01%	0.02%	0	3.6%	2	0.18%	4	4.9%	2	0.08%	0	1.7%
59-US	3	2	1	0.02%	0.11%	0.02%	2	12.3%	5	0.21%	62	42.5%	3	0.09%	171	30.5%
60-CH	1	0	1	0.02%	0.01%	0.02%	0	4.0%	3	0.20%	9	20.2%	2	0.08%	0	1.7%
64-CH	3	1	2	0.03%	0.01%	0.03%	43	15.7%	4	0.21%	74	27.6%	4	0.09%	487	50.1%
70-CA	1	0	1	0.02%	0.01%	0.02%	0	2.6%	1	0.19%	0	8.4%	2	0.08%	18	24.2%
72-CA	1	0	1	0.03%	0.01%	0.03%	0	7.0%	1	0.20%	0	10.9%	1	0.08%	0	5.4%
73-FR	1	0	1	0.02%	0.01%	0.02%	0	3.5%	1	0.18%	0	4.6%	1	0.08%	0	7.4%
98-US	8	7	1	0.02%	0.11%	0.02%	89	35.2%	1	0.20%	0	10.9%	1	0.08%	0	6.4%
99-IT	2	0	2	0.03%	0.01%	0.03%	0	13.7%	1	0.19%	0	9.3%	2	0.08%	0	1.7%

**“Regional cluster” (100 km) network indicators\***

Cluster	BR- DEG	BR- IDEG	BR- ODEG	BR- CLS	BR- ICLS	BR- OCLS	BR- BTW	BR- EVC	IP- DEG	IP- CLS	IP- BTW	IP- EVC	RE- DEG	RE- CLS	RE- BTW	RE- EVC
14-FR	11	6	5	0.09%	0.44%	0.09%	127.4	35.3%	5	0.96%	47.9	29.9%	3	0.46%	41.0	36.0%
15-DE	11	5	6	0.09%	0.44%	0.09%	22.2	44.3%	3	0.92%	0.0	23.6%	1	0.05%	0.0	0.0%
16-DE	18	9	9	0.09%	0.45%	0.09%	128.7	59.0%	13	1.15%	66.9	68.0%	6	0.52%	96.6	60.3%
17-US	3	2	1	0.09%	0.41%	0.09%	1.7	9.7%	5	1.00%	12.5	36.8%	3	0.48%	49.7	37.2%
22-DE	4	0	4	0.10%	0.04%	0.10%	0.0	20.8%	7	1.01%	16.8	41.3%	5	0.51%	51.9	68.3%
24-US	3	2	1	0.09%	0.43%	0.09%	1.0	10.3%	2	0.79%	0.0	10.8%	1	0.44%	0.0	14.2%
27-CA	2	0	2	0.10%	0.04%	0.10%	0.0	10.2%	2	0.88%	0.0	17.8%	1	0.44%	0.0	14.2%
34-IT	2	0	2	0.10%	0.04%	0.10%	0.0	10.2%	1	0.79%	0.0	7.5%	2	0.44%	0.0	17.2%
35-US	9	7	2	0.09%	0.46%	0.09%	76.2	35.8%	2	0.93%	0.0	21.0%	1	0.38%	0.0	5.5%
38-US	13	8	5	0.09%	0.46%	0.09%	79.8	50.8%	10	1.12%	51.8	59.0%	3	0.42%	42.0	12.9%
41-US	7	3	4	0.09%	0.41%	0.09%	18.4	26.9%	5	0.98%	7.5	31.0%	1	0.46%	0.0	20.5%
43-US	9	6	3	0.09%	0.46%	0.09%	148.6	32.2%	10	1.12%	32.2	60.9%	3	0.41%	24.9	9.8%
47-US	42	26	16	0.09%	0.53%	0.09%	793.4	100%	24	1.41%	314.9	100%	8	0.54%	228.6	69.0%
48-JP	2	1	1	0.09%	0.38%	0.09%	0.0	7.1%	1	0.78%	0.0	7.0%	1	0.38%	0.0	4.4%
51-US	30	17	13	0.09%	0.49%	0.09%	452.1	73.8%	21	1.37%	294.7	89.7%	6	0.52%	111.6	52.2%
54-KR	5	2	3	0.09%	0.37%	0.09%	1.4	21.5%	2	0.79%	45.0	7.1%	5	0.48%	77.4	34.1%
58-US	2	1	1	0.09%	0.42%	0.09%	0.0	15.5%	1	0.67%	0.0	2.7%	1	0.43%	0.0	12.4%
60-JP	2	0	2	0.10%	0.04%	0.10%	0.0	3.8%	1	0.78%	0.0	7.0%	1	0.46%	0.0	20.5%
62-JP	15	10	5	0.09%	0.46%	0.09%	249.5	45.7%	13	1.20%	234.4	63.6%	10	0.56%	280.4	100%
63-US	3	2	1	0.09%	0.44%	0.09%	94.3	10.4%	2	0.87%	0.0	16.7%	2	0.41%	21.4	10.6%
64-JP	2	1	1	0.09%	0.04%	0.09%	0.0	3.8%	1	0.78%	0.0	7.0%	1	0.46%	0.0	20.5%
66-US	22	11	11	0.09%	0.49%	0.09%	357.2	57.6%	16	1.22%	213.5	68.2%	6	0.49%	84.7	47.7%
67-US	4	1	3	0.09%	0.38%	0.09%	56.9	18.2%	1	0.68%	0.0	3.2%	2	0.45%	0.0	26.0%
7-DE	8	4	4	0.09%	0.43%	0.09%	27.0	37.5%	8	1.08%	104.9	45.2%	2	0.47%	2.9	25.6%
70-IL	4	3	1	0.09%	0.44%	0.09%	32.0	17.5%	4	0.98%	9.8	29.7%	1	0.05%	0.0	0.0%
74-CN	3	1	2	0.10%	0.04%	0.10%	7.5	8.4%	1	0.55%	0.0	0.6%	5	0.52%	61.9	64.7%
78-US	3	1	2	0.09%	0.32%	0.09%	0.0	14.9%	4	0.99%	45.0	28.9%	2	0.36%	1.5	4.2%
8-GB	4	0	4	0.10%	0.04%	0.10%	0.0	19.6%	5	1.01%	6.6	33.3%	6	0.53%	126.6	66.6%
80-US	4	3	1	0.09%	0.41%	0.09%	0.4	17.7%	2	0.88%	0.0	17.6%	1	0.44%	0.0	14.2%
87-TW	5	4	1	0.09%	0.43%	0.09%	80.0	12.9%	1	0.85%	0.0	9.9%	2	0.43%	9.2	12.7%
9-DE	3	0	3	0.10%	0.04%	0.10%	0.0	15.7%	3	0.78%	0.0	17.1%	4	0.48%	60.5	36.1%
94-SG	2	1	1	0.09%	0.39%	0.09%	0.0	11.4%	2	0.89%	0.0	17.5%	1	0.43%	0.0	11.1%
95-AU	3	1	2	0.10%	0.04%	0.10%	0.5	4.2%	2	0.74%	45.0	5.1%	2	0.44%	0.0	17.2%

Network centrality acronyms: degree (DEG), in-degree (IDEG), out-degree (ODEG), closeness (CLS), in-closeness (ICLS), out-closeness (OCLS), betweenness (BTW), eigenvector (EVC) for branch (BR), inter-personal (IP) and research (RE) networks.

**Table 6:** Cluster knowledge network indicators



**“City cluster” set (50 km)**

	RES	ACT	PCT	CLM	BR-DEG	BR-IDEG	BR-ODEG	BR-CLS	BR-ICLS	BR-OCLS	BR-BTW
RES	1.000	0.238	0.902	0.902	0.686	0.676	0.637	0.075	0.391	0.075	0.676
ACT	0.238	1.000	0.171	0.276	0.343	0.314	0.361	0.084	0.062	0.084	0.396
PCT	0.902	0.171	1.000	0.682	0.501	0.513	0.432	0.052	0.309	0.052	0.469
CLM	0.902	0.276	0.682	1.000	0.818	0.784	0.797	0.082	0.385	0.082	0.828
BR-DEG	0.686	0.343	0.501	0.818	1.000	0.979	0.938	0.096	0.540	0.096	0.968
BR-IDEG	0.676	0.314	0.513	0.784	0.979	1.000	0.849	0.025	0.563	0.025	0.940
BR-ODEG	0.637	0.361	0.432	0.797	0.938	0.849	1.000	0.207	0.447	0.207	0.922
BR-CLS	0.075	0.084	0.052	0.082	0.096	0.025	0.207	1.000	-0.043	1.000	0.100
BR-ICLS	0.391	0.062	0.309	0.385	0.540	0.563	0.447	-0.043	1.000	-0.043	0.434
BR-OCLS	0.075	0.084	0.052	0.082	0.096	0.025	0.207	1.000	-0.043	1.000	0.100
BR-BTW	0.676	0.396	0.469	0.828	0.968	0.940	0.922	0.100	0.434	0.100	1.000
BR-EVC	0.617	0.362	0.452	0.734	0.971	0.938	0.933	0.135	0.609	0.135	0.922
IP-DEG	0.718	0.390	0.565	0.781	0.894	0.839	0.902	0.152	0.512	0.152	0.875
IP-CLS	0.280	0.187	0.224	0.271	0.372	0.355	0.365	0.417	0.440	0.417	0.330
IP-BTW	0.822	0.401	0.644	0.883	0.862	0.812	0.864	0.112	0.424	0.112	0.883
IP-EVC	0.599	0.431	0.452	0.680	0.868	0.827	0.853	0.149	0.559	0.149	0.835
RE-DEG	0.801	0.513	0.697	0.774	0.697	0.692	0.637	0.105	0.197	0.105	0.736
RE-CLS	0.168	0.200	0.137	0.138	-0.006	0.014	-0.041	-0.054	-0.194	-0.054	0.068
RE-BTW	0.730	0.665	0.656	0.694	0.611	0.614	0.545	0.130	0.146	0.130	0.635
RE-EVC	0.632	0.528	0.526	0.639	0.551	0.548	0.502	0.120	0.087	0.120	0.605

	BR-EVC	IP-DEG	IP-CLS	IP-BTW	IP-EVC	RE-DEG	RE-CLS	RE-BTW	RE-EVC
RES	0.617	0.718	0.280	0.822	0.599	0.801	0.168	0.730	0.632
ACT	0.362	0.390	0.187	0.401	0.431	0.513	0.200	0.665	0.528
PCT	0.452	0.565	0.224	0.644	0.452	0.697	0.137	0.656	0.526
CLM	0.734	0.781	0.271	0.883	0.680	0.774	0.138	0.694	0.639
BR-DEG	0.971	0.894	0.372	0.862	0.868	-0.697	0.006	0.611	0.551
BR-IDEG	0.938	0.839	0.355	0.812	0.827	0.692	0.014	0.614	0.548
BR-ODEG	0.933	0.902	0.365	0.864	0.853	-0.637	0.041	0.545	0.502
BR-CLS	0.135	0.152	0.417	0.112	0.149	-0.105	0.054	0.130	0.120
BR-ICLS	0.609	0.512	0.440	0.424	0.559	-0.197	0.194	0.146	0.087
BR-OCLS	0.135	0.152	0.417	0.112	0.149	-0.105	0.054	0.130	0.120
BR-BTW	0.922	0.875	0.330	0.883	0.835	0.736	0.068	0.635	0.605
BR-EVC	1.000	0.878	0.411	0.816	0.869	-0.630	0.079	0.548	0.473
IP-DEG	0.878	1.000	0.436	0.943	0.965	-0.729	0.005	0.639	0.516
IP-CLS	0.411	0.436	1.000	0.372	0.510	-0.142	0.060	0.187	0.023
IP-BTW	0.816	0.943	0.372	1.000	0.871	0.819	0.153	0.728	0.620
IP-EVC	0.869	0.965	0.510	0.871	1.000	-0.654	0.023	0.584	0.437
RE-DEG	0.630	0.729	0.142	0.819	0.654	1.000	0.363	0.903	0.873
RE-CLS	-0.079	-0.005	-0.060	0.153	-0.023	0.363	1.000	0.328	0.517
RE-BTW	0.548	0.639	0.187	0.728	0.584	0.903	0.328	1.000	0.833
RE-EVC	0.473	0.516	0.023	0.620	0.437	0.873	0.517	0.833	1.000

**“Regional cluster” set (100 km)**

	RES	ACT	PCT	CLM	BR-DEG	BR-IDEG	BR-ODEG	BR-CLS	BR-ICLS	BR-OCLS	BR-BTW
RES	1.000	0.281	0.936	0.920	0.647	0.670	0.565	-0.070	0.360	-0.070	0.664
ACT	0.281	1.000	0.188	0.319	0.399	0.328	0.481	0.150	0.023	0.150	0.391
PCT	0.936	0.188	1.000	0.750	0.429	0.465	0.344	-0.038	0.272	-0.038	0.459
CLM	0.920	0.319	0.750	1.000	0.827	0.838	0.753	-0.037	0.396	-0.037	0.834
BR-DEG	0.647	0.399	0.429	0.827	1.000	0.982	0.957	-0.007	0.500	-0.007	0.953
BR-IDEG	0.670	0.328	0.465	0.838	0.982	1.000	0.884	-0.116	0.581	-0.116	0.956
BR-ODEG	0.565	0.481	0.344	0.753	0.957	0.884	1.000	0.160	0.340	0.160	0.881
BR-CLS	-0.070	0.150	-0.038	-0.037	0.007	-0.116	0.160	1.000	-0.758	1.000	-0.028
BR-ICLS	0.360	0.023	0.272	0.396	0.500	0.581	0.340	-0.758	1.000	-0.758	0.453
BR-OCLS	-0.070	0.150	-0.038	-0.037	0.007	-0.116	0.160	1.000	-0.758	1.000	-0.028
BR-BTW	0.664	0.391	0.459	0.834	0.953	0.956	0.881	-0.028	0.453	-0.028	1.000
BR-EVC	0.617	0.382	0.425	0.770	0.966	0.942	0.933	-0.042	0.563	-0.042	0.866
IP-DEG	0.674	0.527	0.492	0.822	0.936	0.906	0.915	0.030	0.451	0.030	0.888
IP-CLS	0.599	0.469	0.452	0.696	0.808	0.804	0.757	-0.166	0.548	-0.166	0.741
IP-BTW	0.796	0.412	0.630	0.912	0.899	0.882	0.861	0.042	0.407	0.042	0.900
IP-EVC	0.622	0.517	0.466	0.741	0.872	0.854	0.839	-0.045	0.499	-0.045	0.802
RE-DEG	0.792	0.546	0.672	0.785	0.644	0.593	0.676	0.263	0.092	0.263	0.635
RE-CLS	0.318	0.311	0.265	0.330	0.209	0.171	0.253	0.295	-0.214	0.295	0.262
RE-BTW	0.883	0.423	0.787	0.855	0.677	0.661	0.653	0.156	0.178	0.156	0.693
RE-EVC	0.679	0.581	0.605	0.637	0.477	0.412	0.544	0.392	-0.104	0.392	0.469

	BR-EVC	IP-DEG	IP-CLS	IP-BTW	IP-EVC	RE-DEG	RE-CLS	RE-BTW	RE-EVC
RES	0.617	0.674	0.599	0.796	0.622	0.792	0.318	0.883	0.679
ACT	0.382	0.527	0.469	0.412	0.517	0.546	0.311	0.423	0.581
PCT	0.425	0.492	0.452	0.630	0.466	0.672	0.265	0.787	0.605
CLM	0.770	0.822	0.696	0.912	0.741	0.785	0.330	0.855	0.637
BR-DEG	0.966	0.936	0.808	0.899	0.872	0.644	0.209	0.677	0.477
BR-IDEG	0.942	0.906	0.804	0.882	0.854	0.593	0.171	0.661	0.412
BR-ODEG	0.933	0.915	0.757	0.861	0.839	0.676	0.253	0.653	0.544
BR-CLS	-0.042	0.030	-0.166	0.042	-0.045	0.263	0.295	0.156	0.392
BR-ICLS	0.563	0.451	0.548	0.407	0.499	0.092	-0.214	0.178	-0.104
BR-OCLS	-0.042	0.030	-0.166	0.042	-0.045	0.263	0.295	0.156	0.392
BR-BTW	0.866	0.888	0.741	0.900	0.802	0.635	0.262	0.693	0.469
BR-EVC	1.000	0.917	0.831	0.835	0.891	0.621	0.138	0.643	0.445
IP-DEG	0.917	1.000	0.906	0.918	0.971	0.725	0.297	0.722	0.569
IP-CLS	0.831	0.906	1.000	0.784	0.965	0.563	0.122	0.596	0.397
IP-BTW	0.835	0.918	0.784	1.000	0.831	0.736	0.335	0.763	0.587
IP-EVC	0.891	0.971	0.965	0.831	1.000	0.663	0.215	0.667	0.504
RE-DEG	0.621	0.725	0.563	0.736	0.663	1.000	0.541	0.949	0.919
RE-CLS	0.138	0.297	0.122	0.335	0.215	0.541	1.000	0.474	0.655
RE-BTW	0.643	0.722	0.596	0.763	0.667	0.949	0.474	1.000	0.866
RE-EVC	0.445	0.569	0.397	0.587	0.504	0.919	0.655	0.866	1.000

**Table 7:** Pair-wise correlation of innovation indicators and main network indicators

**Round Model and ANOVA significance levels**

- C1 mc1.1 <- lm(log(clm \* pct) ~ log(res \* act))  
mc1.2 <- lmer(log(clm \* pct) ~ log(res \* act) + (1|cou))
- C2 mc2.1 <- lmer(log(clm \* pct) ~ res \* log(res \* act) + (1|cou))\*\*\*  
mc2.21 <- lmer(log(clm \* pct) ~ br\_deg \* log(res \* act) + (1|cou))  
mc2.22 <- lmer(log(clm \* pct) ~ br\_ideg \* log(res \* act) + (1|cou))  
mc2.23 <- lmer(log(clm \* pct) ~ br\_odeg \* log(res \* act) + (1|cou))  
mc2.24 <- lmer(log(clm \* pct) ~ re\_deg \* log(res \* act) + (1|cou))

```

mc2.25 <- lmer(log(clm * pct) ~ ip_deg * log(res * act) + (1|cou))***
mc2.31 <- lmer(log(clm * pct) ~ br_cls * log(res * act) + (1|cou))
mc2.32 <- lmer(log(clm * pct) ~ br_ics * log(res * act) + (1|cou))***
mc2.33 <- lmer(log(clm * pct) ~ br_ocs * log(res * act) + (1|cou))
mc2.34 <- lmer(log(clm * pct) ~ re_cls * log(res * act) + (1|cou))***
mc2.35 <- lmer(log(clm * pct) ~ ip_cls * log(res * act) + (1|cou))

C3
mc3.1 <- lmer(log(clm * pct) ~ (res + ip_deg) * log(res * act) + (1|cou))***
mc3.2 <- lmer(log(clm * pct) ~ (res + br_ics) * log(res * act) + (1|cou))
mc3.3 <- lmer(log(clm * pct) ~ (res + re_cls) * log(res * act) + (1|cou))
mc3.4 <- lmer(log(clm * pct) ~ (res + br_ics + re_cls) * log(res * act) + (1|cou))
mc3.5 <- lmer(log(clm * pct) ~ (res + ip_deg + re_cls) * log(res * act) + (1|cou))***
mc3.6 <- lmer(log(clm * pct) ~ (res + ip_deg + br_ics) * log(res * act) + (1|cou))
mc3.7 <- lmer(log(clm * pct) ~ (res + ip_deg + br_ics + res_cls) * log(res * act) + (1|cou))***
mc3.8 <- lmer(log(clm * pct) ~ (ip_deg + br_ics + re_cls) * log(res * act) + (1|cou))

C4
mc4.1 <- lmer(log(clm * pct) ~ (res + inp_cls + br_ics + res_deg) * log(res * act) + (1|COU))
mc4.2 <- lmer(log(clm * pct) ~ (res + inp_deg + br_ideg + res_deg) * log(res * act) + (1|COU))***
mc4.3 <- lmer(log(clm * pct) ~ (res + inp_cls + br_ics + res_cls) * log(res * act) + (1|COU))

R1
mr1.1 <- lm(log(clm * pct) ~ log(res * act))
mr1.2 <- lmer(log(clm * pct) ~ log(res * act) + (1|cou))

R2
mr2.1 <- lmer(log(clm * pct) ~ res * log(res * act) + (1|cou))***
mr2.21 <- lmer(log(clm * pct) ~ br_deg * log(res * act) + (1|cou))
mr2.22 <- lmer(log(clm * pct) ~ br_ideg * log(res * act) + (1|cou))***
mr2.23 <- lmer(log(clm * pct) ~ br_odeg * log(res * act) + (1|cou))
mr2.24 <- lmer(log(clm * pct) ~ res_deg * log(res * act) + (1|cou))***
mr2.25 <- lmer(log(clm * pct) ~ inp_deg * log(res * act) + (1|cou))
mr2.31 <- lmer(log(clm * pct) ~ br_cls * log(res * act) + (1|cou))
mr2.32 <- lmer(log(clm * pct) ~ br_ics * log(res * act) + (1|cou))***
mr2.33 <- lmer(log(clm * pct) ~ br_ocs * log(res * act) + (1|cou))
mr2.34 <- lmer(log(clm * pct) ~ res_cls * log(res * act) + (1|cou))
mr2.35 <- lmer(log(clm * pct) ~ inp_cls * log(res * act) + (1|cou))***

R3
mr3.1 <- lmer(log(clm * pct) ~ (res + res_deg) * log(res * act) + (1|cou))***
mr3.2 <- lmer(log(clm * pct) ~ (res + br_ics) * log(res * act) + (1|cou))
mr3.3 <- lmer(log(clm * pct) ~ (res + inp_cls) * log(res * act) + (1|cou))
mr3.4 <- lmer(log(clm * pct) ~ (res + br_ics + inp_cls) * log(res * act) + (1|cou))
mr3.5 <- lmer(log(clm * pct) ~ (res + res_deg + inp_cls) * log(res * act) + (1|cou))***
mr3.6 <- lmer(log(clm * pct) ~ (res + res_deg + br_ics) * log(res * act) + (1|cou))
mr3.7 <- lmer(log(clm * pct) ~ (res + res_deg + br_ics + inp_cls) * log(res * act) + (1|cou))***
mr3.8 <- lmer(log(clm * pct) ~ (res_deg + br_ics + inp_cls) * log(res * act) + (1|cou))

R4
mr4.1 <- lmer(log(clm * pct) ~ (res + inp_deg + br_ics + res_cls) * log(res * act) + (1|COU))
mr4.2 <- lmer(log(clm * pct) ~ (res + inp_deg + br_ideg + res_deg) * log(res * act) + (1|COU))***
mr4.3 <- lmer(log(clm * pct) ~ (res + inp_cls + br_ics + res_cls) * log(res * act) + (1|COU))

```

Model acronyms: model for city cluster set (mc), model for regional cluster set (mr), linear model (lm), multi-level linear model (lmer), country (cou). For other acronyms refer to table 2 and table 6. Model notation adapted from R code. \*\*\* ANOVA  $\chi^2$  significance level of < 0.1%.

**Table 8:** Model estimation rounds

## References

- Acs, Z. J., Audretsch, D. B., & Feldman, M. P. (1994). R&D spillovers and innovative activity. *Managerial and Decision Economics*, 15(2), 131–138.
- Ambos, T. C., Ambos, B., & Schlegelmilch, B. B. (2006). Learning from foreign subsidiaries: An empirical investigation of headquarters' benefits from reverse knowledge transfers. *International Business Review*, 15(3), 294–312.
- Anselin, L., Varga, A., & Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics*, 42(3), 422–448.
- Arundel, A., & Kabla, I. (1998). What percentage of innovations are patented? empirical estimates for European firms. *Research Policy*, 27, 127–141. [http://doi.org/10.1016/S0048-7333\(98\)00033-X](http://doi.org/10.1016/S0048-7333(98)00033-X)
- Asheim, B., & Isaksen, A. (2002). Regional innovation systems: the integration of local “sticky” and global “ubiquitous” knowledge. *The Journal of Technology Transfer*, 1–14.
- Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American Economic Review*, 630–640.
- Audretsch, D. B., Lehmann, E. E., & Wright, M. (2014). Technology transfer in a global economy. *Journal of Technology Transfer*, 39(3), 301–312.
- Autant-Bernard, C. (2001). Science and knowledge flows: evidence from the French case. *Research Policy*, 30(7), 1069–1078.
- Autio, E., Kenney, M., Mustar, P., Siegel, D., & Wright, M. (2014). Entrepreneurial innovation: The importance of context. *Research Policy*, 43(7), 1097–1108.
- Awate, S., Larsen, M. M., & Mudambi, R. (2014). Accessing vs sourcing knowledge: A comparative study of R&D internationalization between emerging and advanced economy firms. *Journal of International Business Studies*.
- Barnett, G. A., Park, H. W., Jiang, K., Tang, C., & Aguillo, I. F. (2014). A multi-level network analysis of web-citations among the world's universities. *Scientometrics*, 99(1), 5–26.
- Bates, D., Machler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. Retrieved from <http://cran.r-project.org/package=lme4>
- Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28(1), 31–56.
- Bekkers, R., & Bodas Freitas, I. M. (2008). Analysing knowledge transfer channels between universities and industry: To what degree do sectors also matter? *Research Policy*, 37(10), 1837–1853.
- Boschma, R. (2005). Proximity and Innovation: A Critical Assessment. *Regional Studies*, 39, 61–74.
- Breschi, S., & Lissoni, F. (2001). Knowledge Spillovers and Local Innovation Systems: A Critical Survey. *Industrial and Corporate Change*, 10(4), 975–1005.
- Breyer, C., Birkner, C., Meiss, J., Goldschmidt, J. C., & Riede, M. (2013). A top-down analysis: Determining photovoltaics R&D investments from patent analysis and R&D headcount. *Energy Policy*, 62, 1570–1580.
- Bukvova, H. (2010). Studying Research Collaboration: A Literature Review, 10.
- Camagni, R., & Capello, R. (2002). Milieux Innovateurs and Collective Learning: From Concepts to Measurement. In Z. Acs, H. F. de Groot, & P. Nijkamp (Eds.), *The Emergence of the Knowledge Economy SE - 2* (pp. 15–45). Springer, Berlin Heidelberg.
- Capello, R. (2009). Spatial spillovers and regional growth: a cognitive approach. *European Planning Studies*, 17(5), 639–658.
- Carlino, G. & Kerr, W.R. (2014). Agglomeration and Innovation, Working Paper 15-007, Harvard Business School.
- Castellani, D., Jimenez, A., & Zanfei, A. (2013). How remote are R&D labs[quest] Distance factors and international innovative activities. *J Int Bus Stud*, 44(7), 649–675.
- Cohendet, P., & Amin, A. (2006). Epistemic communities and communities of practice in the knowledge-based firm1. In *New Frontiers in the Economics of Innovation and New Technology: Essays in Honour of Paul A. David* (p. 296). Edward Elgar Publishing.
- Cooke, P. (2001). Regional innovation systems, clusters, and the knowledge economy. *Industrial and Corporate Change*, 10(4), 945–974.
- Cooke, P. (2007). *Regional knowledge economies: markets, clusters and innovation*. Edward Elgar Publishing.
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Sy*, 1695. Retrieved from <http://igraph.org>
- D'Este, P., Guy, F., & Iammarino, S. (2012). Shaping the formation of university–industry research collaborations: what type of proximity does really matter? *Journal of Economic Geography*, lbs010.
- De la Tour, A., Glachant, M., & Meniere, Y. (2011). Innovation and international technology transfer: The case of the Chinese photovoltaic industry. *Energy Policy*, 39(2), 761–770.
- De Prato, G., & Nepelski, D. (2014). Global technological collaboration network: network analysis of international co-

- inventions. *Journal of Technology Transfer*.
- De Rassenfosse, G., & van Pottelsberghe de la Potterie, B. (2009). A policy insight into the R&D–patent relationship. *Research Policy*, 38(5), 779–792.
- Delgado, M., Porter, M. E., & Stern, S. (2014). *Defining clusters of related industries*. National Bureau of Economic Research.
- Dewald, U., & Fromhold-Eisebith, M. (2015). Trajectories of sustainability transitions in scale-transcending innovation systems: The case of photovoltaics. *Environmental Innovation and Societal Transitions*.
- Dodgson, M. (1992). The strategic management of R&D collaboration. *Technology Analysis & Strategic Management*, 4(3), 227–244.
- Dodgson, M. (1993). Learning, Trust, and Technological Collaboration. *Human Relations*, 46(1), 77–95.
- Dunning, J. H. (2000). The eclectic paradigm as an envelope for economic and business theories of MNE activity. *International Business Review*, 9(2), 163–190.
- Edquist, C. (1997). *Systems of innovation: technologies, institutions, and organizations*. Psychology Press.
- Ertur, C., & Koch, W. (2011). A contribution to the theory and empirics of Schumpeterian growth with worldwide interactions. *Journal of Economic Growth*, 16(3), 215–255.
- Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and “Mode 2” to a Triple Helix of university–industry–government relations. *Research Policy*, 29(2), 109–123.
- Fazio, G., & Lavecchia, L. (2013). Social capital formatio across space: proximity and trust in European regions. *International Regional Science Review*, 26(3), 296–321.
- Fischer, M., & Getis, A. (2010). Introduction. In M. M. Fischer & A. Getis (Eds.), *Handbook of Applied Spatial Analysis SE - 1* (pp. 1–24). Springer Berlin Heidelberg.
- Fotheringham, A. S., Brunson, C., & Charlton, M. (2003). *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons.
- Freeman, C. (1995). The “National System of Innovation” in historical perspective. *Cambridge Journal of Economics*, 19(1), 5–24.
- Fritsch, M. (2004). Cooperation and the efficiency of regional R&D activities. *Cambridge Journal of Economics*, 28(6), 829–846.
- Frost, T. S. (2001). The geographic sources of foreign subsidiaries’ innovations. *Strategic Management Journal*, 22(2), 101–123.
- Frost, T. S., & Zhou, C. (2005). R&D co-practice and “reverse”knowledge integration in multinational firms. *Journal of International Business Studies*, 36(6), 676–687.
- Furman, J. L., Porter, M. E., & Stern, S. (2013). The determinants of national innovative capacity, 31(2002), 899–933.
- Gertler, M. S. (2003). Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). *Journal of Economic Geography*, 3(1), 75–99.
- Gertler, M. S., & Levitte, Y. M. (2005). Local nodes in global networks: the geography of knowledge flows in biotechnology innovation. *Industry and Innovation*, 12(4), 487–507.
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 1360–1380.
- Grau, T., Huo, M., & Neuhoff, K. (2012). Survey of photovoltaic industry and policy in Germany and China. *Energy Policy*, 51, 20–37.
- Hagedoorn, J., & Cloudt, M. (2003). Measuring innovative performance: Is there an advantage in using multiple indicators? *Research Policy*, 32(8), 1365–1379.
- Hertog, P. den. (2000). Knowledge-intensive business services as co-producers of innovation. *International Journal of Innovation Management*, 4(04), 491–528.
- Hervas-Oliver, J.-L., & Albors-Garrigos, J. (2013). *Are Technological Gatekeepers Constraining my Cluster ? Unfolding the paradox of gatekeepers resilience across cluster life cycle stages* (No. 13-04).
- Huallacháin, B. Ó., & Lee, D.-S. (2014). Urban centers and networks of co-invention in American biotechnology. *The Annals of Regional Science*, 52(3), 799–823.
- Iammarino, S., & McCann, P. (2006). The structure and evolution of industrial clusters: Transactions, technology and knowledge spillovers. *Research Policy*, 35(7), 1018–1036.
- Jaffe, A. B. (1989). Real effects of academic research. *The American Economic Review*, 957–970.
- Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics*, 108(3), 577–598.
- Jensen, M. B., Johnson, B., Lorenz, E., & Lundvall, B. Å. (2007). Forms of knowledge and modes of innovation. *Research Policy*, 36(5), 680–693.
- Jones, C. I. (1995). R&D-Based Models of Economic Growth. *Journal of Political Economy*, 103(4).
- Karlsson, C. (2010). *Handbook of research on cluster theory* (Vol. 1). Edward Elgar Publishing.

- Kim, K., & Kim, Y. (2015). Role of policy in innovation and international trade of renewable energy technology: Empirical study of solar PV and wind power technology. *Renewable and Sustainable Energy Reviews*, 44, 717–727.
- Kleinknecht, A., Montfort, K. V. A. N., & Brouwer, E. (2002). The non-trivial choice between innovation indicators, *Economics of Innovation and Technology*, 11(2), 109–121.
- Klitkou, A., & Coenen, L. (2013). The emergence of the Norwegian solar photovoltaic industry in a regional perspective. *European Planning Studies*, 21(11), 1796–1819.
- Klitkou, A., & Godoe, H. (2013). The Norwegian PV manufacturing industry in a Triple Helix perspective. *Energy Policy*, 61(1), 1586–1594.
- Kwon, K.-S., Park, H. W., So, M., & Leydesdorff, L. (2012). Has globalization strengthened South Korea's national research system? National and international dynamics of the Triple Helix of scientific co-authorship relationships in South Korea. *Scientometrics*, 90(1), 163–176.
- Lanjouw, J. O., & Schankerman, M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators. *The Economic Journal*, 114(495), 441–465.
- Leydesdorff, L., & Alkemade, F. (2014). Geographic and Technological Perspectives on “Photovoltaic Cells:” Patents as Instruments for Exploring Innovation Dynamics. *arXiv Preprint arXiv: ...*, 1–44. Retrieved from <http://arxiv.org/abs/1401.2778>
- Leydesdorff, L., & Persson, O. (2010). Mapping the geography of science: Distribution patterns and networks of relations among cities and institutes. *Journal of the American Society for Information Science and Technology*, 61(8), 1622–1634.
- Leydesdorff, L., & Sun, Y. (2009). National and international dimensions of the Triple Helix in Japan: University–industry–government versus international coauthorship relations. *Journal of the American Society for Information Science and Technology*, 60(4), 778–788.
- Leydesdorff, L., Wagner, C. S., & Adams, J. (2013). International Collaboration in Science : The Global Map and the Network. *El Profesional de La Informacion*, 22(Feb), 87–96.
- Lo, C. C., Wang, C. H., & Huang, C. C. (2013). The national innovation system in the Taiwanese photovoltaic industry: A multiple stakeholder perspective. *Technological Forecasting and Social Change*, 80(5), 893–906.
- Locke, R. M., & Wellhausen, R. L. (2014). *Production in the innovation economy*. MIT Press.
- Looy, B. Van, Landoni, P., Callaert, J., Pottelsberghe, B. Van, Sapsalis, E., & Debackere, K. (2011). Entrepreneurial effectiveness of European universities : An empirical assessment of antecedents and trade-offs. *Research Policy*, 40(4), 553–564.
- Lundvall, B.-A. (1992). National systems of innovation: An analytical framework. *London: Pinter*.
- Luo, S., Lovely, M. E., & Popp, D. (2013). Intellectual Returnees As Drivers of Indigenous Innovation: Evidence From the Chinese Photovoltaic Industry, NBER Working Papers 13-05, US Census-Bureau.
- Maggioni, M. A., Nosvelli, M., & Uberti, T. E. (2007). Space versus networks in the geography of innovation: A European analysis. *Papers in Regional Science*, 86(3), 471–493.
- Maggioni, M. A., & Uberti, T. E. (2009). Knowledge networks across Europe: which distance matters? *The Annals of Regional Science*, 43(3), 691–720.
- Malerba, F., & Orsenigo, L. (1997). Technological regimes and sectoral patterns of innovative activities. *Industrial and Corporate Change*, 6(1), 83–118.
- Masciarelli, F., Laursen, K., & Prencipe, A. (2010). Trapped by over-embeddedness: the effects of regional social capital on internationalization. DRUID Working Papers.
- Morgan, K. J. (2004). The Exaggerated Death of Geography: Learning, Proximity and Territorial Innovation Systems, *Journal of Economic Geography* 4(1), 3–21.
- Mudambi, R., & Navarra, P. (2004). Is knowledge power ? Knowledge flows , subsidiary power and rent-seeking within MNCs, *Journal of International Business Studies*, 35 (5), 385–406.
- Nelson, R. R. (1993). *National Innovation Systems: a Comparative Study*. New York: Oxford University Press.
- Ooms, J., James, D., DeRoy, S., Wickham, H., & Horner, J. (2015). RMySQL: Database Interface and MySQL Driver for R. Retrieved from <http://cran.r-project.org/package=RMySQL>
- Ortega, J. L., & Aguillo, I. F. (2009). Mapping world-class universities on the web. *Information Processing and Management*, 45(2), 272–279.
- Parzen, E. (1962). On Estimation of a Probability Density Function and Mode. *Annals of Mathematical Statistics*, 33(3), 1065–1076.
- Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D’Este, P., ... Hughes, A. (2013). Academic engagement and commercialisation: A review of the literature on university–industry relations. *Research Policy*, 42(2), 423–442.
- Polanyi, M. (1958) *Personal knowledge: towards a Post-Critical Philosophy*. University of Chicago Press.
- Ponds, R., Van Oort, F., & Frenken, K. (2007). The geographical and institutional proximity of research collaboration.

- Papers in Regional Science*, 86(3), 423–443.
- Ponds, R., van Oort, F., & Frenken, K. (2010). Innovation, spillovers and university-industry collaboration: An extended knowledge production function approach. *Journal of Economic Geography*, 10(2), 231–255.
- Porter, M. E. (1998). Clusters and the new economics of competition, *Harvard Business Review* (nov/dec 1998), 77-90.
- R Core Team. (2015). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from <http://www.r-project.org/>
- Ranga, M., & Etzkowitz, H. (2012). Triple Helix Systems : An Analytical Framework for Innovation Policy and Practice in the Knowledge Society. *Industry & Higher Education*, 27 (3), 237-262.
- Raspe, O., & van Oort, F.G. (2011). Firm heterogeneity , productivity and spatially bounded knowledge externalities. *Socio-Economic Planning Sciences*, 45(1), 38–47.
- Raspe, O., & van Oort, F. (2009). Growth of new firms and spatially bounded knowledge externalities. *The Annals of Regional Science*, 46(3), 495–518.
- Richardson, H.W. (1995). Economies and Diseconomies of Agglomeration. In: H. Giersch (Ed), *Urban Agglomeration and Economic Growth*, berlin/Heidelberg, Springer, pp. 123-155.
- Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5), S71-S102.
- Rosenblatt, M. (1956). Remarks on Some Nonparametric Estimates of a Density Function. *The Annals of Mathematical Statistics*, 27(3), 832–837.
- Sassen, S. (2002). *Global networks, linked cities*. Psychology Press.
- Schlierf, K., & Meyer, M. (2013). Situating knowledge intermediation: Insights from science shops and knowledge brokers. *Science and Public Policy*, 40 (4), 430-441.
- Simmie, J. (2003). Innovation and Urban Regions as National and International Nodes for the Transfer and Sharing of Knowledge. *Regional Studies*, 37(October), 607–620.
- Tabuchi, T. (1998). Urban Agglomeration and Dispersion: A Synthesis of Alonso and Krugman. *Journal of Urban Economics*, 44(3), 333–351.
- Tödting, F., & Trippel, M. (2005). One size fits all? Towards a differentiated regional innovation policy research. *Research Policy*, 34(8) 1203-1219.
- Torre, A. and Wallet (Eds) (2014). *Regional Development and Proximity Relations*, Edward Elgar.
- Van Geenhuizen, M., & Nijkamp, P. (2012). Knowledge virtualization and local connectedness among young globalized high-tech companies. *Technological Forecasting and Social Change*, 79(7), 1179–1191.
- Vidican, G., McElvaney, L., Samulewicz, D., & Al-Saleh, Y. (2012). An empirical examination of the development of a solar innovation system in the United Arab Emirates. *Energy for Sustainable Development*, 16(2), 179–188.
- Wasserman, S., & Faust, C. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge University Press.
- West, J. (2014). Too little, too early: California’s transient advantage in the photovoltaic solar industry. *The Journal of Technology Transfer*, 39(3), 487–501.
- Wickham, H. (2015). rvest: Easily Harvest (Scrape) Web Pages. Retrieved from <http://cran.r-project.org/package=rvest>
- Williamson, O. E. (1981). The Economics of Organization : The Transaction Cost Approach. *The American Journal of Sociology*, 87(3), 548–577.
- Wu, C.-Y. (2014). Comparisons of technological innovation capabilities in the solar photovoltaic industries of Taiwan, China, and Korea. *Scientometrics*, 98(1), 429–446.
- Ye, F. Y., Yu, S. S., & Leydesdorff, L. (2013). The Triple Helix of University-Industry-Government Relations at the Country Level, and its Dynamic Evolution under the Pressures of Globalization. *Journal of the American Society for Information Science and Technology*, 1–21.
- Zheng, C., & Kammen, D. M. (2014). An innovation-focused roadmap for a sustainable global photovoltaic industry. *Energy Policy*, 67, 159–169.