Science and Public Policy Advance Access published November 20, 2014

Science and Public Policy (2014) pp. 1–19 doi:10.1093/scipol/scu070

Characterizing and comparing innovation systems by different 'modes' of knowledge production: A proximity approach

Sjoerd Hardeman[1,*], Koen Frenken[2,3], Önder Nomaler[4] and Anne L. J. Ter Wal[5]

[1]*Unit of Econometrics and Applied Statistics, Joint Research Centre, European Commission, Via Enrico Fermi 2749, TP 361, 21027 Ispra (VA), Italy*

[2]*Innovation Studies, Copernicus Institute of Sustainable Development, Utrecht University, PO Box 80115, 3508TC, Utrecht, The Netherlands*

[3]*Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE), Lund University, Sölvegatan 16, S-22100, Lund, Sweden*

[4]*School of Innovation Sciences, Eindhoven University of Technology, PO Box 513, 5600 MB, Eindhoven, The Netherlands*

[5]*Imperial College Business School, Department of Innovation & Entrepreneurship, Tanaka Building, South Kensington Campus, London SW7 2AZ, United Kingdom*

**Corresponding author. Email: sjoerd.hardeman@jrc.ec.europa.eu.*

Though the concept of innovation systems has become influential in both academia and policy-making, an analytical approach to understanding innovation systems is still lacking. In particular, there is no analytical framework to measure 'Mode 1' and 'Mode 2' knowledge production. We propose a framework based on the proximity concept. Mode 1 and Mode 2 knowledge production are characterized by collaborations with cognitive, organizational, social, institutional and geographical proximity, and distance, respectively. Using a gravity model approach we apply our framework to the case of type 2 diabetes research and provide a characterization of the global innovation system and a comparative analysis of the North American and European innovation systems. Our main results hold that although collaborative research on type 2 diabetes generally follows a logic of proximity and hence is not characterized as Mode 2, important differences and similarities exist between the North American and European innovation systems.

Keywords: mode 2 knowledge production; collaboration; co-publication; diabetes; triple helix; network.

1. Introduction

Few doubt that the organization of scientific knowledge production has changed substantially over the past decades. Universities now interact more closely with industry and other societal stakeholders to legitimate public funding and raise more private funding (Gibbons et al. 1994). Although, historically, these hybrid collaborations were already quite common in some fields of science, it has been argued that only recently have these forms of collaboration become ubiquitous. Indeed, this trend has been evidenced by publication data showing that universities increasingly co-publish with other institutional actors including firms, governments and hospitals (Hicks and Katz 1996; Adams et al. 2005).

From a geographical perspective, one can expect that the nature and extent of such cross-institutional interactions differ across territories to the extent that territories control the institutions that structure such hybrid collaborations. Thus, a national innovation system (Freeman 1987; Lundvall 1988), can be defined as a national

system whose organizations interact across institutional spheres for the purpose of knowledge production and innovation. *Mutatis mutandis*, one can also speak of a regional innovation system (Cooke et al. 1998) or an international innovation system (Carlsson 2006). More generally, one can speak of a territorial innovation system with its boundaries defined by geographical areas that have some degree of institutional specificity (Morgan 2004).

A comprehensive assessment of the distributed nature of knowledge production has been proposed by Gibbons et al. (1994). They considered the increased interaction across institutional spheres as just one aspect out of many more aspects that characterize today's knowledge production processes. They introduced the distinction between the traditional university mode of knowledge production termed 'Mode 1', and the emerging distributed mode of knowledge production termed 'Mode 2'. Gibbons et al. (1994: 34) summarised their central thesis concerning Mode 2 knowledge production as follows:

> ...not only is the average number of authors per paper increasing, but much more significantly, so are the diversity of specialisms and disciplines involved in the writing of a single paper and the range of institutions and organizations from which the authors originate. In addition, the geographical distribution of these institutions continues to broaden. In mode 2, not only are more actors involved in the genesis of knowledge, but they remain socially distributed.

Thus, their thesis not only emphasizes university–industry–government collaboration similar to the territorial innovation system concepts, but it also highlights a trend towards globalization (Castells 1996) and interdisciplinarity (Barry et al. 2008) in collaborative knowledge production.

Though the concepts of innovation systems and Mode 2 knowledge production have become very influential in academic research and policy-making circles alike (Lundvall 2007; Hessels and Van Lente 2008), their use has been limited to qualitative research. We argue that the lack of quantitative empirical research emerges from a lack of operational concepts that capture the various characteristics of Mode 2, as a result of which there is a lack of methodological standardization and empirical understanding of the innovation system and Mode 2 concepts. We believe that, ultimately, the operational difficulties are rooted in the a-theoretical notions underlying the two concepts. What is more, the proliferation of alternative concepts that aim to capture the changing nature of knowledge production—the network society (Castells 1996), the triple helix of university–industry–government relations (Leydesdorff and Etzkowitz 1996), academic capitalism (Slaughter and Leslie 1997), open innovation (Chesbrough 2003), global pipelines (Bathelt et al. 2004), logics of interdiscplinarity (Barry et al. 2008) and search regimes (Bonaccorsi 2008)—further complicate the quest for a common understanding and more cumulative research programmes.

What is lacking then is an analytical approach that allows for a systematic understanding of the nature and extent to which organizations interact in innovation systems, and how such systems can be compared across territories. The aim of this paper is therefore to propose an analytical framework suitable for a systematic assessment of some of the main propositions made within the literature on innovation systems and Mode 2 knowledge production that allows for both a general characterization and a comparison of territorial innovation systems.

First, as innovation system and Mode 2 knowledge production are essentially concepts that refer to 'interactive learning' (Lundvall 1988) and 'distributed' knowledge production (Gibbons et al. 1994), we propose to base such a framework on the proximity concept (Rallet 1993; Rallet and Torre 1999; Knoben and Oerlemans 2006). The proximity dimensions we adopt are taken from Boschma (2005) and Balland (2012) who distinguish between cognitive, organizational, social, institutional and geographical proximity. As will be shown, these dimensions map almost one-to-one to various aspects which are emphasized under the Mode 2 knowledge production concept. That is, actors engaged in collaborative knowledge production can be positioned along the several Mode 2 dimensions as being more or less proximate, where proximity corresponds to Mode 1 and its opposite (distance) to Mode 2 knowledge production. In doing so, we develop an analytical framework for the study of territorial innovation systems based on the various dimensions of collaboration.

Second, we apply our framework to the case of knowledge production in the field of type 2 diabetes, analysing the worldwide patterns of collaboration in this field as well as providing a comparative analysis of the North American (USA and Canada) and European (EU15 (the 15 member states of the EU prior to 1 May 2004) and Switzerland) innovation systems in the field of type 2 diabetes. Our system delineation is akin to the notion of a technological system (Carlsson et al. 2002; Hekkert et al. 2007) in that we delineate our innovation system in terms of the actors who work on solutions to a common problem (here, type 2 diabetes). We use co-publication data to indicate the collaborations between actors in the innovation system. As such, the problem at stake concerns type 2 diabetes as addressed throughout the scientific literature. The actors involved are the organizations that concern themselves with providing evidence on solutions to this problem, and their interactions are reflected by collaborations among organizations as measured by co-publications.[1]

Overall, the objective of this study is to propose the proximity concept within a network analytic approach as a useful analytical tool to characterize and compare innovation systems in terms of different modes of knowledge production. As such, the main contribution of this study

is two-fold. One contribution is that we show how the proximity framework can be used empirically to test general propositions that follow from a characterization of territorial innovation systems in terms of Mode 1 and Mode 2 knowledge production. The other main contribution is that we show how the proximity framework can be empirically used to compare innovation systems in terms of their 'Mode 2-ness'.

Accordingly, our main empirical results are also two-fold. Our first empirical result holds that the global innovation system in the field of type 2 diabetes research is generally characterized by Mode 1 instead of Mode 2 knowledge production as evidenced by the importance of proximity rather than distance in all five dimensions. Second, differences and similarities exist between the North American and European innovation systems in the field of type 2 diabetes research. In particular, we find that social and organizational proximity play a relatively smaller role in Europe as compared to North America while cognitive and institutional proximity are equally important in Europe and North America. Hence, we conclude that Europe is more Mode 2 than North America in terms of the geographical, social and organizational aspects to collaborative research and less Mode 2 in terms of being relatively confined to take place within national boundaries, although Europe and North America are fairly equal in terms of their Mode 2-ness when it comes to the institutional and cognitive aspects of collaborative knowledge production.

2. Theoretical framework

If anything characterized the change in scientific knowledge production over the past century, it has been its increasingly distributed nature. This trend was noted by Price (1963) who noticed an increase in the number of authors on scientific papers. He described this trend at the time as a transformation from 'little science to big science'. Since then, the number of authors per paper has steadily increased with the mean number of authors per paper currently exceeding 3.5 in science and engineering and 2.0 in social sciences (Wuchty et al. 2007).[2] The trend in increasing levels of collaboration is accompanied by an increase in the share of university–industry–government relations (Hicks and Katz 1996; Adams et al. 2005) and increasing internationalization (Adams et al. 2005; Frenken et al. 2009).

To explain the pattern of interaction, a proximity approach is useful as it emerged from research in innovation networks (Rallet 1993; Rallet and Torre 1999). The proximity concept can be applied to interaction in science using the same five dimensions as distinguished in the study of innovation networks: cognitive, organizational, social, institutional and geographical proximity (Boschma 2005; Frenken et al. 2009).

2.1 Cognitive proximity. The effective transfer of knowledge in research collaboration requires absorptive capacity to identify, interpret and exploit the new knowledge (Cohen and Levinthal 1990; Nooteboom 1999). In effect, the ease of knowledge transfer and mutual learning between actors may depend on the similarity of their knowledge bases (Lane and Lubatkin 1998). The capacity of actors to exchange and combine their knowledge requires cognitive proximity. That is, the knowledge bases of actors should be similar enough in order to communicate, understand and process scientific knowledge successfully. The importance of cognitive proximity is evident from the disciplinary nature of most scientific research, publishing and teaching.

2.2 Organizational proximity. Organizational proximity has been defined as the extent to which networks occur within the context of an organizational arrangement (Boschma 2005). Typically, organizational proximity thus refers to the extent to which any two actors are under shared hierarchical control. Historically, universities have emerged as the prime organizational vehicle to organize interaction between scientists, where hierarchy is typically delegated to a *primus inter pares*. Yet, other organizational forms have emerged in science including public research agencies and industrial laboratories. Organizational proximity is argued to facilitate the establishing of collaboration networks, because it reduces uncertainty and opportunism in collaboration projects through collegiality and shared goal orientations.

2.3 Social proximity. The notion of social proximity has its roots in the embeddedness literature (Granovetter 1985). This literature indicates that interactions are always embedded in a social context and that, in turn, social relations affect the outcomes of interactions. In the context of science, social proximity may refer to the extent that two actors have established a friendly relation in the past (in previous projects, as colleagues, as friends, or otherwise). As for organizational proximity, social proximity reduces the uncertainty and opportunism in collaboration, as opportunistic behaviour will lead to reputational loss within an actor's social network (Dasgupta and David 1994). In addition, social proximity between actors may stimulate commitment and mutual trust, both of which may trigger the initiation and continuation of collaborative engagements.

2.4 Institutional proximity. Whereas social proximity is defined in terms of socially embedded relations between actors at the micro-level, institutional proximity is associated with institutions at the macro-level. As such, actors are institutionally proximate once they operate under the same set of norms and values. Both formal and informal

institutions structure behaviour by providing particular incentives. In science, universities, industries, governments and hospitals all operate under different institutional regimes, thus giving rise to incentive incompatibility problems (Dasgupta and David 1994). For example, firms have an incentive to appropriate knowledge, while universities have an incentive to publish research instantaneously. Collaboration thus benefits from institutional proximity as fewer conflicts are expected to arise when collaborators have similar incentives. This explains why cross-institutional collaborations (viz. 'triple helix interactions') are difficult to organize.

2.5 Geographical proximity. The final dimension to be distinguished is geographical proximity. There is a strong claim that geographical proximity is still an important driver of network formation despite the tendency for innovation systems to internationalize (Castells 1996; Carlsson 2006). Indeed, the majority of scientific collaborations take place between actors who are geographically proximate, and generally, within the same region or country (Katz 1994; Hoekman et al. 2009, 2010). Geographical proximity is beneficial for research as effective learning requires face-to-face interaction to transfer tacit knowledge (Collins 1985). Such interaction is easier (and cheaper) to organize when agents are co-located in space. Once the other four forms of proximity have been defined, geographical proximity can be defined in a restricted manner as the inverse of physical distance between actors in absolute (e.g. kilometres) or relative terms (e.g. travel time) (Boschma 2005).[3] For analytical purposes, it is essential to define geographical proximity in such a restricted manner, in order to isolate it from the other dimensions of proximity.

As proximity is an analytical concept, it offers some specific advantages in empirical work explaining the structure of networks. First, by incorporating multiple proximity dimensions in a single explanatory framework, one can test which forms of proximity are determining the patterns in collaboration networks. When including only one proximity dimension in the analysis, findings typically show the strong explanatory power of that dimension. However, due to correlation between proximities, one can only assess the effect of a dimension if other dimensions are controlled for (Boschma 2005; Frenken et al. 2009). Second, one can extend the list of relevant proximity dimensions from Boschma's (2005) five dimensions to any number of dimensions without changing the meaning of each dimension. For example, linguistic and cultural proximity dimensions may be introduced. Thus, the proximity dimensions are analytically orthogonal even though empirically, many dimensions of proximity may turn out to be correlated.

Following the proximity concept, the distinction between Mode 1 and Mode 2 knowledge production can now be made analytically. Mode 1 stands for knowledge production in which actors are distributed, yet proximate, while Mode 2 knowledge production stands for distributed knowledge production processes, in which actors are distant. The proposed definition of Mode 1 coincides with the ivory tower image of scientific knowledge production, a mode which is disciplinary (cognitive proximity), within university departments (organizational proximity), in personal networks (social proximity), under a strict set of academic norms (institutional proximity) and co-located within the walls of the laboratory site (geographical proximity). Mode 2, by contrast, is characterized by Gibbons et al. (1994) as transdisciplinary (cognitive distance), cross-organizational (organizational distance), in temporary and open networks (social distance), with various, possibly conflicting, goals (institutional distance), and crossing national borders and physical space (geographical distance). A close reading of Gibbons et al. (1994) provides further support for this interpretation of the Mode 2 concept. For each of the proximity dimensions, quotes from Gibbons et al. (1994) can be found that express the nature of Mode 2 knowledge production as collaborative research between distant actors (see Table 1).

From a proximity perspective, one can further qualify the Mode 2 concept as proposed by Gibbons and colleagues. If one were to define Mode 2 knowledge production in a strict sense, as collaborations in which the actors are distant in all dimensions, one can expect to observe very few instances of such modes of collaboration. More often, one would expect to observe that actors are proximate in one dimension as a means of managing the difficulties and conflicts that arise from being distant in the other four dimensions (Ponds et al. 2007). Accordingly, one could develop a more refined typology of Mode 2 knowledge production. For example, Mode 2 knowledge production within geographic clusters would make use of geographical proximity as an organizing principle, while Mode 2 knowledge production within a dedicated organization would make use of organizational proximity as an organizing principle.

The proximity framework also aptly highlights the differences between the Mode 2 concept on the one hand and the more specific terminology proposed by other scholars. The innovation system concept stressed inter-organizational learning, cross-institutional interaction and proximity within a particular territory. As such, the innovation system concept (Freeman 1987; Lundvall 1988) emphasizes the importance of bridging organizational and institutional distances while benefitting from geographical proximity. Thus, while rich in scope, the innovation system concept does not explicitly include the cognitive and social dimensions of collaborative knowledge production.

Other concepts have been more focused on a single dimension. For example, the network society concept (Castells 1996) emphasizes geographical distance, the

Table 1. Translation of 'Mode 2' knowledge production to collaborative knowledge production along five proximity dimensions

Mode 2 knowledge production	Expressed by Gibbons et al. (1994) as:	Related to
Transdisciplinarity	'*...a novel environment in which knowledge flows more easily across disciplinary boundaries...*' (p. 20) in which '*integration is not provided by disciplinary structures...but is envisaged and provided from the outset in the context of usage, or application in the broad sense...*' (p. 27) and '*...disciplines are no longer the only locus of the most interesting problems, nor are they the homes to which scientists must return for recognition or rewards*' (p. 30).	Cognitive distance
Societal contextualization	'*...the organization of research more open and flexible*' (p. 20) '*...with knowledge becoming socially distributed to ever wider segments of society*' (p. 34). Here, '*the previous one-way communication process from scientific experts to the lay public perceived to be scientifically illiterate and in need of education by experts has been supplanted by politically backed demands for accountability of science and technology and new public discussions in which experts have to communicate a more 'vernacular' science than ever before*' (p. 36).	Organizational distance
Social distributedness	'*...preference given to collaborative rather than individual performance and excellence judged by the ability of individuals to make a sustained contribution in open, flexible types of organization in which they may only work temporarily*' (p. 30).	Social distance
Institutional hybridization	'*...a closer integration of the process of discovery with that of fabrication*' (p. 19) in which '*...institutional differences between, say, universities and industry, seem to be less and less relevant*' (p. 30). '*Thus while different kinds of institutions are able to maintain their own distinctive character and functions, they continually generate new forms of communication. This partially explains the emergence of hybrid new communities, consisting of people who have been socialized in different subsystems (...), but who subsequently learn different (...) modes of behaviour, knowledge and social competence that originally they did not possess*' (p. 37).	Institutional distance
Diversity of sites	'*...the diffusion over a wide range of potential sites of knowledge production...*' (p. 17).	Geographical distance

triple helix concept (Leydesdorff and Etzkowitz 1996; Etzkowitz and Leydesdorff 2000) and the academic capitalism concept (Slaughter and Leslie 1997) focus on institutional distance, the open innovation concept (Chesbrough 2003) is essentially addressing organizational distance, and the logics of interdisciplinarity concept (Barry et al. 2008) obviously deals with the role of cognitive distance. More encompassing are the notions of global pipelines (Bathelt et al. 2004) stressing geographical, organizational and cognitive distance and the notion of search regimes (Bonaccorsi 2008; Bonaccorsi 2010) focusing on the role of cognitive and institutional distances in different disciplines and changes therein over time.

The main advantage of conceptualizing an innovation system in terms of collaborative knowledge production among proximate (Mode 1) or distant actors (Mode 2) holds that each single actor can be characterized as a coordinate in five-dimensional space using only the information on the actors involved. What is required is to operationalize the cognitive, organizational, institutional, social and geographical attributes of each actor such that their mutual distance can be established in five-dimensional space. This renders the empirical operationalization of the Mode 2 concept straightforward (obviously at the expense of the richness of the qualitative descriptions put forward by Gibbons and colleagues). First, if proximity is an important driver of collaboration between organizations of an innovation system on all five dimensions, then this innovation system is characterized as Mode 1 rather than Mode 2. In principle, from distinguishing different aspects of Mode 2 knowledge production along different proximities, any innovation system can be characterized in terms of its 'Mode 2-ness' with reference to the particular proximity dimensions that do not play a role in collaborative research. Second, turning to a comparison of territorial innovation systems, the less proximity in a focal innovation system plays a role in steering collaborative knowledge production between organizations as compared to another innovation system, the more that focal innovation system is characterized by following a Mode 2 logic. This latter analysis is especially relevant in the light of science policy seeking to steer particular forms of collaborative research (Dosi et al. 2006; Bonaccorsi 2007; Chessa et al. 2013).

3. Data and methods

3.1 Case selection

The choice of diabetes as a case to illustrate our framework empirically, resides first and foremost in the reality of the problem. Diabetes affects millions of people around the globe and is expected to affect even more people in the near future (Danaei et al. 2011; Hurley 2011). Within the scientific literature then, contributions to solving this

problem have grown tremendously. The fact that diabetes is a chronic disease further underlines its importance. Diabetes not only affects many people, as a chronic disease it also affects many people for longer periods of time and possibly with major consequences.

Diabetes, and especially its type 2 variant,[4] constitutes a very complex disease involving many interacting factors such as: genetics, lifestyle and the (industrialized) environment (Zimmet et al. 2001). However, not only are the aspects involved in the constitution of this disease varied, as a consequence so are the people and organizations occupying themselves with finding solutions to this problem. What medical professionals call translational medicine (Woolf 2008) seems to be especially accurate for diabetes, that is, as a description of medical science that concerns itself with diabetes duly takes into account the whole process from the laboratory bench to the patient's bedside involving different actors (National Institutes of Health 2004). As such, the nature of diabetes as a scientific problem is immediately enmeshed with societal undertones. Thus, the provision of solutions can be expected to be organized in a Mode 2 fashion (Gibbons et al. 1994; Nowotny et al. 2003).

Overall, we thus consider the field of type 2 diabetes research to constitute an exemplary case wherefore territorial innovation systems can be expected to be organized in a Mode 2 fashion: that is, along a logic of distance instead of proximity. Alternatively, if territorial innovation systems in the field of type 2 diabetes are not found to be organized along a logic of distance, this casts considerable doubt on research more generally (i.e. beyond the field of type 2 diabetes research itself) being organized in a Mode 2 fashion. That is to say, if even a field like type 2 diabetes research is not organized in a Mode 2 fashion, then we would also have low expectations about other fields being organized in a Mode 2 fashion.

3.2 Dependent variable

The dependent variable in the analysis is the intensity of collaborative science between each pair of organizations. Co-publications are scientific papers that are produced by multiple organizations and are often used as indicators of collaborations in science (Katz and Martin 1997; Frenken et al. 2009). Generally, co-publications concern co-authored papers by scholars each working for different organizations. In fewer cases, co-publications concern single-authored papers by scholars with multiple affiliations. In both cases, multiple organizations can be said to have been involved in the production of scientific knowledge.

It must further be noted that co-publications representing collaborative science are only a proxy of research collaboration, since not all research collaborations may end up in a scientific publication, and, vice versa, not all organizations mentioned on a paper may have had an active role in the production of that particular knowledge. Yet, as long as large sets of data are used, these exceptions are no longer expected to influence the conclusions that can be drawn from the data analysis. For this reason, co-publications have been an accepted indicator of collaborative science (Lundberg et al. 2006).

We used Elsevier's Scopus database to construct our dependent variable. We proceeded in a number of steps (Hardeman 2013). First, in order to identify and extract all bibliometric records representing documents that are concerned with research on type 2 diabetes we constructed a search query based on a list of tags that capture the different names used to address this health problem (see Appendix, A4). Extracting bibliometric information pertaining to a particular research field or discipline is in itself far from straightforward. The list that we used is adapted from discussions that we had with experts from this research field and is complemented by terms denoting type 2 diabetes as they are provided in the medical classification systems of the International Classification of Diseases (World Health Organization 2011), the Medical Subject Headings (MeSH) (US National Library of Medicine 2011), and EMTREE (Elsevier Pharma Development Group 2009). Using the search query defined, we extracted 72,725 uniquely coded bibliometric records that represent scientific publications concerned with type 2 diabetes for the period 1996–2008.

Second, every record lists one or more organizations as author affiliations. For each publication record, the information elements can include:

- the name of an organization
- an organization ID
- a sub-organization ID
- the country in which the organization is located
- the city and/or region in which the organization is located
- a more fine-grained description of the location of an organization (e.g. a street, zip code or post box)

If we split the publications record-wise for every organization represented there we can identify 186,719 unique publication-organization pairs.

To identify unique organizations we first made use of the ID Scopus assigns to an organization listed as an affiliation of an author of a document. To assess Scopus' consistency in assigning unique IDs to unique organizations, we randomly checked 105 such IDs across 18,390 records (9.8% of all information elements). This check involved making sure that the different names attributed to each unique ID are indeed representing the same organization. We performed this check manually and concluded that as only 1.8% of all records represent a deviating name. In general, Scopus' affiliation IDs are consistent across records (at least in our case). This lends support to our

approach of taking Scopus' affiliation IDs as our starting point to identify organizations.

However, although the affiliation IDs assigned by Scopus are internally consistent, this does not imply that different affiliation IDs cannot refer to the same (overarching) organization. Thus, in order to ensure that different IDs indeed reflect different organizations we thus had to unify our organizational level data (Van Raan 2005). We used a set of three rules to unify our organization-level data. If any two affiliation IDs belong to:

- the same institutional sphere
- the same hierarchical meta-structure
- the same geographical area

Then these organization IDs are taken to reflect the same organization. Institutional spheres were assigned on the basis of the mission statements taken from the websites of the organizations representing the affiliation IDs. Similar to Parsons' (1956a; 1956b) idea of bracketing up society into sub-spheres, we distinguish among four such institutional spheres: industry, care, academia and government. Likewise, from their websites we assign all affiliation IDs to their overarching hierarchical structure. Given that a single (overarching) organization can be located at different physical sites we define the organization at the branch level. That is, every set of unique institutional-hierarchical entities were clustered according to their geographical location (see Leydesdorff and Persson (2010) for a discussion on using bibliometric data to map the geography of science). Here we took 50 km of separation to delineate one branch from another.[5] Following this three-fold procedure we unified all affiliation IDs that occur more than nine times in our data set and eventually obtained 1,218 distinct organizations that can be characterized as a coordinate in five-dimensional space.[6]

Finally, for each pair of unique organizations, we counted the number of times they were co-occurring on a paper during the period 2003–8. Given that we have a total of 1,218 organization branches, we have 741,153 observations of organization pairs. Since collaborations are undirected interactions (i.e. the number of collaborations between organization i and organization j is the same as the number of collaborations between organization j and organization i), we only use half of the total collaboration matrix. What is more, we only take inter-organizational collaborations into account and thus dispense with the diagonal of the full matrix (i.e. those instances in which organization i is the same as organization j). Hence, given:

$$N = \frac{(n^2 - n)}{2}$$

where N is the number of organization pairs and n the number of organizations, we end up with 741,153 observations at the global level. It must be noted that most observations are zero implying that most organization pairs that could, in principle collaborate, did not actually collaborate at all during the period under investigation.

3.3 Independent variables

In order to explain collaboration intensity between each organization pair, we propose five independent variables covering the five proximity dimensions described (see Table 2 for a formal description of the independent variables). First, geographical proximity is operationalized as the inverse of the distance in kilometres separating two organizations. Apart from measuring geographical proximity in terms of the inverse of kilometric distance, we also include a dummy variable measuring whether or not any two organizations are from the same country. While the former operationalization of geographical proximity comes closest to the idea put forward by Boschma (2005), the latter operationalization of geographical proximity captures the role of national boundaries deemed important within the concept of national systems of innovation.

Second, social proximity is operationalized as the number of prior ties between any two organizations, measured as the log of the number of co-publications in the period 1996–2002.[7] We acknowledge that one would ideally have more fine-grained data on social ties between organizations, for example about labour mobility flows, friendship relations or ties among former colleagues (Breschi and Lissoni 2009). In the absence of such data, we take social proximity as approximated by past collaboration activities, as this operationalization comes very close to the idea of flexible networks deemed important within Mode 2 knowledge production (Gibbons et al. 1994). That is, if social proximity positively relates to collaboration, one would rather speak of inflexible networks and hence Mode 1 knowledge production. Reversely, if social proximity has a negative effect on collaboration, one can speak of flexible networks as in Mode 2 knowledge production. It should be noted that social proximity, when used as a determinant of collaboration, does not necessarily reflect social relations as they occur at the time of collaborations, but much more the persistence of past social relations.

Third, starting from the premise that organizations are more cognitively proximate when they often publish in the same academic journals, we measure cognitive proximity as the cosine of the overlap in journals in which any two organizations published in the period 1996–2002. Let X_i be a vector of length $N = 1995$, where N is the number of all journals in which type 2 diabetes research is published. The Kth element of X_i indicates the number of papers published by institute i in journal K. For a dyad comprising institutes i and j, cognitive proximity is simply the cosine of the angle

Table 2. Description of variables

Variable	Description
No. of co-publications$_{ij}$	Number of papers in period 2003–8 on which organization i and organization j both appear
Intra-country dummy$_{ij}$	Dummy equals 1 if both organization i and organization j are from same country
Geographical proximity$_{ij}$	Inverse of distance (in km) (plus 1) between organization i and organization j
Social proximity$_{ij}$	Log of number of papers in period 1996–2002 in which both organization i and organization j appear
Cognitive proximity$_{ij}$	Cosine of overlap in journals in which both organization i and organization j publish in period 1996–2002
Institutional proximity$_{ij}$	Dummy = 1 if both organization i and organization j are from same institutional sphere (i.e. academia, industry, government or care)
Organizational proximity$_{ij}$	Dummy = 1 if both organization i and organization j belong to same hierarchical meta-structure
Mass$_{ij}$	Log of total number of publications of organization i times total number of publications of organization j in period 2003–8
Transitivity$_{ij}$	Number of organizations with which both organization i and organization j co-publish in period 2003–8
EU dummy$_{ij}$	Dummy = 1 if both organizations are from EU
Expected collaboration propensity$_{ij}$	Square root of number of dyads formed by organization i in period 2003–8 times number of dyads formed by organization j in period 2003–8

between vectors X_i and X_j. Using journal source information instead of patent class information, this measure of cognitive proximity closely resembles existing measures of technological (Jaffe 1986) or knowledge relatedness (Breschi et al. 2003).

Fourth, institutional proximity is represented by a dummy variable denoting whether or not any two organizations belong to the same institutional sphere. In delineating organizations we collected data on the mission statements of the organizations involved. Using the websites of organizations we were thus able to assign every organization to a unique institutional sphere. It follows that any pair of organizations can readily be characterized in terms of institutional proximity once the two organizations belong to the same institutional sphere: either academia, industry, government or care (Ponds et al. 2007).

Finally, organizational proximity is measured as a dummy variable indicating whether or not two organizations belong to the same overarching hierarchical meta-structure. In the context of our study organizational proximity can be of two kinds. It involves either a characterization of the relation between a university (assigned to the institutional sphere of academia) and its associated university hospital (assigned to the institutional sphere of care) or a characterization of the relation between two organizations of the same overarching hierarchical meta-structure but located at different physical sites. As such, our operationalization comes very close to a transaction cost interpretation of organizational proximity along hierarchical lines (Williamson 1981).[8]

3.4 Methods: A gravity model approach

To analyse the determinants of co-publication activity between any two organizations, we apply a gravity equation specification of the kind proposed by Ponds et al. (2007) and later adopted by Hoekman et al. (2009) and Maggioni and Uberti (2009) among others. In a gravity model, the gravitational force between two objects is assumed to be positively dependent on the mass of the objects and negatively (positively) on the distance (proximity) between them. In our case this means that the collaboration intensity between two organizations is dependent on their size (as approximated by their total number of publications) and the various proximity measures.

Alternatively, we could have opted for other network analytic approaches to assess collaboration patterns in the field of type 2 diabetes research such as the multiple regression quadratic assignment procedure, exponential random graph models, and stochastic actor-oriented models (Broekel et al. 2014). We choose to adopt a gravity modelling approach for three reasons. One reason is practical: a gravity modelling approach is not restricted by the number of observations while the alternative methods are much more restricted in that sense. Another reason is that our analysis is comparative in nature and thus does not involve a dynamic analysis as in the case of stochastic actor-oriented models. Finally, and most importantly, many organizations in the context of our case of type 2 diabetes research are sizable to the extent that agency at this aggregate level is contested, if not completely absent. In other words, organizations in the field of type 2 diabetes research are much more like countries or regions in terms of the agency involved (or the lack thereof) rather than the 'real' actors often considered in more sociological social network analyses (Snijders 2011).

As in other gravity equation specifications used to model collaborative science, we apply a zero-inflated negative binomial regression model since we deal with count data characterized by over-dispersion (Burger et al. 2009).

In the zero-inflated part of our regression models we only include one variable that captures the expectation on any two organizations collaborating (i.e. 'expected collaboration propensity') (see Table 2). As such we assume that, in principle, any two organizations should be able to collaborate provided both organizations have collaborative capacity in the first place. The negative binomial part of our regression models then explains the extent to which the collaboration intensity between any two organizations can be explained by proximity provided that the organizations involved can, in principle, collaborate.[9]

Given that our observations are at the dyadic level, each organization affects multiple observations. Known as Galton's problem (Tylor 1889), it can be argued that observations are not statistically independent, thus potentially leading to an underestimation of standard errors (Lincoln 1984). In addition, co-publication networks generally show a high degree of clustering (Newman 2001). Consequently, collaboration as measured by co-publication need not necessarily be driven by a strict mutual (or dyadic) proximity rationale, rather by a rationale in which any two organizations are primarily connected by both being connected to a third organization. To correct for these issues we include a structural variable (i.e. 'transitivity') that accounts for the number of collaborators that any two organizations have in common (Lincoln 1984; Stuart 1998) (see Table 2). In addition all the models that were used report on robust standard errors.

4. Results

From the observation that 85% of all publications in type 2 diabetes in the period 1996–2008 concern co-publications, we conclude that the global innovation system in this field of research is indeed best characterized as an interactive (i.e. collaborative) system. Compared to other, more general, research fields, type 2 diabetes research seems to be among the most collaborative ones. Indeed, whereas Wuchty et al. (2007) found that in the year 2000 for the general field of science and engineering—the most collaborative field of research they identified—around 80–85% of all publications involve co-publications, type 2 diabetes research can definitely be considered to be a collaborative enterprise.

In order to compare different territorial innovation systems, we took the example of North America (USA and Canada) and Europe (EU15 and Switzerland). Table 3 shows some general output characteristics of the global, European and North American innovation systems in the field of type 2 diabetes research. As reported before, the global innovation system produced 72,725 publications in the period 1996–2008. Europe (29,868 publications) produces somewhat more publications than North America (24,976 publications). Overall, the USA produces by far the largest number of publications (22,701), followed by Great Britain (7,785 publications), Japan (5,974 publications), and Germany (5,061 publications). Most European countries are part of the top-20 of countries producing the largest number of publications; the exceptions being Austria (730 publications, ranked 24th), Greece (617 publications, ranked 26th), Ireland (291 publications, ranked 32nd) and Portugal (178 publications, ranked 41st). Canada, the other North American country, produced 2,930 publications and is therefore ranked 7th. Overall, the European and North American innovation systems are roughly comparable in terms of the number of publications.

Table 3. Quantities of research produced by global, European, North American, and top-20 national innovation systems in type 2 diabetes research (full, single count)

Territory	Number of publications
Global	72,725
EU15 + Switzerland	29,868
USA + Canada	24,976
USA	22,701
Great Britain	7,785
Japan	5,974
Germany	5,061
Italy	3,766
France	3,494
Canada	2,930
Australia	2,380
Spain	2,230
Sweden	2,162
Netherlands	2,101
China	2,098
Denmark	1,908
India	1,462
Switzerland	1,197
Finland	1,173
Belgium	1,116
Poland	994
Turkey	898
South Korea	882

Unfortunately, we do not have data about the impact of each individual paper published by European and North American researchers in the period 1996–2008, let alone the quality of collaborations between organizations. However, as the journal impact factors of the journals listed in Scopus are publicly available, we had a closer look at the most recent impact factors of journals in which European and North American researchers publish. Table 4 lists some statistics about the quality of journals in which European and North American papers are published. The use of journal impact factors to evaluate research is highly contested (Seglen 1997). Nevertheless, and although journal impact factors say next to nothing about the quality of an individual paper

Table 4. Quality of journals in which research from Europe and North America is published

	Number of publications (%)	Weighted average SNIP 2011	Aggregate SNIP 2011	Number of publications with SNIP 2011 ≥ 1 (%)
Europe	25,765 (86%)	1.42	36,528	15,562 (60%)
North America	21,670 (87%)	1.83	39,551	15,174 (70%)

Table 5. Descriptive statistics of global type 2 diabetes innovation system (n = 741,153)

Variable	Mean	Std. Dev.	Min.	Max.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. No. of co-publications$_{ij}$	0.052	0.602	0.000	161.000	1.000									
2. Intra-country dummy$_{ij}$	0.095	0.294	0.000	1.000	0.159	1.000								
3. Geographical proximity$_{ij}$	0.001	0.022	0.000	1.000	0.255	0.163	1.000							
4. Social proximity$_{ij}$	0.007	0.084	0.000	3.761	0.440	0.155	0.227	1.000						
5. Cognitive proximity$_{ij}$	0.072	0.134	0.000	1.000	0.146	0.160	0.052	0.203	1.000					
6. Institutional proximity$_{ij}$	0.407	0.491	0.000	1.000	0.012	−0.008	−0.007	0.013	0.008	1.000				
7. Organizational proximity$_{ij}$	0.001	0.036	0.000	1.000	0.124	0.103	0.191	0.117	0.039	0.027	1.000			
8. Mass$_{ij}$	6.498	1.087	4.605	12.180	0.136	0.047	0.013	0.132	0.380	0.075	−0.009	1.000		
9. Transitivity$_{ij}$	0.259	1.101	0.000	45.000	0.332	0.236	0.093	0.530	0.393	0.026	0.041	0.357	1.000	
10. Expected collaboration propensity$_{ij}$	26.888	21.167	0.000	289.239	0.196	0.122	0.022	0.178	0.429	0.025	0.003	0.792	0.482	1.000

published in that journal, overall they provide a rough indication of the quality of the scientific media outlets in which researchers of a particular innovation system can publish their results.

It should be noted that journal impact information is not available for all publications. The second column in Table 4 shows, however, that the number and percentage of publications for which such information is available are roughly the same for Europe (25,765; 86%) and North America (21,670; 87%). The next three columns show that on average, North America publishes in slightly higher quality journals than does Europe. All three columns provide information on journal impact derived from the source normalized impact per paper (SNIP) (Moed 2011). The weighted average SNIP (third column Table 4) and the aggregate SNIP (4th column Table 4) are both higher for North America than for Europe. However, in total, Europe and North America publish a roughly equal number of publications in journals where SNIP > 1 (15,562 publications and 15,164 publications, respectively); although this gives North America publishing a larger share of publications (70%) in journals with SNIP ≥ 1 than Europe (60%). Overall, and although there are differences in the quantity and quality of research produced in Europe as compared to North America, we have no reasons to expect that the order of magnitude is such that these differences are likely to have a large impact on collaborative research in both innovation systems being organized differently.

Turning to our main empirical results, Table 5 first shows the descriptive statistics of and correlations among the variables included in the analysis of the global innovation system. A first observation is that organizational proximity is very rare. From the mean of this dummy variable, one can read that only 0.1% of organization pairs are organizationally proximate. This means that the large majority of organization branches also belong to different overarching organizations. Concerning correlations, the intra-country dummy is obviously correlated with geographical proximity as both reflect proximity in physical space, albeit in fundamentally different ways. More interesting, the highest positive correlations are found between some of the other proximity variables. Geographical proximity is correlated with social and organizational proximity. The first correlation seems to suggest that social proximity is more easily maintained when actors are geographically proximate. The second correlation reflects the fact that organizations belonging to the same parent organization are often co-located. In particular, this holds true for academic hospitals and their corresponding universities. Further, social proximity is correlated with cognitive proximity, which seems to suggest that repeated ties occur more often within disciplines than across disciplinary boundaries. This is in line with the Mode 1 versus Mode 2 distinction made by Gibbons et al. (1994), who emphasized that transdisciplinary projects often occur in one-off projects.

Table 6 shows the same set of descriptive statistics, but compares the North American and European innovation systems in type 2 diabetes research. Given that 359 organizations operate in the field of type 2 diabetes

Table 6. Descriptive statistics of European and North American innovation system (n = 163,051 (EU and NA); n = 98,790 (EU); n = 64,261 (NA))

Europe and North America combined	Mean	Std. Dev.	Min.	Max.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. No. of co-publications$_{ij}$	0.155	1.160	0.000	161.000	1.000										
2. Intra-country dummy$_{ij}$	0.391	0.488	0.000	1.000	0.113	1.000									
3. Geographical proximity$_{ij}$	0.005	0.039	0.000	1.000	0.248	0.107	1.000								
4. Social proximity$_{ij}$	0.022	0.149	0.000	3.761	0.472	0.120	0.245	1.000							
5. Cognitive proximity$_{ij}$	0.117	0.168	0.000	1.000	0.176	0.097	0.051	0.257	1.000						
6. Institutional proximity$_{ij}$	0.381	0.486	0.000	1.000	0.023	−0.019	−0.012	0.024	0.016	1.000					
7. Organizational proximity$_{ij}$	0.005	0.071	0.000	1.000	0.113	0.088	0.193	0.124	0.046	0.060	1.000				
8. Mass$_{ij}$	6.710	1.142	4.605	11.933	0.200	−0.008	0.017	0.200	0.423	0.081	−0.032	1.000			
9. Transitivity$_{ij}$	0.650	1.883	0.000	45.000	0.365	0.179	0.088	0.569	0.444	0.052	0.037	0.453	1.000		
10. Expected collaboration propensity$_{ij}$	34.363	24.170	0.000	276.597	0.266	0.073	0.024	0.247	0.433	0.064	−0.016	0.832	0.566	1.000	
11. EU dummy	0.606	0.489	0.000	1.000	−0.022	−0.686	−0.006	−0.035	0.065	0.016	−0.049	−0.015	−0.132	−0.091	1.000

Europe	Mean	Std. Dev.	Min.	Max.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. No. of co-publications$_{ij}$	0.134	1.092	0.000	104.000	1.000									
2. Intra-country dummy$_{ij}$	0.121	0.326	0.000	1.000	0.212	1.000								
3. Geographical proximity$_{ij}$	0.004	0.038	0.000	1.000	0.278	0.222	1.000							
4. Social proximity$_{ij}$	0.018	0.135	0.000	3.045	0.484	0.211	0.291	1.000						
5. Cognitive proximity$_{ij}$	0.109	0.156	0.000	1.000	0.173	0.159	0.064	0.250	1.000					
6. Institutional proximity$_{ij}$	0.387	0.487	0.000	1.000	0.017	0.015	-0.016	0.020	0.028	1.000				
7. Organizational proximity$_{ij}$	0.002	0.048	0.000	1.000	0.194	0.124	0.350	0.230	0.058	0.013	1.000			
8. Mass$_{ij}$	6.697	1.082	4.605	11.685	0.169	-0.013	0.022	0.162	0.405	0.050	0.003	1.000		
9. Transitivity$_{ij}$	0.450	1.478	0.000	39.000	0.341	0.198	0.121	0.580	0.409	0.038	0.078	0.370	1.000	
10. Expected collaboration propensity$_{ij}$	32.588	20.733	0.000	251.960	0.218	0.010	0.027	0.191	0.415	0.039	0.009	0.813	0.436	1.000

North America	Mean	Std. Dev.	Min.	Max.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. No. of co-publications$_{ij}$	0.187	1.257	0.000	161.000	1.000									
2. Intra-country dummy$_{ij}$	0.806	0.396	0.000	1.000	0.051	1.000								
3. Geographical proximity$_{ij}$	0.005	0.041	0.000	1.000	0.211	0.047	1.000							
4. Social proximity$_{ij}$	0.028	0.167	0.000	3.761	0.459	0.050	0.193	1.000						
5. Cognitive proximity$_{ij}$	0.131	0.185	0.000	1.000	0.178	−0.020	0.034	0.262	1.000					
6. Institutional proximity$_{ij}$	0.371	0.483	0.000	1.000	0.032	−0.044	−0.006	0.031	0.002	1.000				
7. Organizational proximity$_{ij}$	0.009	0.096	0.000	1.000	0.065	0.047	0.098	0.062	0.035	0.104	1.000			
8. Mass$_{ij}$	6.731	1.228	4.605	11.933	0.236	−0.037	0.011	0.241	0.443	0.124	−0.059	1.000		
9. Transitivity$_{ij}$	0.958	2.342	0.000	45.000	0.395	0.064	0.060	0.569	0.473	0.074	0.008	0.540	1.000	
10. Expected collaboration propensity$_{ij}$	37.092	28.445	0.000	276.597	0.314	0.020	0.020	0.294	0.446	0.098	−0.037	0.862	0.652	1.000

research in North America and 445 in Europe, these two territorial innovation systems are characterized by, respectively, 64,261 and 98,790 organization pairs summing up to a total of 163,051 organization pairs for the North American and European innovation systems in type 2 diabetes research combined. On average, and notwithstanding the extremely skewed distribution of co-publications, any two organizations in North America collaborate more (0.187 times) than any two organizations in Europe (0.134 times).

First, comparing the descriptive statistics between Europe and North America we observe that the means and standard deviations are fairly similar except for the intra-country dummy, social proximity and organizational proximity, for which the means are much larger in North America. The difference in mean for the intra-country dummy is due to the inclusion of far less countries in the North American innovation system (only the USA and Canada) as compared to the European innovation system (15 EU countries and Switzerland). Despite the difference in the mean, we chose to consider this variable in our comparison. However, first, the potentially hampering impact of administrative borders on research collaboration is vastly different (both empirically and theoretically) from that of kilometric distance (Hoekman et al. 2010); and second, especially with respect to European science policy, the role of country borders is deemed to be particularly important as part of the construction of a single European Research Area (Chessa et al. 2013).

The difference in mean for organizational proximity suggests that the North American innovation system is structurally more hierarchically organized than the European innovation system with relatively more

organization branches operating in this field that belong to the same overarching organization. The mean of social proximity being higher in North America than in Europe points at past collaborations having taken place in the former more than in the latter territorial innovation system. For social proximity the difference in mean is on a par with the difference in standard deviation between Europe and North America.

Second, again with some notable exceptions, the correlations among variables are also comparable between Europe and North America. For Europe, however, the number of co-publications between organization pairs correlates substantially more with the intra-country dummy and organizational proximity than for North America. The latter finding is especially striking given that, on average, more organization branches in North America are organizationally proximate than in Europe. Conversely, for North America, institutional proximity correlates substantially more with the number of co-publications than in Europe. This seems to run counter to the prevailing view that North America is better at translating basic research into commercial innovation than is Europe (Dosi et al. 2006).

Table 7 then shows the estimates for three gravity equation models, with all three models successively showing the negative binomial part, the zero-inflated part, and some additional statistics. The latter include the Vuong test (Vuong 1989), testing whether or not the choice of the zero-inflated negative binomial regression model is appropriate. Overall, for all three models, the results of the Vuong test indicate that the zero-inflated negative binomial regression model fits the data best.

Model 1 in Table 7 shows the results for the global analysis taking into account all organizations worldwide that publish on type 2 diabetes. It is clear from the results that all five proximity dimensions are positive and significant, reflecting that all five dimensions contribute to facilitating research collaborations that lead to co-publications. This result suggests that, on the aggregate level, evidence of a Mode 2 type of pattern of collaborative science is generally absent. That is, distant organizations in any of the five dimensions are less prone to collaborate than close organizations. This does not mean that for each individual organization particular forms of distance may not motivate a particular research collaboration. Rather, when aggregating all collaborations, the effect of such motivations disappears, given that in most of collaborations proximity rather than distance is driving the formation of research partnerships. Hence, overall, the global innovation system on type 2 diabetes research is characterized by Mode 1 rather than Mode 2 knowledge production.

Table 7. Zero-inflated negative binomial regression results (dependent variable: No. of co-publications$_{ij}$)

	Model 1: Global				Model 2: EU and North America				Model 3: EU versus North America			
Negative binomial part	Coef.	Std. Err.	Z score	P value	Coef.	Std. Err.	Z score	P value	Coef.	Std. Err.	Z score	P value
Intra-country dummy	1.87	0.02	78.92	0.000	1.28	0.03	44.76	0.000	1.03	0.05	18.84	0.000
Geographical proximity	3.27	0.16	20.90	0.000	3.18	0.16	20.02	0.000	4.45	0.31	14.47	0.000
Social proximity	1.09	0.04	27.70	0.000	0.99	0.04	24.56	0.000	0.87	0.05	18.39	0.000
Cognitive proximity	0.65	0.06	10.13	0.000	0.49	0.07	6.70	0.000	0.20	0.09	2.25	0.024
Institutional proximity	0.15	0.02	8.16	0.000	0.21	0.02	8.42	0.000	1.56	0.15	10.67	0.000
Organizational proximity	1.16	0.11	10.81	0.000	0.90	0.10	9.02	0.000	0.17	0.03	5.54	0.000
Intra-country dummy×EU									1.22	0.07	18.15	0.000
Geographical proximity×EU									−2.72	0.34	−8.11	0.000
Social proximity×EU									−0.14	0.07	−2.16	0.031
Cognitive proximity×EU									0.14	0.12	1.17	0.243
Institutional proximity×EU									−0.02	0.05	−0.43	0.664
Organizational proximity×EU									−1.15	0.18	−6.34	0.000
EU dummy									0.13	0.07	1.95	0.051
Mass	0.39	0.02	25.69	0.000	0.43	0.02	23.87	0.000	0.53	0.02	32.30	0.000
Transitivity	0.03	0.00	6.19	0.000	0.01	0.00	2.11	0.035	0.02	0.00	5.48	0.000
Constant	−5.65	0.13	42.61	0.000	−5.49	0.16	−34.66	0.000	−6.44	0.15	−42.28	0.000
Zero-inflated part												
Expected collaboration prop.	−0.06	0.00	−58.78	0.000	−0.06	0.00	−42.84	0.000	−0.06	0.00	−41.29	0.000
Constant	3.04	0.05	56.28	0.000	2.75	0.07	41.60	0.000	2.56	0.07	39.10	0.000
lnalpha	0.72	0.03	23.50	0.000	0.43	0.04	9.82	0.000	0.26	0.04	6.97	0.000
Vuong statistic	25.49			0.000	19.36			0.000	20.00			0.000
Pseudo log likelihood	−85876.8				−44216.41				−43168.0			
Observations	741,153				163,051				163,051			
Non-zero observations	19,601				11,393				11,393			
McFadden's Adj. R^2	0.25				0.23				0.25			

In models 2 and 3 we only included organizations located in North America and Europe and included interaction terms to analyse whether or not proximity dimensions had a differential effect on establishing collaborations within Europe as compared to within North America. First, in model 2 we regenerate the results of model 1 but then restrict them to organization pairs that are located in Europe and North America only. As with the global innovation system in the field of type 2 diabetes, collaboration is driven by proximity in all five dimensions. Hence, taken together, the European and North American situations are generally characterized by Mode 1 instead of Mode 2 knowledge production.[10]

Second, model 3 introduces a dummy variable for intra-European research collaborations and interaction terms to analyse whether or not proximity dimensions had a differential effect on establishing collaborations within Europe as compared to within North America. The results reveal three striking results in the comparison of the European and North American innovation systems. First, the intra-country interaction effect highlights that the bias towards national rather than international collaboration is much stronger in Europe than in North America (read USA–Canada collaboration). Second, geographical, organizational and social proximity play less of a role in Europe than in North America. Third, cognitive and institutional proximity are equally important in Europe and North America. Taken together, and apart from some similarities, the comparative analysis thus shows that there are indeed significant differences between the two territorial innovation systems.

The interpretation of the results warrants a fine-grained analysis in its own right. However, some suggestive interpretations can already be made. The first result, regarding the relative bias towards national collaboration in Europe, may well reflect larger linguistic and cultural variety within Europe than in North America (Crescenzi et al. 2007). What holds is that the North American innovation system is more integrated than the European innovation system when it comes to crossing national boundaries. This result further supports the doubts that have recently been raised about the accomplishment of a European Research Area (Chessa et al. 2013).

In contrast, European collaborations in type 2 diabetes research are less restricted by geographical, organizational and social proximity. The result of social and organizational proximity being of less importance in Europe than in North America suggests that in North America science structures are more stratified than those in Europe, in the sense that North American collaboration patterns are more responsive to avoiding opportunism. On the one hand, it might be argued though that the lesser role played by organizational proximity in Europe is mainly due to the relative absence of overarching organizations in that system. It should be noted that the average number of organization pairs that are organizationally proximate is substantially higher in North America than in Europe. On the other hand, the single correlation of organizational proximity with the number of co-publications between organization pairs was lower for North America than for Europe. Also, and although warranting much more research, our tentative conclusion seems to be supported by claims that the US innovation system is very stratified (Jones et al. 2008) and the European innovation system is fairly cohesive (Hoekman et al. 2009).

Finally, and especially interesting in the light of science policy debates, the third result holds that no differences are found between the two innovation systems in terms of the roles played by institutional and cognitive proximity. Although both forms of proximity are of importance in shaping interactions in European and North American science, it does not seem to be the case that institutional or disciplinary differences are more easily bridged in either of these two systems. As such, the widely held conviction that Europe is worse in translating basic research into commercial innovation seems not to be supported by our findings (Dosi et al. 2006).

Overall, what holds for our comparative analysis is that: first, both the European and North American innovation system in type 2 diabetes research are generally characterized by Mode 1 instead of Mode 2 knowledge production. Second, neither of the two territorial innovation systems can be considered more Mode 2 than the other innovation system in all its aspects. Rather, differences and similarities exist between the two innovation systems as to the role played by particular proximity dimensions steering collaborative research in Europe as compared to North America.

5. Conclusions

We proposed an analytical framework based on the proximity concept to analyse and compare territorial innovation systems in the research field concerned with type 2 diabetes. Where innovation system analysis tends to focus on inter-organizational and cross-institutional dimensions of collaboration, we propose a richer framework based on five dimensions taken from the work by Boschma (2005) on proximity and collaboration. We have been able to show how the five proximity dimensions map almost one-to-one to various aspects of distributed 'Mode 2' knowledge production (Gibbons et al. 1994). Thus, the proximity framework allows for an analytical operationalization of the Mode 2 concept. In particular, our proposed framework for thinking of innovation systems in terms of Mode 1 versus Mode 2 knowledge production along five different dimensions of proximity allows both for an empirical characterization as well as a comparative analysis of innovation systems.

In the empirical analysis we assessed the extent to which each proximity dimension affected the intensity of collaboration between organizations working on type 2 diabetes. Our main empirical results are that: first, the global innovation system in the field of type 2 diabetes research is generally characterized by Mode 1 rather than Mode 2 knowledge production as evidenced by the importance of proximity rather than distance in all five dimensions. Second, differences and similarities exist between the North American and European innovation systems in the field of type 2 diabetes research. With respect to the latter finding, we observe that while geographical, social and organizational proximity play a relatively smaller role in Europe as compared to North America, and national boundaries play a relatively smaller role in North America as compared to Europe; cognitive and institutional proximity are equally important in Europe and North America. Hence, we conclude that Europe is more Mode 2 than North America in terms of the social and organizational aspects of collaborative knowledge production, while Europe and North America are fairly equal in terms of their Mode 2-ness when it comes to the institutional and cognitive aspects of collaborative knowledge production.

Our framework allows for a systematic characterization and comparison of territorial innovation systems. A next step in applying our framework would be to undertake a comparative analysis in combination with a performance analysis in the form of a 'benchmarking exercise'.[11] By comparing different systems in terms of their modes of knowledge production and relating these modes to differences in performance, hypotheses regarding the functioning of different territorial innovation systems could be analysed (Arndt and Sternberg 2000; Carlsson et al. 2002; Tödtling et al. 2009); for example, in terms global market shares in research output, inventions and innovations (Bonaccorsi 2008). A simple hypothesis would posit that the higher the level of 'Mode-2-ness', the better the performance of a particular system. This claim seems to be inherent to the Mode 2 concept as introduced by Gibbons et al. (1994), as their description of Mode 2 knowledge production suggests that Mode 2 is better able to solve complex societal problems than Mode 1. Yet, it is for this implicit normative reasoning that the Mode 2 concept has been criticized (Godin 1998; Shinn 2002). Alternatively, one could argue that proximity along at least some dimensions is required to reduce uncertainties and avoid conflicts in research collaboration (Boschma 2005; Balland 2012). So long as empirical research is not carried out in a systematic way such that evidence can be compared and accumulated across different units of comparison, the debate remains empirically ill-informed. We hope that, as a first step, our proximity framework can also serve as a tool to assess the performance of different types of innovation systems.

Of course, the conclusions drawn here follow from assessing one particular case and therefore need not be applicable to research at large: neither at the global scale, nor for the two territorial innovation systems compared in this study. As argued before, however, if anywhere, the premises of Mode 2 knowledge production can be expected to hold especially for the field of type 2 diabetes research. Hence, the fact that we do not find strong evidence for Mode 2 knowledge production in this field casts considerable doubt on the prevalence of Mode 2 knowledge production as an organizing principle of contemporary innovation systems in general. All this is not to suggest that policy-makers should focus more on stimulating Mode 2 research *per se*. The extent to which Mode 2 knowledge production is desirable from a normative point of view is a different matter altogether (Shinn 2002). At the very least, however, and which particular organizing principle is most desirable from a normative point of view, our proposed framework for thinking of innovation systems in terms of Mode 1 versus Mode 2 knowledge production along five dimensions of proximity allows for both a systematic characterization and a systematic comparison of innovation systems.

Acknowledgements

We thank Pierre-Alexandre Balland, Robin Cowan, Marie Ferru, Gimon de Graaf, Bertien Hart, Jarno Hoekman, the attendants of the presentation given at the EuroLIO conference on "Geography of innovation"(Saint Étienne, January 2012), and two anonymous referees for their help and suggestions. The usual disclaimer applies. In addition, the opinions expressed are those of the authors only and should not be considered as representative of the official position of the organizations to which they are affiliated.

Funding

This work was supported by the Netherlands Organisation for Scientific Research (NWO) under the VIDI programma (grant number 452-06-005 to K.F. and S.H.).

Notes

1. When looking at collaboration as captured by co-publication data, some may prefer to speak of science systems rather than innovation systems. We chose to use the more common term of innovation system here, as scientific research is an integral component of medical innovation.
2. In contrast, this trend is rather weak in the humanities (Wuchty et al. 2007).
3. Some prefer to speak of physical proximity in this context (Frenken et al. 2009).

4. The prime medical issue of diabetes is described as hyperglycaemia, that is, the bodily condition in which an excessive amount of glucose circulates the blood. A state of hyperglycaemia is problematic in that it is indicative of the blood delivering too little energy for the organs to function properly. When this state continues for longer periods of time, this may lead to severe complications. Among the complications of hyperglycaemia, diabetic coma can be most acute. Other, more common, complications involve a loss of sight and severe foot ulcers. Although largely similar in their complications, we can grossly distinguish two most prevalent types of diabetes (type 1 diabetes and type 2 diabetes). On the one hand, type 1 diabetes is generally taken to reflect a state in which the body is insufficiently capable of producing hormones that enable the transformation of glucose into energy (Tattersall 2009). On the other hand, type 2 diabetes is generally taken to reflect a state in which the body is insufficiently capable of metabolizing (i.e. transforming) insulin properly thus leading to an inadequate bodily uptake of energy (Tattersall 2009). Regardless of the bodily capacity to produce insulin (characteristic of type 1 diabetes), type 2 diabetes is primarily characterized by a resistance or deficiency of the body in using insulin. No treatment has been proposed for type 2 diabetes so far that is fully capable of improving the bodily capacity to metabolize insulin on a continuous basis, in a similar fashion as insulin itself has been proposed as a continuous treatment option for patients with type 1 diabetes.
5. Apart from taking 50 km we also experimented with 30 km and 70 km as our geographical boundary of the organization. These alternative geographical boundaries did not alter the results of our analyses.
6. It should be noted that our delineation of organization branches depends crucially upon our ideas on what constitutes an organization branch in the first place. A definite and objective delineation of organization branches is difficult, if not impossible, to achieve (Hardeman 2013). We take the organization branch within innovation systems as constituting:

   > . . . a dense network at the center of a web of relationships. (Badaracco 1991: 314)

   Hence, our treatment of the organization closely follows a relational perspective on organizations as they are embedded in territories (Dicken and Malmberg 2001). As such we believe it fits perfectly within the larger multi-dimensional proximity framework.
7. The distribution of the number of co-publications between organizations in the period 1996–2002 is extremely skewed. Hence, we log-transformed this variable. As to make log-transformation possible for all values (including zeroes), all values for the number of co-publications in the period 1996–2002 were added with 1. By taking into account past co-publications the argument holds that when past co-publications do not steer contemporary co-publication we can indeed speak of flexible, open, and temporary networks and, hence, Mode 2 knowledge production. We choose this particular time-lag structure (i.e. 1996–2002) in order to leave the two time periods considered (i.e. past and contemporary research collaborations) roughly equal in years and hence what is considered open versus non-open, flexible versus inflexible, temporary versus structural were roughly the same.
8. As for other proximity dimensions, the institutional and organizational proximity are orthogonal. That is, any two organizations (defined at the branch level) may be organizationally proximate yet institutionally distant and vice versa.
9. Zero-inflated negative binomial regression modelling is used when the dependent variable reflects count data with an excessive number of zeroes (Long 1997). In the present paper the excessive zeroes are called structural zeroes: structural because they could not conceivably have taken any value other than zero. Zero-inflated negative binomial regression models allow for a separate modelling of the process generating structural zeroes *vis-à-vis* the process generating the counts of the dependent variable (including zeroes that could have taken a non-zero value). While the zero-inflated part models the process generating structural zeroes independent from the process generating the counts of the dependent variable, the negative binomial part models the counts of the dependent variable conditional on the likelihood that the dependent variable can be non-zero.

   Theoretically, proximity is neither a necessary nor a sufficient condition for collaborative science to take place (Boschma 2005). In principle, collaborative science can take place over longer distances. But, without sufficient resources to collaborate, proximity alone is not enough for collaboration. As such, the process generating excessive zeroes in collaborative science should indeed be modelled differently from the process generating the counts of our dependent variable. Hence, in the zero-inflated part we do not include proximity variables but only a variable that captures the extent to which any two organizations can possibly collaborate in the first place (i.e. 'expected collaboration propensity') and only in the negative binomial part do we specify a gravity model that includes variables on multiple proximity dimensions. If we had included the same variables in the zero-inflated part of the model as we did in the negative binomial part, we would have run the risk

of 'over-deflating' some observations that take a value of zero in their dependent variable. In other words, also including all variables measuring the various proximity dimensions in the zero-inflated part would lead to biased estimates in the main parameters of interest.

Finally, it should be noted that including the same set of variables in the zero-inflated part as in the negative binomial part basically amounts to saying that the mechanisms underlying the generation of structural zeroes is essentially the same as the process generating the counts of the dependent variable; leaving the use of a zero-inflated negative binomial regression modelling approach instead of a simple negative binomial regression modelling approach obsolete (Allison 2012, esp. Chap. 9).

10. In Note 9 we motivated our choice for specifying the zero-inflated part of the regression models differently from the negative binomial part. Notwithstanding our arguments for specifying the regression models in this particular way, as an extra robustness check we estimated two alternative regression models. One included the same set of variables in the zero-inflated part as are included in the negative binomial part. Another left out the zero-inflated part altogether and thus estimated an ordinary negative binomial regression model. The outcomes of these models remain largely the same with most parameters neither changing in direction nor in significance. The exception is the parameter estimating the effect of cognitive proximity which turns out to be no longer significant when we include the same set of variables in the zero-inflated part as we did in the negative binomial part of our models. However, it should be noted that the parameter estimating the effect of cognitive proximity again turns significant when we estimate an ordinary (i.e. not zero-inflated) negative binomial regression model. Taken together, these mixed results warrant a more in-depth discussion of the proper specification of (zero-inflated) negative binomial regression models. Such a discussion is, however, beyond the scope of the current study. That having been said, we thank an anonymous reviewer for pointing out this important issue.
11. From a methodological point of view, another next step in refining the comparative analysis of territorial innovation systems would be to take into account the problem of spatial autocorrelation of flows as a potential bias. Research in spatial econometrics has addressed this issue in the context of R&D collaborations (Scherngell and Lata 2013; Chun 2013). Although, thus far only research that takes regions as the basic unit of analysis has been taken into account, a natural extension of this methodology would be to also consider it in the context of organization-level analysis.

References

Adams, J., Black, G., Clemmons, R. and Stephan, P. (2005) 'Scientific teams and institutional collaborations: Evidence from U.S. universities, 1981–1999', *Research Policy*, 34: 259–85.

Allison, P. (1999) *Logistic Regression Using SAS®: Theory and Application*. Cary, NC: SAS Institute Inc.

Arndt, O. and Sternberg, R. (2000) 'Do manufacturing firms profit from intraregional innovation linkages? An empirical based answer', *European Planning Studies*, 8: 465–85.

Badaracco, J. L. Jr. (1991) 'The boundaries of the firm'. In: Etzioni, A. and Lawrence, P. R. (eds) *Socio-Economics: Toward a New Synthesis*, pp. 293–327. New York: M. E. Sharpe Inc.

Balland, P. A. (2012) 'Proximity and the evolution of collaboration networks: Evidence from R&D projects within the GNSS industry', *Regional Studies*, 46: 741–56.

Barry, A., Born, G. and Weszkalnys, G. (2008) 'Logics of interdisciplinarity', *Economy and Society*, 37: 20–49.

Bathelt, H., Malmberg, A. and Maskell, P. (2004) 'Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation', *Progress in Human Geography*, 28: 31–56.

Bonaccorsi, A. (2007) 'Explaining poor performance of European science: Institutions versus policies', *Science and Public Policy*, 34: 303–16.

—— (2008) 'Search regimes and the industrial dynamics of science', *Minerva*, 46: 285–315.

—— (2010) 'New forms of complementarity in science', *Minerva*, 48: 355–87.

Boschma, R. A. (2005) 'Proximity and innovation: A critical assessment', *Regional Studies*, 39: 61–74.

Breschi, S. and Lissoni, F. (2009) 'Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows', *Journal of Economic Geography*, 9: 439–68.

——, —— and Malerba, F. (2003) 'Knowledge-relatedness in firm technological diversification', *Research Policy*, 32: 69–87.

Broekel, T., Balland, P. A., Burger, M. and Van Oort, F. (2014) 'Modeling knowledge networks in economic geography: a discussion of four empirical strategies', *Annals of Regional Science*, Online first, DOI:10.1007/s00168-014-0616-2.

Burger, M. J., Van Oort, F. G. and Linders, G. J. M. (2009) 'On the specification of the gravity model of trade: Zeros, excess zeros and zero-inflated estimation', *Spatial Economic Analysis*, 4: 167–87.

Carlsson, B. (2006) 'Internationalization of innovation systems: A survey of the literature', *Research Policy*, 35: 56–67.

——, Jacobsson, S., Holmén, M. and Rickne, A. (2002) 'Innovation systems: Analytical and methodological issues', *Research Policy*, 31: 233–45.

Castells, M. (1996) *The Rise of the Network Society*. Oxford, UK: Blackwell.

Chesbrough, H. (2003) *Open Innovation: The New Imperative for Creating and Profiting From Technology*. Cambridge, MA: Harvard Business School Press.

Chessa, A., Morescalchi, A., Pammolli, F., Penner, O. *et al.* (2013) 'Is Europe evolving toward an integrated research area?', *Science*, 339: 650–1.

Chun, Y. (2013) 'Network autocorrelation and spatial filtering'. In: Scherngell, T. (ed.) *The Geography of Networks and R&D Collaborations*, pp. 99–113. Heidelberg, Germany: Springer.

Cohen, W. M. and Levinthal, D. A. (1990) 'Absorptive capacity: A new perspective on learning and innovation', *Administrative Science Quarterly*, 35: 128–52.

Collins, H. M. (1985) *Changing Order: Replication and Induction in Scientific Practice*. London: Sage.

Cooke, P., Uranga, M. G. and Etxebarria, G. (1998) 'Regional systems of innovation: An evolutionary perspective', *Environment and Planning A*, 30: 1563–84.

Crescenzi, R., Rodriguez-Pose, A. and Storper, M. (2007) 'The territorial dynamics of innovation: A Europe–United States comparative analysis', *Journal of Economic Geography*, 7: 673–709.

Danaei, G., Finucane, M. M., Lu, Y., Singh, G. M. *et al.* (2011) 'National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: Systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants', *Lancet*, 378: 31–40.

Dasgupta, P. and David, P. A. (1994) 'Toward a new economics of science', *Research Policy*, 23: 487–521.

Dicken, P. and Malmberg, A. (2001) 'Firms in territories: A relational perspective', *Economic Geography*, 77: 345–63.

Dosi, G., Llerena, P. and Sylos Labini, M. (2006) 'Science–technology–industry links and the 'European paradox': Some notes on the dynamics of scientific and technological research in Europe', *Research Policy*, 35: 1450–64.

Elsevier Pharma Development Group. (2009) *EMTREE: The Life Science Thesaurus*. Version 8.0. Amsterdam: Elsevier.

Etzkowitz, H. and 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: 109–23.

Freeman, C. (1987) *Technology Policy and Economic Performance: Lessons from Japan*. London: Pinter.

Frenken, K., Hardeman, S. and Hoekman, J. (2009) 'Spatial scientometrics: Towards a cumulative research program', *Journal of Informetrics*, 3: 222–32.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S. *et al.* (1994) *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. London: Sage.

Godin, B. (1998) 'Writing performative history: The new New Atlantis?', *Social Studies of Science*, 28: 465–83.

Granovetter, M. (1985) 'Economic action and social structure: The problem of embeddedness', *American Journal of Sociology*, 91: 481–510.

Hardeman, S. (2013) 'Organization level research in scientometrics: A plea for an explicit pragmatic approach', *Scientometrics*, 94: 1175–94.

Hekkert, M. P., Suurs, R. A. A., Negro, S. O., Smits, R. E. H. M. *et al.* (2007) 'Functions of innovation systems: A new approach for analyzing technological change', *Technological Forecasting and Social Change*, 74: 413–32.

Hessels, L. and Van Lente, H. (2008) 'Re-thinking new knowledge production: A literature review and a research agenda', *Research Policy*, 37: 740–60.

Hicks, D. M. and Katz, J. S. (1996) 'Where is science going?', *Science, Technology, and Human Values*, 21: 379–406.

Hoekman, J., Frenken, K., Scherngell, T. and Tijssen, R. (2013) 'Acquisition of European research funds and its effect on international scientific collaboration', *Journal of Economic Geography*, 13: 23–52.

——, —— and Van Oort, F. (2009) 'The geography of collaborative knowledge production in Europe', *Annals of Regional Science*, 43: 721–38.

Hurley, D. (2011) *Diabetes Rising: How a Rare Disease Became a Modern Pandemic, and What To Do About It*. New York: Kaplan Publishing.

Jaffe, A. B. (1986) 'Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market', *American Economic Review*, 76: 984–1001.

Jones, B. F., Wuchty, S. and Uzzy, B. (2008) 'Multi-university research teams: Shifting impact, geography and stratification in science', *Science*, 322: 1259–62.

Katz, J. S. (1994) 'Geographical proximity and scientific collaboration', *Scientometrics*, 31: 31–43.

—— and Martin, B. R. (1997) 'What is research collaboration?', *Research Policy*, 26: 1–18.

Knoben, J. and Oerlemans, L. A. (2006) 'Proximity and inter-organizational collaboration: A literature review', *International Journal of Management Reviews*, 8: 71–89.

Lane, P. J. and Lubatkin, M. (1998) 'Relative absorptive capacity and interorganizational learning', *Strategic Management Journal*, 19: 461–77.

Leydesdorff, L. and Etzkowitz, H. (1996) 'Emergence of a triple helix of university–industry–government relations', *Science and Public Policy*, 23: 279–86.

—— and 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: 1622–34.

Lincoln, J. R. (1984) 'Analysing relations in dyads: Problems, models and an application to interorganizational research', *Sociological Methods and Research*, 13: 45–76.

Long, J. S. (1997) *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage.

Lundberg, J., Thomson, G., Lundkvist, I., Skar, J. *et al.* (2006) 'Collaboration uncovered: Exploring the adequacy of measuring university–industry collaboration through co-authorship and co-funding', *Scientometrics*, 69: 575–89.

Lundvall, B. (1988) 'Innovation as an interactive process: From user–producer interaction to the national system of innovation'. In: Dosi, G., Freeman, C., Nelson, R., Silverberg, G. and Soete, L. (eds) *Technical Change and Economic Theory*, pp. 349–69. London: Pinter.

—— (2007) 'National innovation systems–analytical concept and development tool', *Industry and Innovation*, 14: 95–119.

Maggioni, M. and Uberti, T. (2009) 'Knowledge networks across Europe: Which distance matters?', *Annals of Regional Science*, 43: 691–720.

Moed, H. F. (2011) 'Letter to the Editor: The source normalized impact per paper is a valid and sophisticated indicator of journal citation impact', *Journal of the American Society for Information Science and Technology*, 62: 211–3.

Morgan, K. (2004) 'The exaggerated death of geography: Learning, proximity and territorial innovation systems', *Journal of Economic Geography*, 4: 3–21.

National Institutes of Health. (2004) *From Clinical Trials to Community: The Science of Translating Diabetes and Obesity Research*. Bethesda, MD: National Institutes of Health.

Newman, M. E. J. (2001) 'The structure of scientific collaboration networks', *Proceedings of the National Academy of Sciences of the United States of America*, 98: 404–9.

Nooteboom, B. (1999) *Inter-Firm Alliances: Analysis and Design*. London: Routledge.

Nowotny, H., Scott, P. and Gibbons, M. (2003) 'Introduction: 'Mode 2' revisited: The new production of knowledge', *Minerva*, 41: 179–94.

Parsons, T. (1956a) 'Suggestions for a sociological approach to the theory of organizations – I', *Administrative Science Quarterly*, 1: 63–85.

—— (1956b) 'Suggestions for a sociological approach to the theory of organizations – II', *Administrative Science Quarterly*, 1: 225–39.

Ponds, R., Van Oort, F. G. and Frenken, K. (2007) 'The geographical and institutional proximity of research collaboration', *Papers in Regional Science*, 86: 423–43.

Price, D. D. S. (1963) *Little Science, Big Science*. New York: Columbia University Press.
Rallet, A. (1993) 'Choix de proximité et processus d'innovation technologique', *Revue d'Economie Régionale et Urbaine*, 3: 365–86.
—— and Torre, A. (1999) 'Is geographical proximity necessary in the innovation networks in the era of global economy?', *Geojournal*, 49: 373–80.
Scherngell, T. and Lata, R. (2013) 'Towards an integrated European Research Area? Findings from eigenvector spatially filtered spatial interaction models using European Framework Programme data', *Papers in Regional Science*, 92: 555–77.
Seglen, P. O. (1997) 'Why the impact factor of journals should not be used for evaluating research', *British Medical Journal*, 314: 498–502.
Shinn, T. (2002) 'The triple helix and new production of knowledge: Prepackaged thinking on science and technology', *Social Studies of Science*, 32: 599–614.
Slaughter, S. and Leslie, L. (1997) *Academic Capitalism: Politics, Policies and the Entrepreneurial University*. Baltimore, MD: Johns Hopkins University Press.
Snijders, T. A. (2011) 'Statistical models for social networks', *Annual Review of Sociology*, 37: 131–53.
Stuart, T. (1998) 'Network positions and propensities to collaborate: An investigation of strategic alliance formation in a high-technology industry', *Administrative Science Quarterly*, 43: 668–98.
Tattersall, R. (2009) *Diabetes: The Biography*. New York: OUP.
Tödtling, F., Lehner, P. and Kaufmann, A. (2009) 'Do different types of innovation rely on specific kinds of knowledge interactions?', *Technovation*, 29: 59–71.
Tylor, E. B. (1889) 'On a method of investigating the development of institutions; applied to laws of marriage and descent', *Journal of the Anthropological Institute of Great Britain and Ireland*, 18: 245–72.
Van Raan, A. F. J. (2005) 'For your citations only? Hot topics in bibliometric analysis', *Measurement: Interdisciplinary Research and Perspective*, 3: 50–62.
Vuong, Q. H. (1989) 'Likelihood ratio tests for model selection and non-nested hypotheses', *Econometrica*, 57: 307–33.
Williamson, O. E. (1981) 'The economics of organization: The transaction cost approach', *American Journal of Sociology*, 87: 548–77.
Woolf, S. H. (2008) 'The meaning of translational research and why it matters', *Journal of the American Medical Association*, 299: 211–13.
World Health Organization. (2011) *IC–10: International Statistical Classification of Diseases and Related Health Problems*, 10th revision, Volume 2, Instruction Manual, 2010 edition <http://www.who.int/classifications/icd/ICD10Volume2_en_2010.pdf?ua=1> accessed 4 February 2011.
Wuchty, S., Jones, B. F. and Uzzi, B. (2007) 'The increasing dominance of teams in production of knowledge', *Science*, 316: 1036–9.
Zimmet, P., Alberti, K. G. M. M. and Shaw, J. (2001) 'Global and societal implications of the diabetes epidemic', *Nature*, 414(6865): 782–7.

Appendix: Search query to extract type 2 diabetes publication records

A1. The source of our bibliometric data is the offline version of Elsevier's Scopus which we acquired in June 2009.

A2. In order to retrieve records representing evidence from research on type 2 diabetes we searched for records mentioning in one way or another the following terms in their abstract, title or (indexed or author) keywords: 'non insulin dependent diabetes', 'adult onset diabetes', 'mason type diabetes', 'maturity onset diabetes', 'insulin independent diabetes', 'non ketotic diabetes', 'stable diabetes', 'type 2 diabetes', 'type ii diabetes', 'ketosis resistant diabetes', 'slow onset diabetes', 'mody', 'lipoatrophic diabetes', 'insulin independent diabetes', 'dm2', and 'niddm'.

A3. Note that some terms denoting type 2 diabetes research are rather general, that is, some terms are suspect of having meanings not refering to type 2 diabetes in specific (e.g. 'dm2'). Hence, we first performed a more general search for diabetes research using 'diabetes' and 'diabetic' as search terms only.

A4. More formally then we used the following search query:
{{{diabetes OR diabetic} AND {{adult onset} OR {adultonset} OR {adult-onset} OR {auto somal dominant} OR {autosomal dominant} OR {auto-somal dominant} OR {autosomaldominant} OR {autosomal-dominant} OR {auto-somal-dominant} OR {insulin independent} OR {insulinindependent} OR {insulin-independent} OR {ketosis resistant} OR {ketosisresistant} OR {ketosis-resistant} OR {late onset} OR {lateonset} OR {late-onset} OR {mason type} OR {masontype} OR {mason-type} OR {maturity onset} OR {maturityonset} OR {maturity-onset} OR {non insulin dependent} OR {non insulindependent} OR {non insulin-dependent} OR {non ketotic} OR {noninsulin dependent} OR {non-insulin dependent} OR {noninsulindependent} OR {non-insulindependent} OR {noninsulin-dependent} OR {non-insulin-dependent} OR {nonketotic} OR {non-ketotic} OR {slow onset} OR {slowonset} OR {slow-onset} OR {type 02} OR {type 2} OR {type ii} OR {type-02} OR {type-2} OR {type-ii} OR {aodm } OR {dm 2 } OR {dm2 } OR {dm-2 } OR {mod } OR {mody } OR {ncdmm } OR {niddm } OR {niddy } OR {aodm,} OR {dm 2,} OR {dm2,} OR {dm-2,} OR {mod,} OR {mody,} OR {ncdmm,} OR {niddm,} OR {niddy,} OR {aodm:} OR {dm 2:} OR {dm2:} OR {dm-2:} OR {mod:} OR {mody:} OR {ncdmm:} OR {niddm:} OR {niddy:} OR {aodm;} OR {dm 2;} OR {dm2;} OR {dm-2;} OR {mod;} OR {mody;} OR {ncdmm;} OR {niddm;} OR {niddy;}} OR {{stable diabetes} OR {stable diabetic} OR {diabetes, stable} OR {diabetic, stable} OR {stable-diebetes} OR {stable-diabetic}} OR {{diabetes in young} OR {diabetes in youth} OR {diabetes mellitus in young} OR {diabetes mellitus in youth} OR {diabetes mellitus of the young} OR {diabetes mellitus-in-young} OR {diabetes mellitus-in-youth} OR {diabetes mellitus-of-the-young} OR {diabetes of the young} OR {diabetes-in-young} OR {diabetes-in-youth} OR {diabetes-mellitus in young} OR {diabetes-mellitus in youth} OR {diabetes-mellitus of the young} OR {diabetes-mellitus-in-young} OR {diabetes-

mellitus-in-youth} OR {diabetes-mellitus-of-the-young} OR {diabetes-of-the-young} OR {diabetic in young} OR {diabetic in youth} OR {diabetic of the young} OR {diabetic-in-young} OR {diabetic-in-youth} OR {diabetic-of-the-young} OR {diabetics in young} OR {diabetics in youth} OR {diabetics of the young} OR {diabetics-in-young} OR {diabetics-in-youth} OR {diabetics-of-the-young}} AND {{maturity onset} OR {maturityonset} OR {maturity-onset} OR {non insulin dependent} OR {non insulindependent} OR {non insulin-dependent} OR {noninsulin dependent} OR {non-insulin dependent} OR {noninsulindependent} OR {non-insulindependent} OR {noninsulin-dependent} OR {non-insulin-dependent}}}