Qual Quant (2011) 45:989–1015
DOI 10.1007/s11135-011-9484-3

# On the dynamics of national scientific systems

**Luka Kronegger · Anuška Ferligoj · Patrick Doreian**

Published online: 26 March 2011


**Abstract** Coauthorship links actors at the micro-level of scientists. Through electronic databases we now have enough information to compare entire research disciplines over time. We compare the complete longitudinal coauthorship networks for four research disciplines (biotechnology, mathematics, physics and sociology) for 1986–2005. We examined complete bibliographies of all researchers registered at the national Slovene Research Agency. Known hypotheses were confirmed as were three new hypotheses. There were different coauthoring cultures. However, these cultures changed over time in Slovenia. The number of coauthored publications grew much faster than solo authored productions, especially after independence in 1991 and the integration of Slovenian science into broader EU systems. Trajectories of types of coauthorship differed across the disciplines. Using blockmodeling, we show how coauthorship structures change in all disciplines. The most frequent form was a core-periphery structure with multiple simple cores, a periphery and a semi-periphery. The next most frequent form had this structure but with bridging cores. Bridging cores consolidate the center of a discipline by giving it greater coherence. These consolidated structures appeared at different times in different disciplines, appearing earliest in physics and latest in biotechnology. In 2005, biotechnology had the most consolidated center followed by physics and sociology. All coauthorship networks expanded over time. By far, new recruits went into either the semi-periphery or the periphery in all fields. Two 'lab' fields, biotechnology and physics, have larger semi-peripheries than peripheries. The reverse holds for mathematics and sociology, two 'office' disciplines. Institutional affiliations and shared interests all impact the structure of collaboration in subtle ways.

L. Kronegger (✉) · A. Ferligoj · P. Doreian
Faculty of Social Sciences, University of Ljubljana, Kardeljeva ploščad 5, 1000 Ljubljana, Slovenia
e-mail: luka.kronegger@fdv.uni-lj.si

A. Ferligoj
e-mail: anuska.ferligoj@fdv.uni-lj.si

P. Doreian
Department of Sociology, University of Pittsburgh, 2602 WWPH, Pittsburgh, PA 15260, USA
e-mail: pitpat@pitt.edu

**Keywords** Scientific collaboration · Coauthorship network · Bibliometry · Longitudinal network analysis · Blockmodeling · Core-periphery structure · Mathematics · Physics · Biotechnology · Sociology

## 1 Introduction

Garfield (1955) and Price (1963, 1965) are widely credited with founding contemporary bibliometric studies of science. Following their pioneering efforts, much of the early work in the bibliometry was based on relatively small manually gathered citation and coauthorship networks. The field had new impetus in 1990s with the development of the electronic bibliographic databases and establishment of Garfield's Web of Science (see for example, Garfield 1979). Since then, the number of studies analyzing bibliographic networks on a large scale increased dramatically and the field split into analysis of the directed citation networks and the symmetric coauthorship networks."The number of studies on scientific collaboration has been increasing in the last few decades since researchers, as well as science policy decision makers, have begun to recognize the applicability of these approaches for analytical monitoring of the developments of science (Yasuhiro and Yoshiko 2006)." Coauthorship networks are approached in different ways: some focus on aggregated levels of cooperation among institutions (Corley et al. 2006; Kretschmer et al. 2006) and countries (Gómez et al. 1999) while others focus on the cooperation of individual researchers (Moody 2004). The latter evolved from the analysis of simple network characteristics (Newman 2004) to network modeling based on several approaches including the analysis of scale free random networks and the power law (Barabási et al. 1999, 2002), preferential attachment (Wagner and Leydesdorff 2005), approaches based on information diffusion (Lambiotte and Panzarasa 2009) and other aspects that contribute to the scientific understanding of collaboration (Kretschmer 1997, 1999; Kundra and Kretschmer 1999). Moody reports that for a coauthorship network of sociologists for 1969–1999, based on data extracted from *Sociological Abstracts*, did not form a scale-free network (Moody 2004).

Here, we study complete longitudinal coauthorship networks of Slovenian scientists for four research disciplines: physics, mathematics, biotechnology, and sociology. The basis for our study is Newman's (2004) analysis and comparison of coauthorship networks constructed from three bibliographic databases in biology, physics, and mathematics. Our examination of patterns in collaboration networks of Slovenian scientists draws on his study and adds two substantial components. We include collaboration networks of sociologists in order to extend comparisons of strictly natural and technical sciences to include a representative of the social sciences. Also, we study changes in the structure of collaboration networks over time.

The networks analysed here are *complete* in two ways. First, in the context of describing networks, we have complete information on *all* ties among *all* members of the networks within Slovenia. Second, our data consists of complete bibliographies of *all* registered researchers in Slovenia who work in four disciplines selected for study with the expectation that they would differ from each other and also cover a wide range of disciplinary patterns.

Our article is structured as follows. We first list hypotheses extracted for the literature. We then present an overview of selected static and longitudinal network characteristics for the four disciplines. Next, we consider four consecutive five-year intervals, for all four selected disciplines, and cluster the coauthorship networks using generalized blockmodeling (Doreian et al. 2005). The main structural feature revealed by these analyses is the extent to which these networks all exhibited a multiple core-periphery structure. We then examine how the blockmodel structure of these sequences of collaborative networks are related to the organizational structure of the research institutions in Slovenia, the special research topics in the scientific fields and other factors that are connected to the scientific collaboration. Finally conclusions are drawn and avenues for further work are sketched.

## 2 Research hypotheses

Based on earlier studies of bibliometric databases and broader knowledge concerning scientific collaboration, a set of research hypothesis were extracted to structure our study of collaboration networks in four scientific disciplines.

**H1** Different research disciplines have different publication cultures (Price 1963; Hicks and Katz 1996).

**H2** Researchers from natural and technical sciences collaborate more than researchers from social sciences (Price 1963).

**H3** The number of coauthored publications is growing faster than number of single authored publications in all scientific disciplines (Price 1963).

**H4** The collaboration culture of the natural sciences has been present for a long time in Slovenia. In contrast, collaboration in the social sciences gained its relevance in the last 10 years, mainly because of the pressure towards the internationalization of the Slovenian science.

**H5** Collaboration structure, regardless the research discipline, sooner or later consolidates into a core-periphery structure (Ferligoj and Kronegger 2009).
Consolidation is fostered by integrating each discipline into international policy and collaborative environments. In Slovenia, the important factors were gaining independence in 1991 and the integration of its national science system into European Union (EU) structures. The time of consolidation also depends of the approach taken to research problems including the distinction between teamwork of "lab" disciplines and individual research of "office" disciplines[1].

**H6** Changes of the network structure can be explained by three types of factors: formal factors, defined by organizational structure of research (Perianes-Rodríguez et al. 2010), content of the research (Kuhn 1962) and informal social organization which stands beside the formal organization of the discipline (Crane 1972).

## 3 Data

The data set used here was obtained from two commonly connected sources in Slovenia: (i) the Current Research Information System (SICRIS) which includes the information on all active researchers registered at the Slovenian Research Agency and (ii) the Co-operative On-Line Bibliographic System & Services (COBISS) which contains a database of *all* publications available in Slovenian libraries. Connecting these systems gives a unique *officially maintained* database of complete personal bibliographies of all researchers registered in Slovenia. SICRIS provides additional information on the education, positions and employment of researchers, information on the research groups and the institutions as well as

[1] The disciplines labeled as 'lab' disciplines have a built-in feature of researchers working together in laboratories which necessitates collaborative activity. In contrast, 'office' disciplines do not have this feature and provide greater freedom for individual researchers to work in their own offices.

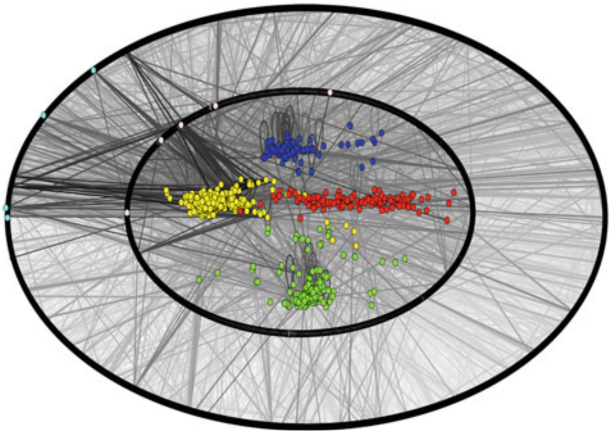


**Fig. 1** Complete coauthorship network of four disciplines. The four groups in the inner part of the network represent physicists on left, mathematicians on the bottom, biotechnologists on the right and sociologists on the top. The inner circle represents researchers who are registered at the Slovenian Research Agency and work in other research fields. The outer circle represents authors who are not researchers registered at the Research Agency

information on both the projects and programs involving Slovenian researchers. Both systems are technically maintained by the Institute of Information Science in Maribor (IZUM).

All researchers registered to work in the fields of physics, mathematics, biotechnology or sociology in Slovenia who were in SICRIS in September 2008 were included in this study. Collaboration between the researchers is operationalized by coauthorship of publications. A symmetric tie between two researchers is measured by coauthorship of relevant scientific contribution[2]. The data set analyzed here is based on all publications issued in the years from 1986 to 2005. The construction of the network differs from similar studies, (e.g. Newman 2004) and (Moody 2004), since the key information for the analysed scientific field about authors and topics is available from the SICRIS database and not in keywords or topic tags of the articles which is usually available bibliographic databases.

The complete coauthorship network of researchers from the four analyzed disciplines with all their coauthors[3] consists of 8,118 units and 84,939 ties. Describing the units (scientists), 250 are physicists, 152 are mathematicians, 105 biotechnologists, and 117 are sociologists. In addition to these 624 units defined for our primary network, there are 1990 authors who are researchers registered at the Slovenian Research Agency and work in other research fields, and 5504 authors who are not registered at the agency and are therefore non-native researchers or, in a small number of cases, unregistered researchers (see Fig. 1). The network has 3899 loops, representing single authored publications and 50246 multiple lines which are summed into the valued connections that present higher number of co-publications.

Essentially, the whole network is one component containing 96% of units. There are 38 nontrivial components of size at least 2 units but they contain only 0.4% (31) units. All but one consist of one or two authors from the primary group of 624 researchers.

[2] Relevant scientific contributions are defined by the Slovenian Research Agency and include original, short or review articles, published scientific conference contributions, monographs or parts of monographs, scientific or documentary films, sounds or video recordings, complete scientific databases, corpus and patents.

[3] For some them we do not have complete information on their personal networks.

**Table 1** Summary statistics of general network properties by disciplines

| | Physics | Math. | Biotech. | Sociology |
|---|---|---|---|---|
| *Primary network* | | | | |
| Number of active authors | 250 | 152 | 105 | 117 |
| Bibl. units per author | 52.5 | 23.9 | 21.4 | 29.9 |
| Coauthors in bibl. unit (single excl.) | 4.6 | 2.8 | 4.6 | 3.7 |
| Single authorships (%) | 5 | 32.9 | 8 | 52.4 |
| Coauthorships within the disc (%) | 72.2 | 29 | 46.5 | 27.6 |
| Coauthorships within the agency (%) | 44.3 | 25.4 | 67.6 | 26.1 |
| Coauthorships outside the agency (%) | 62.7 | 28.2 | 35.4 | 16.4 |
| *Extended network* | | | | |
| Largest component (%) | 97.2 | 80.3 | 97.5 | 90.8 |
| Number of components | 11 | 43 | 7 | 24 |
| Average distance | 4.5 | 5.5 | 4 | 4 |
| Largest distance | 10 | 14 | 8 | 9 |
| Clustering coefficient | 0.246 | 0.493 | 0.419 | 0.490 |
| Density | 0.0019 | 0.0045 | 0.0057 | 0.0098 |
| Average degree | 9.1 | 5.6 | 8.2 | 8.7 |

## 4 General network properties

Our analysis of the coauthorship network properties by research disciplines confirms large differences between the disciplines. The number of physicists is 250, more than twice the number of researchers working in biotechnology or sociology. (See Table 1, line 1.) The differences are even larger in the extended network including secondary units. If we put aside the size of network, physicists, on average collaborate with 18 other authors, mathematicians with 7, biotechnologists with 13 and sociologists with 7. Physicists and biotechnologists seem to have similar pattern of collaboration: they both collaborate with others much more than mathematicians and sociologists whose publication activity reflects a culture of single author publishing. Such patterns can be observed through many network properties; high number of coauthors per bibliographic unit; high numbers of bibliographic units per author for physicists and biotechnologists and low quantities for sociologists and mathematicians. Focusing on this pattern, the first two disciplines could be labeled as “lab” disciplines, and the second two as “office” disciplines. In general, the number of published bibliographic units proportionally follows the sizes of networks with some deviation for biotechnology. Relatively small number of publications per author may be a consequence of the late establishment of the biotechnology as a research discipline in Slovenia[4].

Although the focus of this article is coauthorship, the information on single authored publications in the network is important. Table 1 (lines 5–7) presents the number of publications published with coauthors divided by the number of all publications. In a publication, an author can coauthor with researchers from more than one category of researchers which means the sum of quantities often exceed 100%. Regarding all publications published or copublished by authors from sociology, more than half of publications are single authored and for mathematicians the proportion is above 30%. In contrast, physicists and biotechnologists

[4] Biotechnology developed in the 1970s and is considered as “young discipline”.

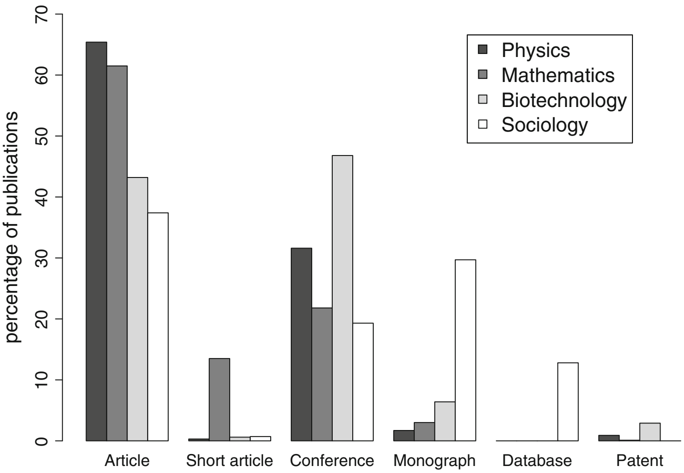



**Fig. 2** Types of publications by disciplines

publish as single authors far less often. (See Table 1, line 4.) Physicists mostly publish in cooperation with colleagues from their discipline and with researchers from abroad, while biotechnologists more often cooperate with Slovenian researchers from other disciplines. Collaboration of mathematicians and sociologists with colleagues within their discipline and researchers from other disciplines follows a similar pattern. The difference between these two disciplines is seen in the mathematicians' higher collaboration with researchers who are not part of the Slovenian science system. (See Table 1.)

In all disciplines, the number of publications per one author seems very high in comparison to the similar analysis of Newman (2004). The main reason for this is the long time period considered here. A secondary reason is that the definition of coauthorship in our study has a broader connotation. Other previously analysed databases do not include information on attendance of conferences, or information on monograph publications. Regarding these differences, we also have to consider differences in types of publications between disciplines shown in Fig. 2. All researchers but biotechnologists primarily publish scientific articles. If we include short scientific articles, the proportion of published articles is the highest among mathematicians followed by physicists, biotechnologists and lastly sociologists. In biotechnology, the most common form of publishing is in published scientific conference proceedings with a contribution or abstract. Sociologists lead in the publication of monographs and with contributions to such publications. They lead also in the production of scientific databases and corpuses. An interesting gap can be observed in the publication of scientific databases or corpuses and patents which indicates clearly different approaches to research productions. Biotechnologists and physicist are patenting more, sociologists more often gather information into databases and corpuses, and mathematicians more often publish their findings in short articles.

The size of networks and coauthorship activity of researchers who work in specific research field, is shown by several network characteristics. The largest component is defined as the largest connected group of individuals in the network. For physicists and biotechnologists, the largest components cover more than 97% of all units in their network. In the network of sociologists this is about 90% and in the network of mathematicians 80% of all authors are in the largest component of their network. Thus, all four disciplines tend to be well connected communities of researchers with collaborations although this is less true for mathematicians. In general, the number of components usually corresponds with the number of nodes in the network. Here, the number of components clearly indicates differences of collaboration structure between "lab" and "office" disciplines and is independent of the size of the network. The distance in the network is defined as length of shortest path between two selected nodes. The distance between two authors who collaborate is 1. Average distances in the four networks considered here generally correspond to those found by Newman (2004). The average distance is the highest among mathematicians, followed by physicists, biotechnologists and sociologists[5].

The global clustering coefficient can be viewed as the average probability of a tie between coauthors of a selected author. For the analysed extended networks, the clustering coefficient is highest for sociologists and mathematicians, slightly lower among biotechnologists and the lowest among physicists. However, interpreting differences in clustering coefficients is complicated by differences between disciplines (Newman 2004) and (later) we examine the actual clustering of these networks via blockmodels. Density is defined as the number of lines in a simple network expressed as proportion of the maximum possible number of links (de Nooy et al. 2005, p. 63). Density should not be confused with the previously mentioned structure of collaboration where we presented data on the number of different coauthored articles. Density is based on the number of different coauthors of selected author. The network of sociologists is the densest followed by networks of biotechnologists, mathematicians and physicists in that order.

Another characteristic is the degree of a vertex in the network. This is the number of lines incident with a vertex (de Nooy et al. 2005, p. 63). Here, degree is the number of different authors with whom a selected author has published. Again the physicists in average collaborate with the highest number of different coauthors. They are followed by sociologists who, on average, collaborate with 8.7 others (when they collaborate) biotechnologists with 8.2 and mathematicians with 5.6 different coauthors. Again, we remind readers that 20 years of publication activity is examined for each discipline.

## 5 Collaboration through time

Here, we present information on coauthorship networks over time for the 20 year period from 1986 to 2005. For some purposes, this is a long time period for analysis of a social system responding only to endogenous changes. However, when the environment within which it operates changes, these external changes merit attention. For the Slovenian science system there were crucial changes that must be considered. First, Slovenia attained independence in 1991. Second, the new country started to be integrated into the scientific environment of the EU. This meant that Slovenia, given the EU formal organization, started to adopt and implement its own science policies. These developments also occurred within broader changes occurring at the global level. For example, there was a dramatic growth of international collaboration (Wagner and Leydesdorff 2005; Gómez et al. 1999). The important change in collaboration culture in Slovenia is clearly indicated by break of the trend in time series of absolute number of published single authors and coauthored publications. As shown in Fig. 3 the change started in early nineties, coincidentally with independence of the country in 1991 and the development of the Internet and electronic communications in 1990s (Laudel 2001).

[5] Some care is needed in interpreting these differences because the networks considered here differ in size.

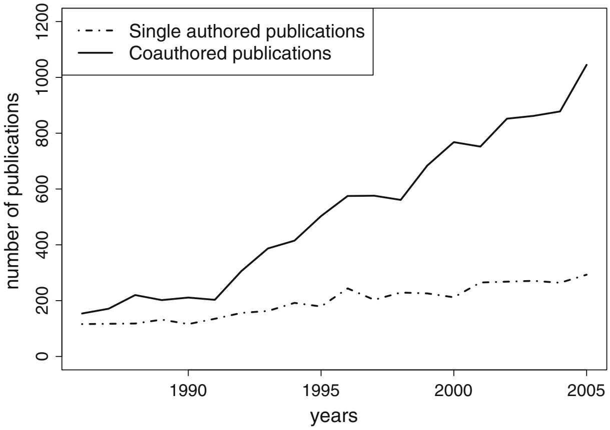



**Fig. 3** Single authored versus coauthored publications in time

Both seem consequential changes in the broader international environment with which any national scientific community must operate.

The time series of collaboration structures presented in Fig. 4 clearly reveal the changes that occurred within the four scientific communities studied here. The values shown in these time series are numbers of articles published as single author or in coauthorship with authors from selected category divided with all articles published in that year. We note that a 'single article' can be coauthored with authors from different categories of authors (from the same discipline, from other disciplines in Slovenia or from authors from abroad). Some of the general characteristics of the four community networks were described in the previous section. (See Table 1.) The most important of these are: (i) the high proportion of single authored publications within sociology and within mathematics; (ii) the high level of collaboration by physicists within their discipline and with authors from abroad and (iii) the collaboration of biotechnologists with researchers from other disciplines within the country. The major difference between mathematicians and sociologists is that mathematicians write more articles with researchers from other countries. Looking closely at the changes in publication behavior over time, there was a dramatic drop in the proportion of single authored publications among mathematicians and sociologists throughout the 20 year period. The proportion of single authored papers by mathematicians gradually decreased for 10 years, while among sociologists it remained high for about 10 years before dropping quite sharply and stabilizing around 2000. For mathematics, the percentage of single authored papers dropped from being slightly above 60% to fall below 30%. In sociology, this percentage rose slightly to above 80% before dropping to slightly below 40%. In sociology single authored publications remain the most frequent form of publishing, while mathematicians started to publish more with others within the discipline and with foreign researchers. For mathematicians, collaboration within the discipline, within the research agency and outside the research agency was stable until just after 1990. Thereafter, all three trajectories climb steadily. Among sociologists, if anything, there was a decline in all three through the mid-1990s before these three trajectories rise. Collaborations within the discipline were about the same as within mathematics at the end of the 20 year period, albeit with a little more fluctuation.

Coauthorship in sociology in Slovenia over time shows a different trend than the one obtained by Moody (2004) who reported steady increase of coauthored articles published in ASR (1936–1999) and Sociological Abstracts (1963–1999).

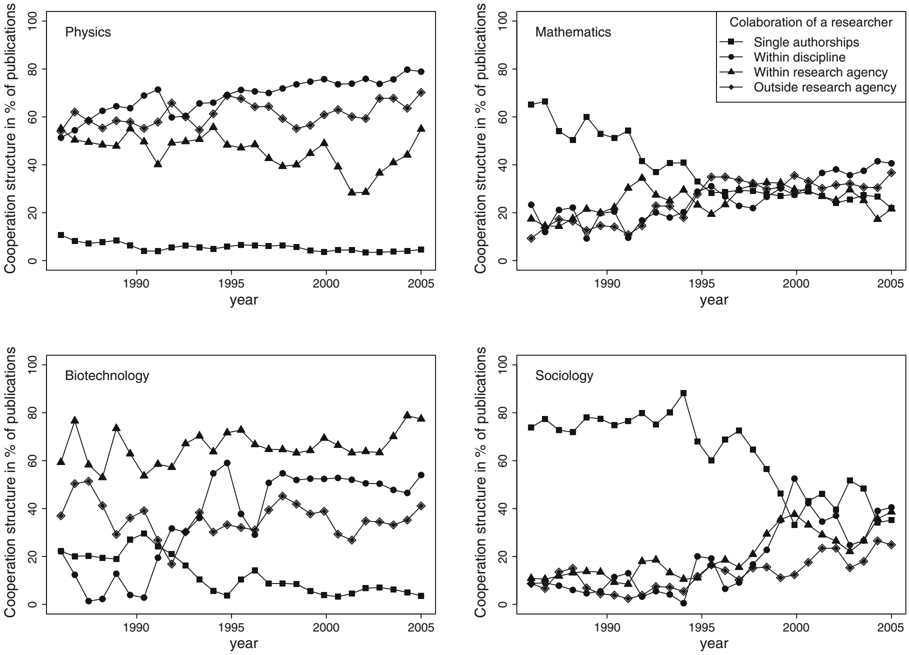



**Fig. 4** Structure of scientific collaboration through time

In physics, the proportion of solo authored productions remained very low through the entire interval with very little change. Within biotechnology, there is an overall decline in solo authored production that ends at about the same level as for physics. This is far lower than the corresponding proportions for mathematics and sociology. The proportion of collaboration within physics rise steadily throughout the 20 years. The levels of collaboration outside the research agency, with more fluctuations, also rises. The proportions for collaboration with researchers from other disciplines who are registered at the Slovenian Research Agency changes with a drop after 2000 before a steady climb through the end of the study period. The trajectories for biotechnology show much more variation than for physics for all forms of collaboration. Collaboration with the researchers registered at the research agency remains high throughout and finishes the highest among the four disciplines. Collaboration within biotechnology is much higher at the end of the 20 year period and the second highest collaborative form for this field at the end of the 20 year interval. It is higher than for both biotechnology and sociology but not physics. This volume of change through time can be viewed as consistent with it being a "young discipline" that is still in the process of defining the form of its collaborative activity.

## Change within four time periods

For a more detailed analysis, we focus on collaborations of scientists, each within their own research discipline, for four 5-year periods. Here we consider only the primary authors, those researchers registered in Slovenia to work in one of the four observed disciplines —physics, mathematics, biotechnology and sociology[6]. We look at the changes in the structure of the coauthorship networks. In order to dampen annual fluctuations in these networks, we divided the twenty-year period into four consecutive five-year intervals. We treat them as snapshots of the form of collaboration networks for the four research disciplines in years 1986 through 1990. The first period (1986–1990) is the period before Slovenia became independent and started to implement its own science policies. The remaining three periods (1991–1995, 1996–2000 and 2001–2005) range from the beginnings of the new implementations to when the Slovene science system was already well integrated into systems of the EU.

We anticipated that the network structures would change over time for each scientific field. We look at some of the overall indicators of network structures before moving on the delineation of these structures in terms of blockmodels. Results for the simple indicators of networks are shown in Table 2. The major items are:

- Every coauthorship network grew in size from one period to the next with increases both in the number of vertices (scientists) and edges (collaborative ties).
- The change in network density differs across the four disciplines:
  - For physics, the density drops steadily across the four periods. In part, this reflects real change and, in part, this is due to this network being the largest for every period.
  - For sociology, the density rises through each period. In large measure, the early increase can be attributed to there being very little collaborative activity early on.
  - The density for biotechnology changes little save for doubling in 1991–1995 before dropping close to the starting level.
  - The density of the mathematics network changes little from the first period to the second before dropping to the same level for the last two periods.
- The average degree rises across all periods for all four disciplines. Not only is the amount of collaboration increasing across the four periods, each author involved in coauthorship activity has, on average, more collaborative partners.
- The average distance also rises across all periods for all four disciplines. This suggests that topics that were once separated, each with its own group of specialists, are becoming more linked over collaborative activities. This is also a necessary consequence of an expanding network.
- The change in the number of components over time differentiates the disciplines:
  - The number of components changes little for physics and sociology across the four periods. However, there are about double the number of components in sociology than in physics suggesting that, in terms of core subject matter, physics has a more coherent subject domain than sociology.
  - The number of components doubles across the four periods for mathematics suggesting a fragmented field.
  - In biotechnology, if anything, the number of components declines modestly suggesting a narrower core focus.
- Change in the largest component, in the sense of the percentage of authors connected in this component, increases across the periods for all four disciplines. Physics starts with the largest component and finishes with largest component. Biotechnology and sociology both start with the smallest largest component and grow at about the same rate to finish

[6] The most important reason for focusing on only this part of networks is that complete networks were constructed only for researchers with complete bibliographies. The second reason is simply computational: direct blockmodeling procedures take huge amounts of computation time.

**Table 2** Network properties through time

| | Network | $t_1$ 1986–1990 | $t_2$ 1991–1995 | $t_3$ 1996–2000 | $t_4$ 2001–2005 |
|---|---|---|---|---|---|
| Physics | Number of vertices | 84 | 125 | 183 | 234 |
| | Number of edges | 173 | 274 | 487 | 686 |
| | Density | 0.050 | 0.035 | 0.029 | 0.025 |
| | Average degree | 4.86 | 5.06 | 6.1 | 6.55 |
| | Largest distance | 9 | 8 | 15 | 16 |
| | Average distance | 2.66 | 3.44 | 4.8 | 5.15 |
| | Number of components | 18 | 20 | 19 | 21 |
| | Largest component (%) | 52.38 | 64.8 | 81.42 | 80.77 |
| | Clustering coefficient | 0.467 | 0.461 | 0.473 | 0.492 |
| | No. of blockmodeling clusters | 8 | 8 | 10 | 11 |
| Mathematics | Number of vertices | 40 | 65 | 96 | 135 |
| | Number of edges | 14 | 42 | 63 | 122 |
| | Density | 0.018 | 0.020 | 0.014 | 0.013 |
| | Average degree | 0.7 | 1.29 | 1.31 | 1.81 |
| | Largest distance | 5 | 5 | 9 | 10 |
| | Average distance | 2.62 | 2.34 | 3.94 | 4.52 |
| | Number of components | 27 | 35 | 49 | 53 |
| | Largest component (%) | 30 | 29.23 | 33.33 | 44.44 |
| | Clustering coefficient | 0 | 0.246 | 0.302 | 0.285 |
| | No. of blockmodeling clusters | 3 | 5 | 7 | 9 |
| Biotechnology | Number of vertices | 16 | 33 | 50 | 79 |
| | Number of edges | 5 | 42 | 48 | 147 |
| | Density | 0.042 | 0.080 | 0.047 | 0.048 |
| | Average degree | 0.63 | 2.55 | 2.32 | 3.72 |
| | Largest distance | 1 | 5 | 6 | 8 |
| | Average distance | 1 | 2.45 | 2.88 | 3.34 |
| | Number of components | 11 | 6 | 9 | 8 |
| | Largest component (%) | 12.5 | 42.42 | 42 | 68.35 |
| | Clustering coefficient | 0 | 0.555 | 0.339 | 0.480 |
| | No. of blockmodeling clusters | 6 | 7 | 7 | 8 |
| Sociology | Number of vertices | 42 | 61 | 88 | 111 |
| | Number of edges | 8 | 26 | 124 | 199 |
| | Density | 0.009 | 0.014 | 0.032 | 0.033 |
| | Average degree | 0.38 | 0.85 | 2.82 | 3.59 |
| | Largest distance | 3 | 4 | 8 | 7 |
| | Average distance | 1.57 | 1.74 | 3.14 | 3.37 |
| | Number of components | 35 | 40 | 41 | 36 |
| | Largest component (%) | 11.9 | 11.48 | 52.27 | 65.77 |
| | Clustering coefficient | 0.429 | 0.500 | 0.589 | 0.539 |
| | No. of blockmodeling clusters | 5 | 6 | 7 | 7 |

next in size to physics. Mathematics has the slowest growth in the size of the largest component.

– Sociology and mathematics, have increasing clustering coefficients in the first three periods which fall slightly in the last one. Biotechnology has the most fluctuating clustering coefficient while physics starts with the highest such coefficient, which slightly declines before rising to its final level in the last period. Sociology finishes with the highest clustering coefficient and mathematics ends with the lowest value of the four disciplines.
– The number of positions (blockmodeling clusters) are at their maximum values in the last time point. The change in the number of positions is the most dramatic for mathematics while this number changes the least for biotechnology.

## 6 Searching for structures with generalized blockmodeling

The basic intuition underlying our analyses is that every scientific discipline has a core-periphery structure (e.g. Ferligoj and Kronegger 2009). Because a 'core-periphery structure' is a much used concept within social science, we specify precisely the meaning of this term as we use it here. For the rest of this paragraph, the term 'scientist' means 'scientist in a specific scientific discipline'. A core is defined as a set of cohesive scientists whose members all collaborate with each other. In blockmodeling terms, a core position has a diagonal block that is complete. There may be multiple cores in the overall structure. If the members of each core coauthor only with other members of their own core, then all cores are of the same type. We view this type as simple cores. However, when there is a core whose members also coauthor in a systematic fashion with members of other cores, we view this a *bridging* core. (There can be *bridging individuals*.) Again, in blockmodeling terms, this is operationalized by having the off-diagonal blocks for a bridging core as complete blocks. The set of cores can be viewed as forming the 'center' of the network. A semi-periphery is occupied by scientists who are involved in at least one coauthoring scientific publication with the others inside the scientific field but do so in a very different fashion compared to members of cores. A small number of scattered coauthorship ties exist within the semi-periphery. This implies that its diagonal block is very sparse and is closer to a null block. Some members of the semi-periphery can also publish with scientists in cores but, again, these coauthorship ties are few and form no systematic pattern. These off-diagonal blocks between the semi-periphery and all cores are also very sparse and close to being null blocks. Finally, the periphery is made up of scientists in who do not collaborate with any other scientists in their field. In blockmodeling terms, the row and column of blocks for the periphery are null. The only permitted block types are null and complete (allowing for near-null blocks) and their locations are known. We clustered the networks assuming the just describe core-periphery structure by fitting pre-specified blockmodels in a deductive fashion within the generalized blockmodeling approach (Doreian et al. 2005).

The clusters of scientists are called also 'positions'. The structure of the pre-specified blockmodel is critical. For the core positions, complete blocks are specified on the diagonal of the blockmodel. The blocks for ties between cores are either null of complete. For the semi-periphery, all blocks are sparse and the blocks for the periphery are exactly null. In delineating the core-periphery as pre-specified blockmodels, we used structural equivalence and examined partitions with between two and twelve positions. To display the results of a blockmodeling analysis, it is standard to present a square array where the rows and columns in the relational matrix have been permuted so that the units of each cluster are placed with each other. Also, each cluster is separated from other clusters and the boundaries are marked with solid lines. By convention, the units of the first cluster (position) is presented first, then the units of the second cluster and so on. The number of clusters for each network (for each discipline in each time interval) was determined visually and by the drop of the criterion function (Doreian et al. 2005). The number of clusters in networks through the four time periods is presented in Table 2 (last line). It varies from 3 to 9 for mathematics, from 5 to 7 for sociology, from 6 to 8 for biotechnology and from 8 to 11 clusters for physics. Generally, the number of clusters rises with the size of network, which is logical consequence of the personal limits of each researcher being able to cooperate with a limited number of coauthors and produce a limited number of publications (Price 1963).

The origins of generalized blockmodeling are found in the study of positions and role systems (Ferligoj et al. 2010). And it is reasonable to think of scientists as playing roles within the structure of their field. So, finding structures in the network helps us to reveal how scientists collaborate within science and help generate its intrinsic structure. We present the blockmodel structures for each discipline in each of the four consecutive five-year periods. With these structures delineated, we go on to examine the mechanisms fostering change in the structure of these networks. We start by considering the structures of the coauthorship networks of Slovenian sociologists.

Sociology

The blockmodel partitions for sociology are shown in Fig. 5 where a black square represents the presence of at least one joint publication for two researchers and a white square represents the absence of any joint publication. Figure 5 presents the blockmodeling structure of the sociological coauthorship network in each of the four time periods. The blockmodel for the first period, $t_1$, has five positions. Three are small simple cores (the first three positions), the fourth position is the semi-periphery (with four researchers) and fifth is a large periphery composed of researchers who did not cooperate within the discipline inside Slovenia. This blockmodel does not quite have the expected core-periphery form. The one departure involves the second position whose member coauthored with all but one member of the periphery. This pattern of ties to actors in another position does have a bridging form. In the second period, $t_2$, the network had exactly the form of a core-periphery structure. There are four simple core positions, each with a complete diagonal block, which are not directly connected to each other. None of these are bridging cores. The next position is an enlarged semi-periphery. Its diagonal block is sparse with a sprinkling of ties and there are just seven coauthorship ties with members of cores. (There is one tie to each of the first three cores and four to the fourth core.) The final position is the periphery. The number of scientists in the periphery is about the same as at $t_1$ but, proportionally, it appears to be a smaller part of the network. The center of the network is slightly larger.

At the third period, $t_3$, the sociological scientific community reveals a dramatic change in the structure of collaboration. In this period, 1995–2000, there were five cores. The first, second, fourth and fifth positions are all cohesive with a small number of ties between them. The third position is a bridging core: its two members collaborate with each other and with all members of the first two cores. The fifth position is the semi-periphery and the periphery is in the last position of this blockmodel. This broad structure remains present in the last time period, $t_4$, with four simple cores and a bridging core as the fourth position. These five cores are the center of the network and, together, they contain only one more sociologist compared to $t_3$. The semi-periphery is the next position and became larger and the periphery in the last position. While the network is much larger at $t_4$, it is clear that the new authors who were drawn into the discipline's publishing activity went primarily into the semi-periphery and periphery.

Physics

The blockmodel structures for physics are shown in the four panels of Fig. 6, with one blockmodel for each of the four periods. At $t_1$ (top left panel), the first five positions are simple cores. Of these, the third, fourth and fifth are maximally dense internally with all of their members collaborating with each other. The fifth core is a bridging core. Both of its members coauthor with members of the first two simple cores. One of them also collaborates with a majority of the physicists in the fifth simple core. There are also a small number of collaborative ties involving members of the second and fifth cores. Otherwise, there are no collaborative ties between members of different cores. This suggests the presence of different research foci

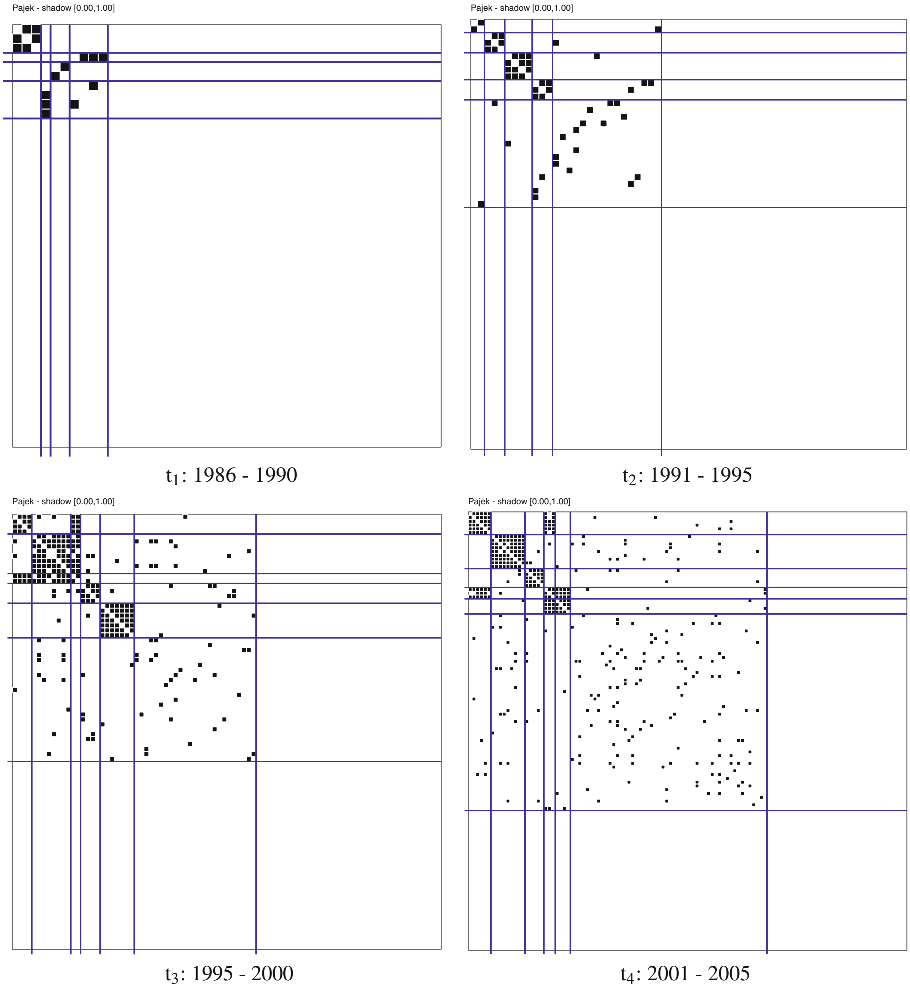
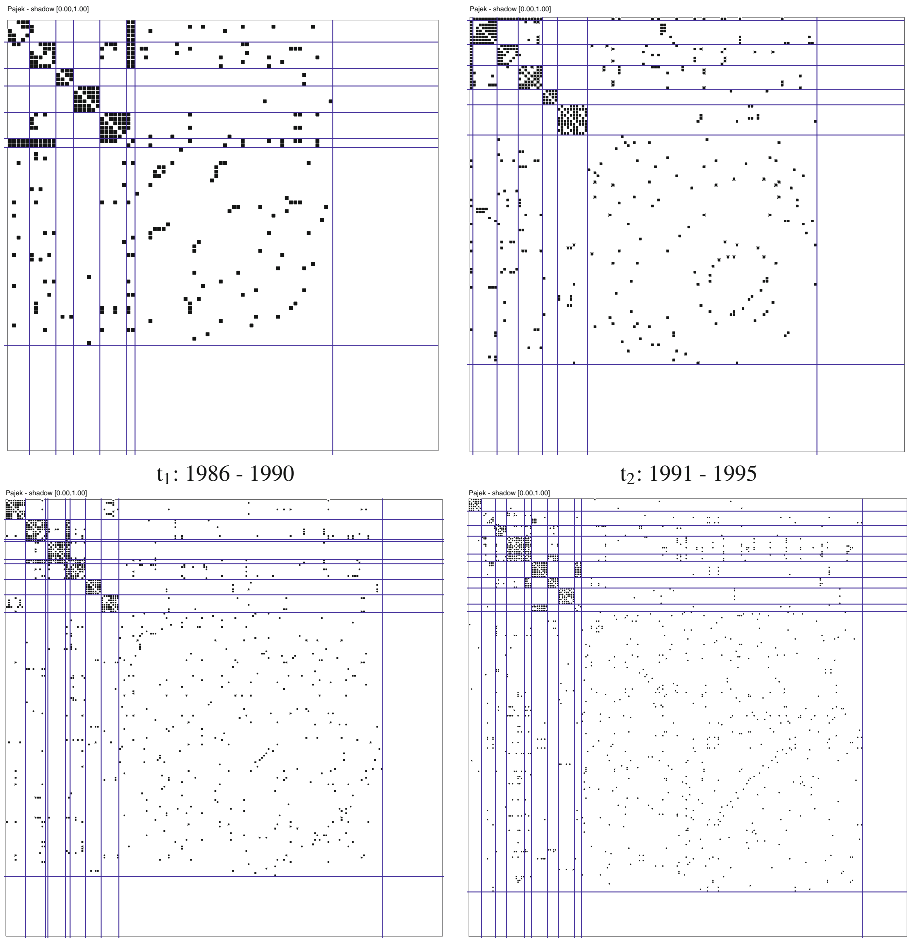



**Fig. 5** Core-periphery blockmodels of sociologists for 4 time periods

for the members of these cores. The seventh position is the semi-periphery. Compared to the sociological coauthorship network, the semi-periphery has many more members and it has a larger presence of coauthorship ties. Regarding the latter, the corresponding diagonal block for the sociologists does not approach this amount of collaborative ties until the final period, $t_4$. The semi-periphery for physics is larger than the periphery (the eighth position of the physics blockmodel), another contrast with sociology.

The blockmodel for $t_2$ also has five simple cores and one bridging core. The bridging core is the first position and is occupied by one physicist who collaborates with all members in first three simple cores. All of the five simple cores are larger than all of the cores for $t_1$. Indeed, the size of the center of this network increased by more than 60%. In contrast the center of the sociological network increased by one sociologist. Consistent with the pattern for physics at $t_1$, there are coauthorship ties between some members of two cores (the first and third). Apart from this small number of ties, the cores have no coauthorship ties suggesting




$t_1$: 1986 - 1990 $t_2$: 1991 - 1995

$t_3$: 1995 - 2000 $t_4$: 2001 - 2005

**Fig. 6** Core-periphery blockmodels of physicists for 4 time periods

the persistence of distinct research areas. Both the semi-periphery (seventh position) and the periphery (eighth position) are larger and the semi-periphery remains larger than the periphery. As was the case for sociology, most of the new entrants into the coauthorship network are located in the semi-periphery and the periphery.

The number of cores increases to eight in the blockmodel for $t_3$. However six of them are simple cores one bridging core, in a clear form, and additional bridging core connected to one of the researchers in the fifth core. This third position is held by researcher who could be part of the second core, but was because of lack of any other connections clustered into separate one. The fifth (bridging) core has authorships with four out of 7 other cores. A clear collaboration structure between clusters of the network can also be observed with collaboration between members of cores 2–6, and members of cores 1 and 8. Members of the seventh core do no collaborate with any author from other cores accept with a few authors from semi-periphery. The semi-periphery, the ninth position in the blockmodel, has expanded dramatically and is much larger than the periphery in the eighth position. Again, bulk of the expansion takes the form of recruiting new physicists into the semi-periphery.

At $t_4$ we focus our attention initially on the first nine positions. All but the first and the ninth are clear cores having dense diagonal blocks. The first and eighth positions have the same profile as in earlier periods being cores without collaborative ties to other cores. The fourth and fifth positions are especially interesting in the sense that they can be described in two ways. One is simply to claim that, in essence, they could be merged and viewed as a single large core, the largest identified thus far. Alternatively, the two can be separated by the presence of many collaborative ties of one (the fifth core) to the seventh core that have nothing in common with the members of the fourth core. We prefer the second impression and we view the fifth core as a bridging core. Not only does it have a large presence of coauthorship with members of the fourth core, there are many coauthorship ties with members of the seventh core. The ninth position, with few internal coauthorship ties, has many collaborative ties with members of the sixth position (another core). In terms of descriptors, this position is best described as being almost a bridging core. We note also that at $t_4$ there are additional coauthorship ties between the cores beyond those described thus far. The center of the network expanded by 50% compared to $t_3$. However, this expansion is much smaller than for the semi-periphery (in the tenth position) which became huge. The periphery (the eleventh position) is about the same size but it has shrunk proportionately.

Mathematics

The blockmodels for the mathematicians are different from those of the sociologists and physicists by seeming simpler. They are shown in Fig. 7. At $t_1$, only three positions are delineated. The first position can be viewed as a simple core lacking just two coauthorship ties to make it a complete diagonal block. The second position is the semi-periphery where there are some coauthorships having no pattern. The third position is a large periphery that is most of the coauthorship network. The overall impression is that mathematicians did not coauthor with other mathematicians in this period. In terms of network structure, fragmentation reigned at $t_1$.

The structure at $t_2$ is different.The first four delineated positions are simple cores. We note that these cores are very small, much smaller than the cores of the sociological and physicist networks. The first core is a 'pure' simple core having no ties to mathematicians in other cores (nor to mathematicians in the semi-periphery). The fifth position has one mathematician with coauthorship ties with members of three cores. This is similar to the single physicist who bridged cores systematically at $t_2$. As such, this position is a bridging core. The semi-periphery (the sixth position) has a small sprinkling of ties in its diagonal blocks and a small number of links to mathematicians in cores but without systematic patterns. This semi-periphery is smaller than the periphery in the last position of the blockmodel. However, both the semi-periphery and periphery have expanded.

At $t_3$, the blockmodel structure reverts to a simpler core-periphery structure. There are five simple cores. The largest has five members and the smallest has two members. There are only two coauthorship ties between the cores. Again, this suggests five distinct research foci. Both the semi-periphery and periphery have expanded with the periphery still the largest position. The overall structure does not change much at $t_4$ although there are now seven identified simple (and small) cores having only six coauthorship ties between them. Again, the semi-periphery and periphery expanded with the periphery remaining the largest position.

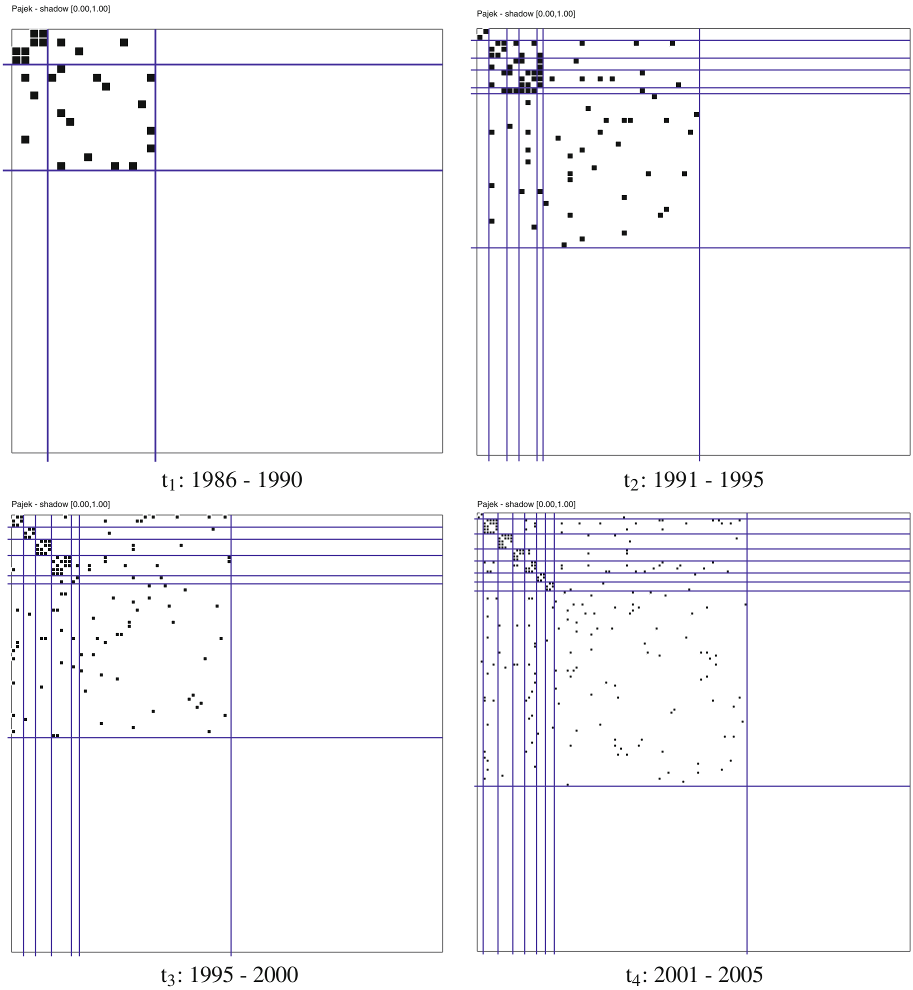



**Fig. 7** Core - periphery blockmodels of mathematicians for 4 time periods

Biotechnology

Figure 8 presents the four blockmodels for biotechnology. For $t_1$, the blockmodel structure is startlingly simple. It has only six pairs of coauthoring biotechnologists but these six pairs of biotechnologists each form a simple core. Each diagonal block is complete and this is the form presented in Fig. 8. There is no semi-periphery and the periphery is larger than the center of the network. At $t_2$ there was a dramatic change to a clear core-periphery structure. There are five simple cores ranging in size from 2 to 6 and only one coauthorship tie exists between one pair of simple cores. There is no bridging core. However, the center comprises about half of the coauthorship network, a distinctive feature for biotechnology compared to the other three disciplines. The semi-periphery occupies the sixth position. The periphery is very small. In essence, the same structure is present at $t_3$. However, the cores have remained small while both the semi-periphery and periphery have expanded. The overall blockmodel

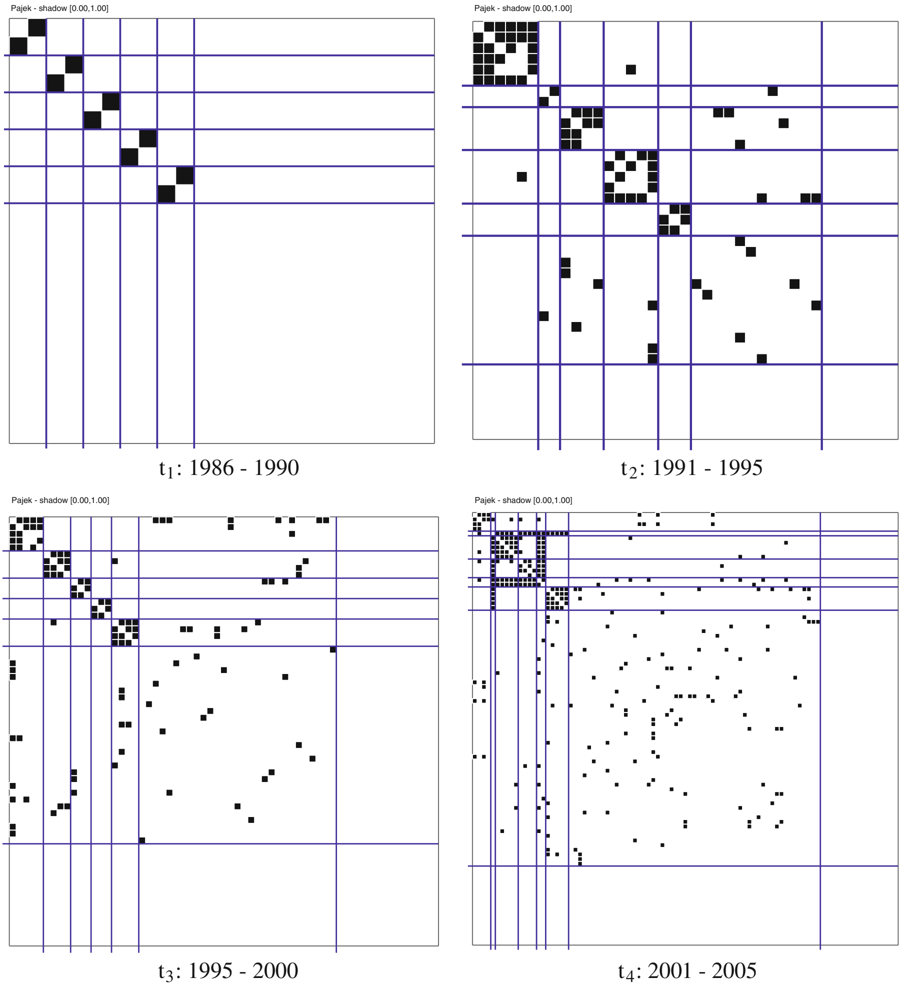



**Fig. 8** Core-periphery blockmodeling of biotechnologists in 4 time intervals

structure for $t_4$ is dramatically different. The central part of the network has six cores. The first, third, fourth and sixth cores are simple cores. *Both* the second and fourth cores are bridging cores. The second position of the blockmodel shown in Fig. 8 is occupied by one biotechnologist. This researcher has one coauthorship tie with a member of the first core and ties to all the members of the remaining cores (except one member of the second core). All members of the fourth core have a coauthorship relation with every member of the second and third cores. The seventh position is a large semi-periphery and the last position is the periphery.

The single biotechnologist in the second core at $t_4$ plays the same role as the single physicist at $t_1$ and $t_2$ in the coauthorship network for physics and the single mathematician at $t_2$ for mathematics. There are two obvious differences for biotechnology compared to the other three disciplines. One difference is that the consolidation, in the form of bridging cores, occurred only in the last period for biotechnology. The second is that biotechnology has two bridging cores. This suggests a greater consolidation among the cores for this field. The central part of the biotechnology network both grew and became more consolidated suggesting that, despite being a 'new' discipline, its core is far more coherent in terms of content than for the other three fields.

### 6.1 Partial summary of the temporal blockmodels

For all disciplines, no single blockmodel structure is present at all periods and there are both similarities and differences between the structures of the disciplines at different points of time. The most frequently present structure is a simple core-periphery structure where there are multiple simple cores, a semi-periphery and a periphery. This simple structure is present: at $t_1$, and $t_2$ for sociology; at $t_3$ and $t_4$ for physics; at $t_1, t_3$ and $t_4$ for mathematics and at $t_1, t_2$ and $t_3$ for biotechnology. In general, within this structural form, the number of cores increases over time for each discipline. The second most frequent blockmodel structure is the core-periphery form with bridging cores. It occurs at $t_3$ and $t_4$ in sociology, at $t_1$ and $t_2$ for physics, at $t_2$ for mathematics and at $t_4$ for biotechnology. Bridging cores tend to be smaller than the simple cores they bridge. Of some interest is that the first appearance of this structure is latest for biotechnology, the youngest of the four disciplines in Slovenia, and it is earliest for physics (at $t_1$) and mathematics (at $t_2$). We use the term 'consolidated center' to describe the idea of having both simple and bridging cores. Such a consolidated structure seems structurally important because it creates a more coherent form for the disciplinary center and facilitates the exchange of ideas across small specialty cores. Such a consolidation need not be stable. While it was present for physics at $t_1$, and $t_2$ it was absent at the last two periods. (However there is a hint of bridging at $t_4$ for physics.) This form was present in mathematics only for $t_2$. It was present for sociology at $t_3$ and $t_4$ and for biotechnology at $t_4$. It will be interesting to see if the simpler structural form, with only simple cores, will appear in these two disciplines after 2005. Given that biotechnology at $t_4$ has the most consolidated disciplinary center (with two bridging cores) this may imply a different structural future for the field compared to sociology. The two 'lab' fields had the periphery smaller than the semi-periphery while the two 'office' fields had peripheries as the largest positions.

## 7 Understanding the changes in disciplinary structures

In the previous section, we presented the blockmodel structures of the coauthorship networks for four disciplines for four consecutive 5-year periods. All four disciplines have structures that, in the main, change from period to period. These structures are not static and, while the structural forms we have detected have interest value of their own, it is necessary to present a less descriptive account of these changes. This amounts to an attempt to present a possible causal account of these changes. We tackle this in three ways: (i) we look closely at the movement of scientists between positions to see if the cores are the same through time or whether scientists move between them; (ii) we examine institutional collaboration as one driver of change and (iii) we look at the content of research domains and publications.

In doing this, we focus on sociology for two primary reasons. First, doing this for all four disciplines would lead to an extremely long document and, second, we are most familiar with this field. Inevitably, disciplinary specific information will be required to account for some of the changes in the forms of collaboration networks. This section, then, is a preliminary demonstration of how the understanding of structural change can be constructed. The detailed examination for physics, mathematics and biotechnology is a core part of our future agenda.

Evolution of blocks

Given the empirical finding that a consolidated core-periphery blockmodel structure appeared first in $t_3$ for sociology, we follow the changes in the memberships of positions from this period to the last period ($t_4$). At both periods there were 7 positions: 5 cores, a semi-periphery and periphery. Of particular interest are the cores of the network. The first panel of Fig. 9 displays the structure at the start point ($t_3$) for this detailed examination. Five cores are presented together with a fragment from part of the semi-periphery. The collaborative ties for members of the five cores are marked by the core's number. The use of the label 6 is for the ties that are present in the semi-periphery. We have not labeled the ties between positions because they are not relevant to our discussion of cores. The second panel of Fig. 9 presents the blockmodel structure at $t_4$ with all of the ties within and between cores. However, rather than label the ties themselves, as we did for $t_3$, we put labels into the diagonal of the panel for $t_4$. The labels 1 through 5 represent the numbers of the cores that were identified for $t_3$. The label 6 denotes sociologists who were present in the semi-periphery at the former time point, the label 7 is used for sociologists who are members of the periphery, and the label 8 for the ones who were new to the coauthorship network at $t_4$. We look at these sociologists first. Examination of the second panel of Fig. 9 shows that just one member of the periphery managed to move into a core. This suggests that moving from the periphery at one point of time to a core at a later point in time is highly unlikely, which means that single authors or authors that collaborate only with others outside the discipline stay in that position. Only two of the sociologists new to the system (denoted by number 8) at $t_4$ move into a core (the third) implying that newcomers to the network do not fare much better than periphery members in moving into a research core.

We turn now to consider the composition of the clusters. First, Cluster 1 from $t_3$ had vanished by $t_4$. The members of this core went either to the semi-periphery, the periphery or left the system entirely. Cluster 2 from $t_3$ largely persisted at $t_4$. It lost one member who moved to the third core at $t_4$ and another who moved into the semi-periphery in the second panel of Fig. 9. Although this core lost two members, it gained three members. Two of the new members came from the $t_3$ Cluster 3 (which is the whole core and was a bridging core for cores 1 and 2) and one from the semi-periphery in $t_3$. Cluster 4 (of $t_3$) had vanished also by $t_4$. The only difference between this vanishing core and $t_3$'s first core is that one member did remain in a core, the third at $t_4$. This core is, in essence, new in terms of its composition. It has one member of the $t_3$ Cluster 2 and the solo survivor of the old Cluster 4. These two sociologists are joined by the only two complete newcomers to the network who reached a core. The last member of this new core came from the old semi-periphery (Cluster 6 at $t_3$). The old Cluster 5, as a core, split into two parts at $t_4$. Three of its members remained in a core (the fourth as shown in Fig. 9) and another three of its members were joined by three other sociologists from semi-periphery to form the first core at $t_4$ (see Fig. 9). The remaining member of the old Cluster 5 moved to the semi-periphery identified at $t_4$. At $t_4$, another new cluster emerged. Two members of this fifth cluster come from semi-periphery, one comes from periphery and one remains from cluster number five. The cluster 4 at $t_4$ is a bridging cluster for clusters 1 and 5.

Our earlier results (in Sect. 6) showed that there were cores present in all periods and that these cores varied in number. Those results implicitly raised the issues of whether extant cores truly persisted and where new cores came from. Given the results in this section, we can now address these issues. Some cores simply disintegrate. In the main, members of these cores move either to the periphery or leave the system. (It is possible that, while they do not collaborate with Slovene sociologists, they may coauthor scientific productions with

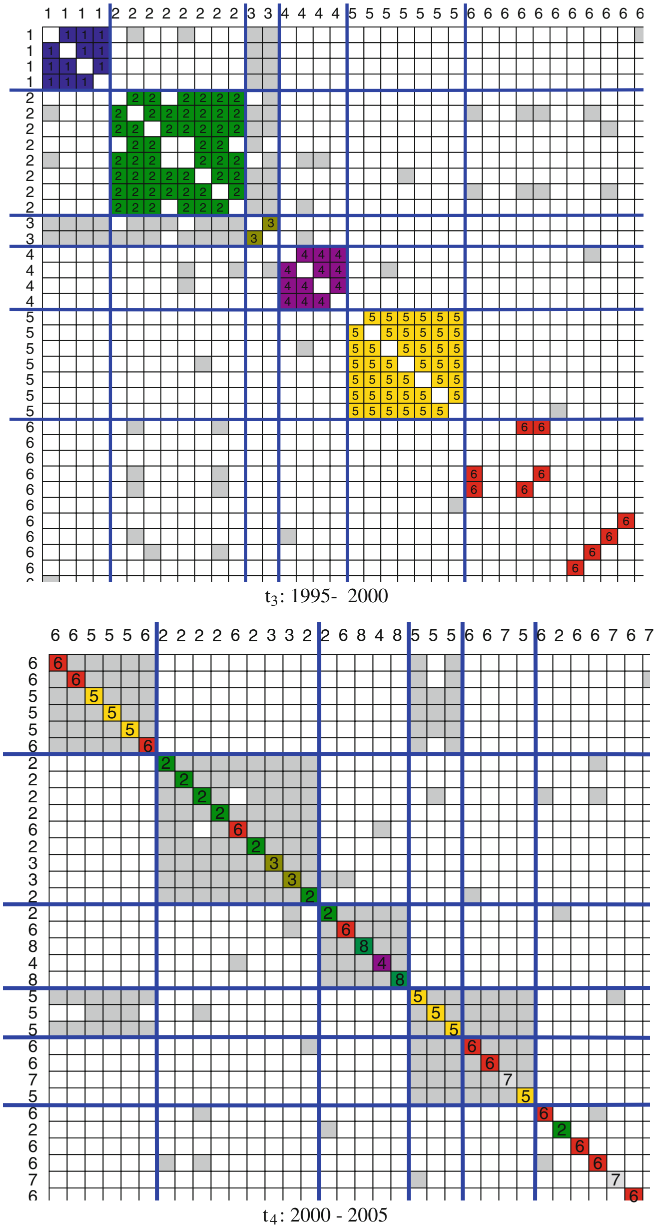



**Fig. 9** Comparison of the core cluster membership in $t_3$ and $t_4$ for the network of sociologists. Different numbers in the squares present cluster membership in the period $t_3$. The numbers 1 to 5 indicate members of 5 core clusters, number 6 indicates authors who are part of semi-periphery, and number 7 authors who belong to periphery cluster. In bottom part of the figure in last period, $t_4$, the number 8 indicates authors who were not yet present in $t_3$

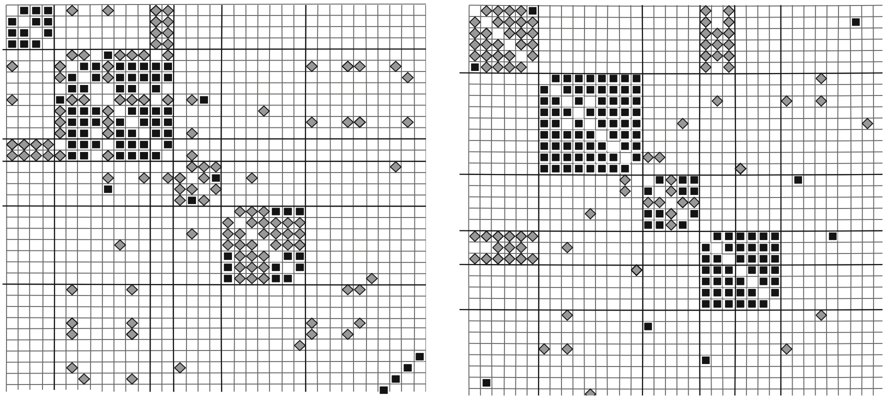



**Fig. 10** Comparing the core clusters by research group equality (sociologists in $t_3$ and $t_4$). *Black squares* on the ties between two researchers denote working in the same research group while the *grey diamonds* represent ties between pairs of sociologists working in different research groups

researchers in other discipline or with sociologists outside Slovenia.) It is possible for a core to persist largely intact, as was the case with Cluster 2 at $t_3$. They may lose a small number of members and recruit a small number of new members. Cores can split and become two new cores. In the Slovenian case, coauthor ties were maintained between the two new cores. New cores can form with a seeding of scientists who have departed from other cores. Based on the data for Slovene sociologists, it is unlikely that a new core will be formed either by newcomers to the systems (who are likely to be younger) or members of the semi-periphery, and this is even less likely from periphery.

## Institutional collaboration

One obvious speculation is that scientists collaborating in producing scientific publications do so at the same research setting. Even though such a speculation has a 'pre-Internet' connotation, it has some intuitive appeal because face-to-face interaction is facilitated by being in the same physical location. The question addressed here is simple: Do the sociologists in core clusters work at the same research unit (e.g., department, research center)? Here, we look closely at two periods, $t_3$ and $t_4$ as we did in Fig. 9. Fig. 10 presents the part of the coauthorship networks with only cores[7] but does so with a slightly different representation of the ties. The coauthorship ties for a pairs of scientists working in the same research unit are denoted by black squares. Coauthorship ties for sociologists working in different organizational units are represented by grey diamonds.

The ties between members of the first core shown in Fig. 10 at $t_3$ (left panel) conforms exactly to the expectation of scientists working at the same institution: they form a complete block with all members coauthoring with each other. Despite the fact that researchers of this cluster come from the same research group this core vanishes in the next time point.

Most of the members of the second core at $t_3$ work in the same research group with the remaining two of them being located in different research settings. The second core at $t_4$ consists of members working in the same research group. Let us remember that the second core at $t_4$ now consists of part of the second cluster and the third cluster from previous time

[7] We have included some of the ties for the next position shown in Fig. 5.

period. The two researchers from the third core at $t_3$ are very interesting: they collaborate with most of the members of the second core who are also from the same research setting and also with all members of the first core without working in the same place. It is an example of consolidation of the research group.

The ties between members of the fourth core form a complete diagonal block but there is only one coauthorship tie that involves researchers from the same place. As mentioned before, the first and the fourth cores at $t_3$ disappear at $t_4$, although they have very different organizational structure: the first consisted of members working in the same research unit and the fourth consisted of members working in different places. In this case, the organizational structure of the research effects the network structure in different ways regarding continuity.

The fifth core at $t_3$, while larger, displays the same pattern with the majority of ties involve coauthors from different research settings. This core subsequently divided into two parts: researchers working in different research units were merged into the first cluster at $t_4$ and the three of them working in the same research unit formed the fourth core in $t_4$.

The detailed examination in this subsection shows that the expectation of sociologists coauthoring because they work in the same place is too simplistic and that collaborations occur across research settings. Of course, this is not an earth shattering result but it forms the foundation for looking more closely at how researchers in different research settings come to coauthor scientific productions.

## Content driven collaboration

One obvious potential mechanism driving research collaboration is that researchers who collaborate wihin a discipline also share the same scientific interests within the broader field (e.g. Moody 2004). However, this does not imply that coauthoring researchers have to share many interests. All that is required is that they share enough interests in order to work together. Even a single shared interest may be enough to support a collaboration. To examine the extent to which the volume of shared interests drives collaboration, we considered items from the set of keywords available in the bibliographic database and titles of publications. Because there are many content words, we first clustered them into 100 clusters. Given these clusters, we operationalized the 'strength' of a tie between two researchers as the percentage of the overlapping clusters of words used in their publications. For this procedure we could only use words in Slovene that were available for 2370 or 82% of all bibliographic units published by sociologists[8]. Words were lemmatized (Erjavec et al. 2005) and clustered, using the k-means method, according to their occurrence in the scientific productions.

Figure 11 contains the same coauthorship data as in Figs. 9 and 10 only the ties are represented in terms of their strength (common interests). Again, we consider only $t_3$ and $t_4$ to show the impact of common interests on coauthorship ties. The darker the representation of the tie, the greater the overlap of sociological interests. In terms of dynamics, the fourth cluster is particularly interesting. This cluster was present at $t_3$ but had dissolved by $t_4$. The shading of ties in the diagonal block form this core at $t_3$ are light representing low levels of common interests. In contrast, the second cluster, present at both $t_3$ and $t_4$, shows much greater overlap of common interest. Moreover, the shading for this cluster at $t_4$ shows a greater overlap in interests at the later time point. The age of core does not automatically imply greater overlaps in the interests of sociologists who collaborate. The newly formed (third) core at $t_4$ came into existence with high levels of common interests. Even though the

[8] In the database most of publications is tagged with keywords in Slovene regardless the language in the text of publication.

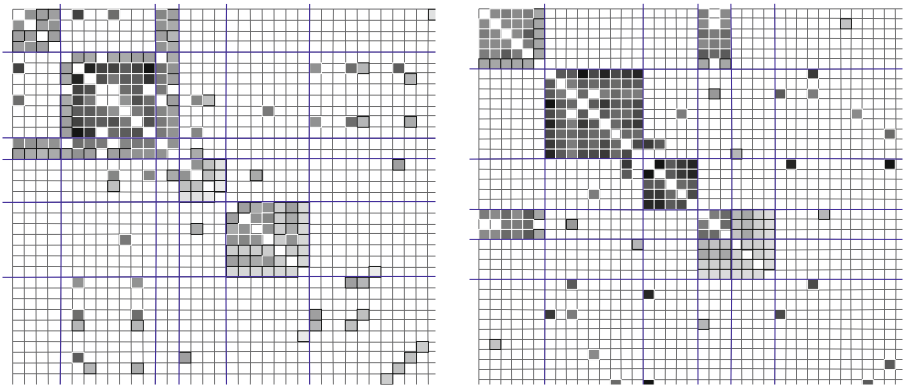



**Fig. 11** Comparing the obtained core clusters for the common content (sociologists in $t_3$ and $t_4$). The strength of the color on a tie presents the percentage of overlapping clusters of words used in authors' publications

fifth core cluster at $t_3$ had split by $t_4$, there was still some cooperation among colleagues in different cores that was driven, in part, by the overlaps of interests.

These results make it clear that there is no simple link between the existence of a scientific core and the overlap in interests of the researchers in a discipline when a temporal perspective is adopted. This makes eminent sense. Some collaborations can be driven for a narrow topic that does not demand that the participants share a lot of interests. At the other extreme, long term members of a broad research tradition are more likely to share more common interests and so work together. Moreover, some research topics become 'hot' and so draw scholars to them. Some of these hot topics are fruitful enough that they persist while others do not live up to their promise and lose adherents. As individuals, some scholars have the same set of interests for long periods of time while others move from topic to topic. Lumping them all of these variations together into a 'single coauthorship network' and focusing only on the whole runs the risk of obscuring factors that drive the change in coauthorship patterns over time.

It seems clear that the levels of interest for the same research topic, as measured by the usage of common words in titles and keywords of the published bibliographic units, do have an impact on the overall structure of scientific collaboration. Our results suggest that low overlap of common interests in a scientific core are consistent with short term persistence while high overlap is strongly linked to coauthorship ties that persist longer in time. Given that the commitment to interests by scientists can vary over time, the overlap of these interests has an impact not only on the duration of particular collaboration but also on the overall structure of coauthorship in science.

## 8 Software

All of the described clustering of attribute and relational data procedures are implemented in Pajek - program for analysis and visualization of large networks (Batagelj and Mrvar 2003). It is freely available, for non-commercial use, at: http://pajek.imfm.si. All calculations besides clustering were made with using R and some visualizations were made with an R package for Generalized and classical blockmodeling of valued networks (Ziberna 2007).

## 9 Conclusion

The unique information database of all publications of the Slovenian researchers provides an extraordinary opportunity to study coauthorship networks of entire disciplines of a country through time.

In the article we tested some known hypotheses (H1, H2, and H3) and some new ones (H4, H5, and H6). We confirmed the hypothesis of different publication cultures between researchers who work in the natural or technical sciences like physics or biotechnology and those who work in the social sciences like sociology. It also became clear that the differences among disciplines do not depend only on the subject of research but also on the nature of the work. In so called "lab" disciplines, collaboration of scientists and publishing coauthoring productions has been present for a long time. In contrast, disciplines where research groups and teamwork are not so crucial for scientific activity have less coauthorship activity. Yet while, overall, there have been changes from publishing as single authors to publishing in cooperation with other scientists in recent years, not all fields were effected in the same way. For the period we considered, the level of single authored publications in physics remained very low while there were changes in different form of who the collaboration partners were. In biotechnology, the initial level of solo authorship was and declined to levels comparable to physics. The remaining disciplines, mathematics and sociology, had very high levels of solo authorship initially and these levels declined dramatically for both fields. However, these declines in the single authored publishing culture occurred in different ways. For mathematics, the steady decrease started in 1980 and stabilized around 1995. The decrease of single authorships among sociologists remained high through 1995 and then dropped dramatically.

Given the presence of coauthorship, the structural forms of these collaborations for each field merited attention. We applied generalized blockmodeling on network slices in four five-year consecutive periods for four disciplines. The disciplines had different detailed structures at different times. However, all can could be characterized by a clear core-periphery structure with small multiple cores, comprised of scientists coauthoring with all, or most, colleagues in their core, a large semi-periphery made up of authors who coauthors a little but have no systematic patterns or presence of collaboration and periphery of authors who do not coauthor with scientists in the same field within Slovenia. In the main, they publish only as single authors but some of these scientists can coauthor productions with researchers from other disciplines or researchers from outside Slovenia. Comparison of blockmodels between disciplines through time showed that clear core-periphery structure is not always present in coauthorship networks at the early time periods.

Of particular interest is a subset of core-periphery structures having bridging cores made up of researchers who collaborate systematically with members of other cores. Separate cores with no or few collaborative ties between them point to fields whose centers focus on different main disciplinary problems. But disciplines having a bridging core have a much more consolidated and coherent center where content areas are linked. However, these consolidated centers are not present at every period and they need not be stable. Such a consolidated center appeared earliest for physics in Slovenia and was present for the first two periods that we studied. It was not present in the third period but there were hints of a reappearance of this form in Physics. For mathematics it appeared once in the second period and then vanished. This consolidated structure appeared in sociology later, in the third period, and then persisted into the fourth. For biotechnology, this form did not appear until that last period and, while this was the latest appearance for a discipline, it occurred in a particularly strong form with two bridging cores. At the end of the period studied, biotechnology and physics had more coherent centers than mathematics.

The coauthorship network expanded for all disciplines throughout the 20 years we examined. In the main, this expansion took the form of recruiting new scientists to the semi-periphery or periphery of their fields. The ‘lab’ disciplines, physics and biotechnology, had peripheries that were much smaller than the semi-peripheries and the reverse was true for the ’office’ disciplines of mathematics and sociology.

In the last part of the article we explored some of the driving forces for the forms taken by scientific coauthorship networks. For this, we focused on the sociological network with respect of possible migration patterns between positions in the overall structure, research group membership and similarity of scientific interests. All three forces do have an impact but they are not simple impacts. Some scientific cores persist, some split into multiple cores and some completely disintegrate. The obvious expectation that cores are centered on specific work environments does not hold. Some do but some do not. But even though scientific cores need not be filled by specialists from the same research locations but they are more likely to persist when they do. The strength of the overlap in scientific interests has some impact on the formation of scientific cores but it does not appear to guarantee their survival.

Here, we studied the three drivers of network structure and the obvious next step is examine them in conjunction. Clearly, there is a need to study the changes in the other three disciplines in the same fashion. We will also consider these and other mechanisms at the level of actors in the SIENA models to test hypotheses statistically. This will require the combination of a micro-level analyses with the scientists and macro-level studies of the overall structure of a discipline. The role of consolidated center for scientific disciplines seems particularly important and with regard to the content and productivity in each discipline.

## References


Barabási, A.L., Albert, R., Jeong, H.: Mean-field theory for scale-free random networks. Physica A. **272**, 173–187 (1999)

Barabási, A.L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., Vicsek, T.: Evolution of the social network of scientific collaborations. Phys. A: Stat. Mech. Appl. **311**(3-4), 590–614 (2002)

Batagelj, V., Mrvar, A.: Analysis and visualization of large networks. In: Jünger, M., Mutzel, P. (eds.) Graph drawing software, pp. 77–103. Springer, Berlin (2003)

Corley, E.A., Boardman, P.C., Bozeman, B.: Design and management of multi-institutional research collaborations: Theoretical implications from two case studies. Res. Policy **35**, 975–993 (2006)

Crane, D.: Invisible colleges: diffusion of knowledge in scientific communities. University of Chicago Press, Chicago (1972)

de Nooy, W., Mrvar, A., Batagelj, V.: Exploratory social network analysis with pajek. Cambridge University Press, Cambridge (2005)

Doreian, P., Batagelj, V., Ferligoj, A.: Generalized Blockmodeling. Cambridge University Press, Cambridge (2005)

Erjavec, T., Ignat, C., Pouliquen, B., Steinberger, R.: Massive multi-lingual corpus compilation: acquis communautaire and totale. In: Proceedings of the 2nd Language and Technology Conference, Poznan, pp. 32–36 (2005)

Ferligoj, A., Kronegger, L.: Clustering of attribute and/or relational data. Metodološki Zvezki. Adv. Methodol. Stat. **6**(2), 135–153 (2009)

Ferligoj, A., Doreian, P., Batagelj, V.: Positions and roles. In: Scott, J., Carrington, P. (eds.) Sage Handbook of Social Network Analysis. Sage Publications, Newbury Park, pp. 434–446 (2010)

Garfield, E.: Citation indexes for science. Science 122, 108–111 (1955)

Garfield, E.: Citation indexing: Its theory and application in science, technology, and humanities. ISI Press, Philadelphia (1979)

Gómez, I., Fernández, M.T., Sebastián, J.: Analysis of the structure of international scientific ooperation networks through bibliometric indicators. Scientometrics **44**(3), 441–457 (1999)

Hicks, D.M., Katz, J.S.: Where is science going?. Sci. Technol. Hum. Values **21**(4), 379–406 (1996)

Kretschmer, H.: Patterns of behaviour in coautorship networks of invisible colleges. Scientometrics **40**(3), 579–591 (1997)

Kretschmer, H.: A new model of scientific collaboration Part 1. Theoretical approach.. Scientometrics **46**(3), 501–518 (1999)

Kretschmer, H., Hoffmann, U., Kretschmer, T.: Collaboration structures between german immunology institutions, and gender visibility, as reflected in the web. Res. Eval. **15**(2), 117–126 (2006)

Kuhn, T.S.: The Structure of Scientific Revolutions. University of Chicago Press, Chicago (1962)

Kundra, R., Kretschmer, H.: A new model of scientific collaboration Part 2. Collaboration of patterns in Indina medicine. Scientometrics **46**(3), 519–528 (1999)

Lambiotte, R, Panzarasa, P: Communities, knowledge creation, and information diffusion. J. Informetr. **3**(3), 180–190 (2009)

Laudel, G: Collaboration, creativity and rewards: why and how scientists collaborate. Int. J. Technol. Manag. **22**, 762–781 (2001)

Moody, J: The structure of a social science collaboration network. Am. Sociol. Rev. **69**, 213–238 (2004)

Newman, M.E.J.: Coauthorship networks and patterns of scientific collaboration. Proc. Natl. Acad. Sci. USA. **101**(Suppl. 1), 5200–5205 (2004)

Perianes-Rodríguez, A., Olmeda-Gómez, C., Moya-Anegón, F.: Detecting, identifying and visualizing research groups in co-authorship networks. Scientometrics **82**(2), 307–319 (2010)

Price, D.S.: Little science, big science and beyond. Columbia University Press, New York (1963)

Price, D.S.: Networks of scientific papers. Science **149**, 510–515 (1965)

Wagner, C., Leydesdorff, L.: Network structure, self-organization, and the growth of international collaboration in science. Res. Policy **34**(10), 1608–1618 (2005)

Yasuhiro, Y., Yoshiko, O.: Patterns of scientific collaboration between Japan and France. Inter-sectoral Analysis Using Probabilistic Partnership Index. Scientometrics **68**, 303–324 (2006)

Ziberna, A.: Generalized blockmodeling of valued networks. Soc. Netw. **29**(1), 105–126 (2007)

Qual Quant (2011) 45:1017–1022
DOI 10.1007/s11135-011-9485-2

# Dynamic building blocks for science: comment on Kronegger, Ferligoj, and Doreian

**Ryan Light · James Moody**

Published online: 29 April 2011


**Abstract** There are two very good reasons to study the social organization of science, and Kronegger, Ferligoj and Doreian's paper exemplify both (henceforth K, F, D). First, we rarely have such rich and detailed data in most other areas of social life. Because science is a written social sphere where credit and authorship matter greatly, we have rich data on scientific careers that allows us to develop new research methodologies. Second, with a nod to the sociology of science, the ever-increasing importance of scientific discovery to national economies, the extension of science through metaphor (or evidence) to other types of organizations with emphasis on the development of new ideas, career trajectories and so forth, underscore the substantive importance of studying science.

**Keywords** Blockmodels · Social network analysis · Sociology of Science

## 1 Substantive contributions and conundrums

A naive view of science suggests clear and distinct communities: since scientists hold PhDs in a single discipline, it is easy to imagine a set of clearly defined tribes working in well-honed problem areas. But, of course, the situation in reality is much less clear, and every analysis of science requires an imposition of boundaries (Gieryn 1999). The seemingly clear boundaries around sociological research, for example, are compromised by interdisciplinary co-authorship or substantive concern trumping disciplinary loyalty resulting in cross-pollinating publications. The case of four scientific disciplines presented by K, F, D provides one of the most comprehensive and well-bounded datasets for the analysis of disciplinary structures to date. As the authors state (pp. 4), the data are complete containing all Slovenian co-authors

R. Light (✉)
Department of Sociology, University of Oregon, Eugene, OR 97403, USA
e-mail: light@uoregon.edu

J. Moody
Department of Sociology, Duke University, Durham, NC 27708, USA
e-mail: jmoody77@soc.duke.edu

within these disciplines expanded by scientists registered by the Slovenian Research Agency. The fact that K, F, D select on officially registered scientists is important not only because it offers a unique picture of Slovenian science contributing to the global picture of science, but also, as the authors note (pp. 4–5), because it avoids the problem of selecting on topic or database that imposes or infers disciplines. In sum, K, F, D provide a unique opportunity to evaluate a complete non-U.S. set of scientists in four disciplines that traverse the laboratory-office divide.

K, F, D use the strength of their data to describe the shifting structure of these four disciplines. Much of the broad disciplinary patterns correspond with previous findings and general expectations. Lab sciences in Slovenia, as one would expect, preceded office sciences in adopting a dominant collaboration practice. Most interesting, perhaps, are the two main structures they discover across disciplines and time. Using block models, the authors discover a dominant core-periphery structure with multiple simple cores (pp. 19). Here, the core consists of co-authors within the same disciplinary core, the semi periphery consists of scientists who rarely co-author inside the disciplinary core, and the periphery consists of scientists who do not co-author within their field (see pp. 12). The existence of this core-periphery structure is not a surprise (see Merton 1968), nor are the differences between the disciplines which by and large conform to the descriptive statistics regarding rates of sole-publication, clearly showing that team science is becoming the norm across disciplines.[1]

But the other common structure they discover is unexpected: a core-periphery with bridging cores. These bridges, the authors argue (pp. 19), lead to a consolidated center as the bridging cores encourage the pollination of ideas across seemingly discrete research centers. Interestingly, these bridging cores do not appear in any consistent pattern across disciplines, but vary over time. Bridges, for example, appear in t3 and t4 for sociology, but in t2 for mathematics, suggesting that bridging and consolidated centers are not the end-result of a more mature science, but instead may respond to endogenous disciplinary pressures.

Like those scholars that try to bridge cores, many scholars attempt to transcend disciplinary boundaries by working with scholars outside of their home discipline or publishing in non-disciplinary journals. Within the Slovenian science data, these scholars would be on the periphery of their discipline. The detailed analysis of Slovenian sociology suggests that this type of boundary-work is excluded from the core discipline as sole-authors and those collaborating solely outside of the discipline remain in the core (pp. 20). This finding is difficult to interpret at the individual-level: Do scholars interested in work outside of the discipline ignore the core or does the core ignore them? However, for the discipline itself, a barrier exists as practitioners on the outskirts remain on the outskirts and serves as a potential indication of a resistance towards the push for interdisciplinary communication and collaboration across academic disciplines (Jacobs and Frickel 2009, pp. 44). Answering these questions is difficult for analyses that restrict vision to within disciplines.

What is lost in analyses that deploy a strict disciplinary boundary if increasing resources and attention are being spent on interdisciplinary research? While its somewhat unfair to simultaneously extol the value of clear boundaries while encouraging the examination of interstitial work, it is important for sociologist of science to think carefully about these less-bounded aspects of science. Future research may benefit from not only exploring such institutionally bounded groups, but those that form in interdisciplinary settings or topics. If, for example, co-authorship differs between economists, sociologists and epidemiologists

[1] While KFD suggest the pattern in Slovenia for Sociology is different from the west as found in Moody (2004), their tables seem to also suggest a strong increase in the rate of collaboration for sociology, if not as steady. While single authorship was common (∼80%) in the early 1990s, it drops significantly to 40–50% in the later periods. Note as well that the trend reported in our earlier work was from the later 1960s forward.

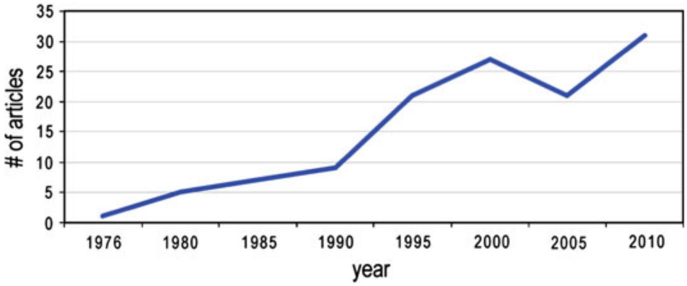



**Fig. 1** Growth of Block models

within HIV/AIDS research, then this may have a profound effect on the speed of the generation and establishment of new ideas and potential solutions to this global health crisis. Also, the analysis of scholars addressing interdisciplinary problems offers a more robust check on the historical effects of previous coauthorship structures as the substantive area is held constant.

## 2 Methodological contributions and the location of disciplinary structures

Perhaps the most unique aspect of this paper is the detailed use of block modeling. The beauty of the generalized block modeling framework is that positions are not limited to be pure cliques or other deeply constrained structures. Instead, similarity across nodes in their pattern of relations can be identified. Thus, unlike standard community detection algorithms (Blondel et al. 2008; Porter et al. 2009) that find only block-diagonal structures (density within groups); the generalized block model can (in principle) identify bridges, isolates, and other substantively interesting patterns. This piece represents a clear example of the value of using this approach.

Research on or using block modeling has grown at a significant rate since the introduction of the idea in the citation classics by Lorrain and White (1971) and White et al. (1976) despite a vast proliferation of competing structure locating procedures. As Fig. 1 depicts, 122 papers indexed in the Web of Science contain some variant of block model in the topic or title. The growth through the 1980s was relatively slow with an increase in the 1990s corresponding, perhaps, with the spread of the personal computer and general social network analysis (see Freeman 2004).

The effect of this growth is not limited to social network analysis or methodological papers, but extends to multiple, diverse substantive areas. As Table 1 indicates, while White et al. (1976) remains the most cited paper on the topic of block modeling with over 450 citations, we can see the top ten citation classics in the field extend to substantive papers on culture (e.g. Anheier et al. 1995), business or trade networks (e.g. Smith and White 1992; Walker et al. 1997; Shan et al. 1994) and, of course, statistical and social scientific methods (e.g. Wasserman and Pattison 1996; Arabie et al. 1978).

On the other hand, the top ten papers cited within this block modeling topic data set consist entirely of methodological papers. Again, White et al. (1976) is the most cited, followed by the classic agenda setting paper by Lorrain and White. While topically these papers are methodological, authorship is diverse from a disciplinary perspective with the presence of sociologists, psychologists, and statisticians. Research on block modeling itself

**Table 1** Top 10 Citations on Blockmodeling

| Authors | Brief citation | Times Cited |
|---|---|---|
| *Top cited articles* | | |
| White, HC, Boorman, SA, Breiger, RL | White HC, 1976, AMER J SOCIOL, V81, P730 | 473 |
| Walker, G, Kogut, B, Shan, WJ | Walker G, 1997, ORGAN SCI, V8, P109 | 259 |
| Shan, WJ, Walker, G, Kogut, B | Shan WJ, 1994, STRATEG MANAGE J, V15, P387 | 200 |
| Wasserman, S, Pattison, P | Wasserman S, 1996, PSYCHOMETRIKA, V61, P401 | 160 |
| Arabie, P, Boorman, SA, Levitt, PR | Arabie P, 1978, J MATH PSYCHOL, V17, P21 | 124 |
| Smith, DA, White, DR | Smith DA, 1992, SOC FORCES, V70, P857 | 74 |
| Anheier, HK, Gerhards, J, Romo, FP | Anheier HK, 1995, AMER J SOCIOL, V100, P859 | 66 |
| Gerlach, ML | Gerlach ML, 1992, ADMIN SCI QUART, V37, P105 | 65 |
| Hoff, PD, Raftery, AE, Handcock, MS | Hoff PD, 2002, J AMER STATIST ASSN, V97, P1090 | 52 |
| Zack, MH, Mckenney, JL | Zack MH, 1995, ORGAN SCI, V6, P394 | 52 |
| *Top articles cited within dataset* | | |
| White, HC, Boorman, SA, Breiger, RL | **White HC, 1976, AM J SOCIOL, V81, P730** | 69 |
| Lorrain, F, White HC | Lorrain F, 1971, J MATH SOCIOL, V1, P49 | 51 |
| Breiger, RL, Boorman, SA, Arabie, P | Breiger RL, 1975, J MATH PSYCHOL, V12, P328 | 46 |
| Arabie, P, Boorman, SA, Levitt, PR | **Arabie P, 1978, J MATH PSYCHOL, V17, P21** | 36 |
| Wasserman, S, Faust, K | Wasserman S, 1994, SOCIAL NETWORK ANAL | 34 |
| Holland, PW, Leinhardt, S | Holland PW, 1981, J AM STAT ASSOC, V76, P33 | 28 |
| Boorman, SA, White, HC | Boorman SA, 1976, AM J SOCIOL, V81, P1384 | 24 |
| White, DR, Reitz, KP | White DR, 1983, SOC NETWORKS, V5, P193 | 20 |
| Wasserman, S, Anderson C | Wasserman S, 1987, SOC NETWORKS, V9, P1 | 20 |
| Holland, PW, Laskey, KB, Leinhardt, S | **Holland PW, 1983, SOC NETWORKS, V5, P109** | 18 |

Bold indicates presence in blockmodel topic data

exemplifies an interdisciplinary project. While still in the nascent stages of its citation life, co-authors Doreian and Ferligoj (along with Vladimir Batagelj), recently published the first comprehensive book on block modeling that has been hailed as a major publication that will necessarily be a touchstone reference for future research (Marsden 2006, pp. 275). This current paper on the network structure of Slovenian science provides an excellent example of the viability of this research program.

The strength of this analysis is complicated only by the awkwardness of dealing with network evolution. As it is difficult (Figure 9 in K, F, D not withstanding), to get a good sense of how position in the network changes over time. Recent work on dynamic networks (Moody 2009; Mucha et al. 2010) has successfully employed a simple trick building on the long-standing tradition of stacking multiple relations (White et al. 1976): simply link each node at time t to itself at t + 1 with identity arcs This then collates the structure, allowing you to see how people change positions over time. We now have solid techniques for clustering such networks to find community structure (Mucha et al. 2010), and extending that work to positional analyses would be a major contribution. As it stands, movement between core, periphery and semi-periphery for these nets is difficult to track.

## 3 Conclusion

The persistence of disciplinary systems seems likely in the coming decades despite institutional encouragement to the contrary, within the United States and elsewhere, from science funding agencies and universities themselves. This push, of course, follows a classic, yet simple finding from social network analysis about how many of the best new ideas develop: Research programs or scientists bridging diverse substantive and methodological areas are often more likely to produce innovation (Burt 2004). Technological change has increased the likelihood that authors will co-author within the international pool of scholars[2]. The authorship of this paper, perhaps, speaks to this interdisciplinarity and internationalism with scholars from the United States and Slovenia, statistics and sociology.

## References

Anheier, H.K., Gerhards, J., Romo, F.P.: Forms of social capital and social structure in cultural fields: Examining Bourdieu's social topography. Am. J. Sociol. **100**(4), 859–903 (1995)

Arabie, P., Boorman, S.A., Scott, A., Levitt, P.R.: Constructing blockmodels: How and why. J. Math. Psy. **17**(1), 21–63 (1978)

Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of community hierarchies in large networks. CORR abs/0803.0476 (2008)

Burt, R.S.: Structural holes and good ideas. Am. J. Soc. **110**(2), 349–399 (2004)

Freeman, L.: The development social network analysis: A study in the sociology of science. Empirical press, vancouver (2004)

Gieryn, T.F.: Cultural Boundaries of Science: Credibility on the Line. University of Chicago Press, Chicago (1999)

Jacobs, J.A., Frickel, S.: Interdisciplinarity: A critical assessment. Ann. Rev. Sociol. **35**, 43–65(2009)

Lorrain, F., White, H.C.: Structural equivalence of individuals in social networks. J. Math. Sociol. **1**(1), 49–80 (1971)

Marsden, P.V.: Book review: generalized blockmodeling. Soc. Network **28**, 275–282 (2006)

Merton, R.K.: The matthew effect in science. Science **159**, 56–63 (1968)

Moody, J.: The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. Am. Soc. Rev. **69**(2), 213–238 (2004)

Moody, J.: Static representations of dynamic networks, Duke University Population Research Working Paper, PWP-DUKE-2009-009 available online: http://papers.ccpr.ucla.edu/download.php?paper=PWP-DUKE-2009-009 (2009)

Mucha, P.J., Richardson, T., Macon, K., Porter, M.A., Onnela, J.P.: Community structure in time-dependent, multiscale, and multiplex networks. Science 328(5980), 876–878 (2010)

Porter, M., Onnela, J., Mucha, P.: Communities in networks. Notices Am. Math. Soc. **56**(9), 1082–1097 (2009)

[2] After all, place is not particularly important to the sociology blockmodels detailed by K,F,D (pg. 23).

Shan, W., Walker, G., Kogut, B.: Interfirm cooperation and startup innovation in the biotechnology industry. Strat. Manag. **15**(5), 387–394 (1994)

Smith, D.A., White, D.R.: Structure and dynamics of the global economy: Network analysis of international trade 1965–1980. Soc. Forces **70**(4), 857–893 (1992)

Walker, G., Kogut, B., Shan, W.: Social capital, structural holes and the formation of an industry network. Organization Sci. **8**(2), 109–125 (1997)

Wasserman, S., Pattison, P.: Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and $p$. Psychometrika **61**(3), 401–425 (1996)

White, H.C., Boorman, S.A., Breiger, R.L.: Social structure from multiple networks. I. Blockmodels of roles and positions. Am. J. Sociol. **81**(4), 730–780 (1976)

Qual Quant (2011) 45:1023–1024
DOI 10.1007/s11135-011-9486-1

# Comment on Kronegger, Ferligoj and Doreian

**Martin Everett**

Published online: 13 May 2011


The paper by Kronegger et al. (2011) DOI 10.1007/s11135-011-9484-3 makes an important contribution to help us understand the dynamics of scientific collaboration in an area of Europe that has in recent history gone through a period of significant change. The existence of a rich dataset which on the face of it seems comprehensive and complete together with detailed attribute data provided the authors with a rich source of information which they have exploited using network analytic techniques. The work is clearly laid out and very well constructed. The six hypotheses they present are clear and precise and they set out about examining these with consummate skill using tools and techniques developed by the authors.

The fact this data extends over a long time period means the dynamical aspects of the changes can be readily seen. Of course it would have been good to have some comparisons from countries that had not been through such radical changes. Independence in 1991 and the move towards European integration in 2004 exactly coincides with the expansion of the internet and the use of electronic communication. It is not clear whether it is the developments in technology or the political changes that account for the increase in collaboration, although clearly a combination of both is the most likely explanation. Data from a country that had more stability over this period would have helped put this in context, but such data may simply not be available. One of the hypotheses has a very strong causal statement, namely H4, which states that increase in collaboration is mainly because of the pressure towards internationalization. This causal aspect, whilst it may have been felt by some of the authors, is not actually directly addressed in the paper and so I do not think that conclusion can be drawn explicitly.

The article makes a number of comparisons between the disciplines but I think more consideration should have been given to the nature of the articles within certain disciplines. For example, there is a very different culture within physics between the theoretical physicists whose work is closer to mathematics and the experimental physicists who typically work on large collaborative experiments, resulting in papers with a large number of co-authors. Equally a mathematician in pure mathematics would typically write a sole authored paper

M. Everett (✉)
School of Social Sciences, University of Manchester, Manchester, M13 9PL, UK
e-mail: martin.everett@manchester.ac.uk

that would be only a few pages long, whereas an applied mathematician may well be involved with a large collaborative team. However, there is of course only so much that can be achieved in one paper and this could well be seen as a distraction from the main findings.

The authors make use of blockmodelling to try and capture certain features of the data. In particular they discuss multiple core-periphery structures and fit the models to this typology. The cores they find seem to be to be relatively small compared to the total number of actors in the network. The peripheries are the isolates and the semi-peripheries seemed to be more defined by the interactions between themselves, rather than the interactions with the core. For a multiple core-periphery structure I would have expected the semi-peripheral actors to interact with the core and not much with each other. If this is the case then it would seem that the data has more of a structure of clustered groups rather than core-periphery and as such a cohesive subgroup analysis could well have been appropriate.

Finally, it appears that the analysis has been done on binary data. However, it would seem that valued data would have provided a far richer source of information. Two aspects come immediately to mind. Firstly, a pair of actors who co-author a number of papers would have a stronger tie that a pair of actors who have only co-authored once. Secondly co-authorship of a paper which only has two authors would be much stronger than co-authorship on a paper with say twenty co-authors. These aspects could only be picked up by using valued data.

In conclusion, I feel that this paper makes an important contribution but that the data itself can be further exploited to give us a deeper understanding of exactly what is going on. But perhaps this is always the case!

Qual Quant (2011) 45:1025–1029
DOI 10.1007/s11135-011-9487-0

# On the dynamics of national scientific systems: a reply

**Patrick Doreian · Anuška Ferligoj · Luka Kronegger**

Published online: 26 March 2011


The comments of Everett (2011) and Light and Moody (2011) confirm our sense that, as a part of studying the dynamics of scientific change, we are tackling an interesting and important set of problems by using a wonderful data set. We appreciate their constructive critiques, especially their prompts to look more closely at the data we have. Their comments also make it clear that we are greatly indebted to the people at Institute of Information Science in Maribor (IZUM) for their maintenance of the Current Research Information System (SICRIS) and Cooperative Online Bibliometric and Services (COBISS) data archives. Without these data sets we could not study total disciplinary networks within the Slovene national science system. We are privileged in having access to these data. We agree with Everett (2011) about the value of comparing this national system with other such systems. However, we do not know of one and hope that our efforts could help promote the idea of creating comparable data sets in other nations. They may, indeed, exist and, if so, we would be more than willing to share ideas with researchers elsewhere regarding national scientific systems.

We agree also with Everett that attention to how data sets are created is merited, especially with regard to the many decisions that must be made in designing them. Some of these decisions are made explicitly while others are made implicitly. We know that some features of the results we report in our paper Kronegger et al. (2011), henceforth KFD, differ somewhat from those of both Newman (2004) and Moody (2004) because of the wider framework used by IZUM for including information. We think that our research program is enhanced by this expansive inclusion of more types of scientific productions. This is a particularly valuable

P. Doreian (✉)
Department of Sociology, University of Pittsburgh, 2602 WWPH, Pittsburgh, PA 15260, USA
e-mail: pitpat@pitt.edu

P. Doreian · A. Ferligoj · L. Kronegger
Faculty of Social Sciences, University of Ljubljana, Kardeljeva ploščad 5, 1000 Ljubljana, Slovenia

A. Ferligoj
e-mail: anuska.ferligoj@fdv.uni-lj.si

L. Kronegger
e-mail: luka.kronegger@fdv.uni-lj.si

feature of the data we use and we plan closer examinations of the different types of scientific productions and their roles in science. Even so, because of the historical importance of articles within science as a form of communication, we agree with Everett that it would be useful to give more attention to the nature of articles that are produced. His examples of differences between the research cultures of theoretical and empirical physicists and between pure and applied mathematical productions are compelling. The scientific productions in our data set are tagged with regard to content and we know the employment locations (universities and research institutes) of the scientists who are involved in coauthoring scientific output. In future work we will pursue a detailed examination of problems addressed in, and the content of, articles and their impact on collaboration. While we accept Everett's assessment that this could be seen as a distraction from the main findings, we think he is pointing us towards an important topic for study.

Everetts examples hint at collaborations between fragments of disciplines and parts of different disciplines. Light and Moody (2011) are more explicit about the problem of disciplinary boundaries for the study of science. Undoubtedly, they are correct in claiming that every analysis of science requires an imposition of boundaries. We note that KFD (Figure 4) provides some information about collaborations outside the research specialties—but this was secondary to the disciplinary results that we reported. We agree with Light and Moody that we benefitted from a design that avoids the problem of selecting a topic or database that imposes or infers disciplines. Light and Moody ask: What is lost in analyses that deploy a strict disciplinary boundary if increasing resources and attention are being spent on interdisciplinary disciplinary research? While it is true that there are increased resources for interdisciplinary work, it seems that disciplinary members still primarily mark their boundaries and turf. So there is much to be gained by studying disciplines. However, Light and Moody question is important and we think that we will be able to provide answers in the future. Given the broader data sets that were drawn upon for KFD, it will be a simple matter (conceptually, if not practically) to couple multiple fields and systematically explore the nature and patterns of coauthorship ties within and between disciplines. This, too, is part of our future agenda.

The notion of a core-periphery structure is used widely in social science studies of social structure. Yet, in many ways, it is used loosely with little attention to specific and multiple forms of, so called, core-periphery structures. The (potential) catalogue of these structures is far from complete. In KFD, we report multiple core-periphery structures and we anticipate that this collection of potential structures will expand with further uses when other disciplines are studied. The idea of a bridging core strikes us a particularly important feature of national science systems, as noted by Light and Moody in reaction to KFD. Everett notes that the (multiple) cores are small. We think this may be a feature of the (relatively) small sizes of scientific systems in Slovenia. Everett is correct to note that one part of some of the Slovene scientific systems (and not just for sociology)—which we called the semi-periphery—is characterized more by ties within it than with ties to any of the cores. This is important for at least three reasons: (i) it is a valid description of the structure of this discipline; (ii) it expands the catalogue of core-periphery types and (iii) it has major implications for blockmodeling. We focus on the last two of these reasons.

An assumption of the study of core-periphery systems seems to be that the real action of a system goes on within the core and through ties between the core and the rest of a social system. Our results suggest that "it ain't necessarily so" and that it is productive for the study of science to explore the contributions of these different parts of scientific systems separately and in conjunction. There is evidence supporting the claim that, for Slovene sociology, there is a greater amount of high quality generalizable research being conducted in the semi-periphery than in the cores. We will be exploring this further in an effort to understand the conditions under which this occurs and the implications that this sort of structural form has for the discipline (or any discipline with this kind of structure.)

In terms of generalized blockmodeling Doreian et al. (2005), and using the direct approach, a blockmodel is specified and fitted by minimizing a criterion function that operationalizes a specific type of equivalence. The criterion function summarizes the total number of inconsistencies between a fitted blockmodel and the closest ideal blockmodel for the type of equivalence that is used. In general, once the blockmodel is fitted, the value of the criterion function plays no further role beyond helping to identify the best blockmodel(s) for the data. Attention is then focused on the delineated blockmodel structure in terms of the block types and their locations. We report this in KFD in our use of multiple cores and the idea of bridging cores. Strictly, the block type for the ties inside the block for the semi-periphery is a null block. In general, while the presence of null blocks in networks has structural significance, there is not a lot that can be written about the form of a null block. However, in the models reported in KFD there are more than a few ties in this null block and, at a minimum, this suggests that we need to pay more attention to the internal structure of this (or any) block when there are a large number of departures from the corresponding ideal block. This can be done by characterizing, for example, the graph theoretical structure of the network inside the semi-periphery positions reported in KFD. This could also be done with a second round of blockmodeling which suggests the value of considering two-step (or multiple-step) blockmodeling both in terms of methods and, perhaps more importantly for studying science, in terms of substance.

Both Everett (2011) and Light and Moody (2011) direct our attention also to the external forces or shocks that can affect any system including scientific systems. Rightly, Everett takes us (mildly) to task on this issue for not disentangling potential external forces and their potential impacts on Slovene scientific systems. We mention Slovenia achieving independence in 1991 and the movement towards integration with wider European institutions along with the upsurge of internet connections internal to nations and between them—but we do not try to separate these forces. The causal standing of our Hypothesis 4 “The collaboration culture of the natural sciences has been present for a long time in Slovenia. In contrast, collaboration in the social sciences gained its relevance in the last ten years, mainly because of the pressure towards the internationalization of the Slovenian science” remains in question given the evidence presented in KFD. This is something we need to address. One step towards doing this is to look more closely, on an annual basis, at collaboration behavior as a complement to summaries for five year periods. We will also be launching a qualitative study of the reasons provided by Slovene scientists for their collaborative behavior via an implementation of a web questionnaire.

All of these responding remarks are little more than a preamble to the deeper topic of network dynamics and network evolution. While the structures revealed by the disciplinary blockmodels are nice characterizations of the structures of four disciplines for four consecutive five-year periods, they are but a set of structural snapshots evenly spaced over time. We did not provide a substantive account for why these structures move between forms described as multiple cores to consolidated cores and, for some disciplines, move back to a simpler form. This could be due to the operation of the external forces that both Everett (2011) and Light and Moody (2011) mention or it could be due to internal disciplinary dynamics. While it is safe to claim that both forces may have been in play, we do need to look at these changes more closely. To that end, we are examining changes in a variety of ways. One was started in KFG by looking a core membership between time points and institute memberships over time. For the former, we documented that some cores at one time point had disintegrated by the next time point (KFD: Figure 9). Members of different cores came together in a new core at the next time point while some members went to the semi-periphery. Even though it is remarkable that some form of a core-periphery structure remained in place despite these dramatic changes perhaps as a self-organized system property we presented no coherent substantive argument for why. We suspect that there were subtle interplays of field based forces, interdisciplinary forces and external forces together with the dynamics of individual career moves occurring while scientists age. These are all topics we intend to pursue given the rich description of content and employment histories in the data. Additionally, we intend to look closely at the implications of the KFD result that the four scientific disciplines expanded primarily by recruiting new members to the semi-periphery and the periphery.

We are at work with a completely different data analytic approach by using SIENA Snijders (2005) to tackle some issues concerning these temporal dynamics. A longer range goal is to couple exponential random graph models (ergms) and blockmodels into a single coherent modeling framework for studying scientific structural change. In this context, the partial history of blockmodeling provided by Light and Moody is instructive. The arrival of blockmodeling (in 1971) was a superb intellectual achievement that was based on the substantive concern of understanding the structure and operation of role systems. Over time, it seems that blockmodeling became known as strictly a methodological approach for discerning the structure of any social network. And while the tools were used as such, it seems that the early promise of a substantive understanding of network structure and dynamics was not realized. This may help account for the Web of Science data reported by Light and Moody (2011) where the most cited items are to the older foundational statements. While a lot of separated blockmodeling studies were conducted, there was little by way of cumulated knowledge of social structure and changes in social structure. This concern motivated the development of Generalized Blockmodeling Doreian et al. (2005) to provide a more general treatment of blockmodeling and to secure its foundations. Included in this document was an expansion of the types of blocks and hence blockmodel structures, as well as the use of pre-specified blockmodels. While it would be nice if this book in its nascent citation life (to use Light and Moodys delightful term) came to have greater influence on current research, the content of the book did have some limitations. In the context of studying scientific systems, two are particularly pertinent.

Everett points one of them out explicitly: While the data on scientific coauthorship are valued data, they were treated as binary data for the blockmodeling. We agree with him that it will be useful, even necessary, to take advantage of the information in the valued network ties. Ziberna (2007) suggests ways in which this can be done and makes it clear that doing so is non-trivial task. Generalized Blockmodeling had a lot to say about structure but included very little regarding the evolution of blockmodel structures nor the evolution of network structure. Light and Moody comment on the awkwardness of dealing with network evolution and added it is difficult (Figure 9 not withstanding) to get a good sense of how positions in the network change over time. Again, we agree. Coupling blockmodeling to ergms, as proposed in Doreian et al. (2009), will take us part of the way towards getting a better sense of the co-evolution of network structure, positions and individual scientific careers.

Light and Moody also point to new techniques that show promise of being valuable in studying change of time of network structures and provide some citations. We will consult those sources because we are convinced that multiple network methods are needed for studying social networks, including national scientific systems. Thus, while KFD used primarily blockmodels for this task, we do not claim that they are the best way of doing so. Certainly, they are not the only tools that will be useful for the task of comprehending structure and structural dynamics. We have a wonderful data set to explore the dynamics of scientific systems and a broadening set of methodological tools. We know also that the only legitimate criterion for success will be the creation of valid substantive understanding that matters. Using only one technique to get there will not be an option.

## References

Doreian, P., Batagelj, V., Ferligoj, A., Generalized Blockmodeling, Cambridge University Press, New York (2005)

Doreian, P., Wang, P., Robins, G., Pattison, P.: Blockmodel traces and exponential graph parameters representing network processes. In: SUNBELT XXIX. International Sunbelt Social Network Conference, San Diego (2009)

Everett, M.: Comment on Kronegger, Ferligoj and Doreian. Qual. Quant. doi:10.1007/s11135-011-9486-1. (this issue) (2011)

Kronegger, L., Ferligoj, A., Doreian, P.: On the dynamics of national scientific systems. Qual. Quant. doi: 10.1007/s11135-011-9484-3. (this issue) (2011)

Light, R., Moody, J.: Dynamic building blocks for science: Comment on Kronegger, Ferligoj and Doreian. Qual. Quant. doi:10.1007/s11135-011-9485-2. (this issue) (2011)

Moody, J.: The structure of a social science collaboration network: disciplinary cohesion from 1963 to 1999. Am. Sociol. Rev. **69**(2), 213–238 (2004)

Newman, M.E.J.: Coauthorship networks and patterns of scientific collaboration. Proc. Natl Acad. Sci. USA 101(Suppl 1), 5200–5205 (2004)

Snijders, T.A.: Models for longitudinal network data. In: Carrington, P., Scott, J.P., Wasserman, S. (eds.) Models and methods in social network analysis, chap. 11, pp. 215–247. Cambridge University Press, New York (2005)

Žiberna, A.: Generalized blockmodeling of valued networks. Soc. Netw. **29**(1), 105–126 (2007)